Predicting and analyzing HIV-1 adaptation to broadly neutralizing antibodies and the host immune system using machine learning

Dissertation

zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Fakultät für Mathematik und Informatik
der Universität des Saarlandes

von

Anna Hake

Saarbrücken, 2022

Tag des Kolloquiums:	20.03.2023
Dekan:	Prof. Dr. Jürgen Steimle
Prüfungsausschuss:	
Vorsitzender:	Prof. Dr. Karl Bringmann
Berichterstatter:	Prof. Dr. Nico Pfeifer Prof. Dr. Olga Kalinina
Akademischer Mitarbeiter:	Dr. Fabian Kern

Acknowledgments

First and foremost, I would like to express my gratitude to my supervisor Nico Pfeifer. He always supported me along my scientific and personal path, sharing his expertise and knowledge with me. Moreover, I couldn't have had a more kind-hearted supervisor. I am deeply thankful to Thomas Lengauer for creating a fantastic research environment in our department and for his long-lasting support even after his retirement. I also want to thank my thesis committee for taking the time to review my thesis and being part of my doctoral examination.

I am thankful to all collaboration partners for our fruitful cooperations and discussions. In particular, I want to thank Rolf Kaiser and Florian Klein for giving me the opportunity to receive clinically relevant feedback on my work. I want to express my gratitude to Hans-Peter Lenhof and Volkhard Helms for teaching me the concepts of bioinformatics and sparking my interest in research. I thank Georg Friedrich, Joachim Büch, and the IST department of the Max Planck Institute for Informatics (MPII) in Saarbrücken (in particular Maik Muschter) for their technical support, and Ruth Schneppen-Christmann for her administrative support. I thank the MPII administration department, especially Alexandra Klasen-Schmitt, Michael Bentz, and Volker Geiss, who always had an open ear for all my concerns and were willing to help. In particular, I want to thank the MPII for their support with regard to childcare opportunities. Knowing your children in good hands enables you to reach your professional goals.

A big thank you goes to all past members of the Computational Biology and Applied Algorithmics Department, who made working such a pleasure. I really enjoyed our lunch and coffee breaks. Thanks to: Alejandro, Basti, Fabian, Matthias, Nadja, Peter, Prabhav, Olga, Tomas, Dilip, Fatemeh, Florian, and Siva. I was very fortunate to had not only competent, but also incredible nice team members: Adrin, Lisa, Matthias, Nora, and Sarvesh. You made the weekly meetings always a pleasure. I also want to thank all the senior researchers of the department who were always happy to share their experience regarding research, career or life topics. Thank you Glenn, Marcel, Markus, Olga, and Tobias.

I am indebted to Peter for introducing me to Nico, which was the start of my statistical learning journey in HIV-1 research. Special thanks go to Nora, Siva, Prabhav, and Peter for taking the time to proofread parts of my thesis! Thank you for staying close friends, even from far away.

Finally, I want to thank my friends and family for their limitless understanding of the challenges on my journey, their constant and unconditional support and love. In particular, I want to thank my parents, my parents-in-law, and my husband for trying to help me in every possible way. I am so grateful for my two independent research projects, my children, who made me grow as a person.

Abstract

Thanks to its extraordinarily high mutation and replication rate, the human immunodeficiency virus type 1 (HIV-1) is able to rapidly adapt to the selection pressure imposed by the host immune system or antiretroviral drug exposure. With neither a cure nor a vaccine at hand, viral control is a major pillar in the combat of the HIV-1 pandemic. Without drug exposure, interindividual differences in viral control are partly influenced by host genetic factors like the human leukocyte antigen (HLA) system, and viral genetic factors like the predominant coreceptor usage of the virus. Thus, a close monitoring of the viral population within the patients and adjustments in the treatment regimens, as well as a continuous development of new drug components are indispensable measures to counteract the emergence of viral escape variants. To this end, a fast and accurate determination of the viral adaptation is essential for a successful treatment.

This thesis is based upon four studies that aim to develop and apply statistical learning methods to (i) predict adaptation of the virus to broadly neutralizing antibodies (bNAbs), a promising new treatment option, (ii) advance antibody-mediated immunotherapy for clinical usage, and (iii) predict viral adaptation to the HLA system to further understand the switch in HIV-1 coreceptor usage.

In total, this thesis comprises several statistical learning approaches to predict HIV-1 adaptation, thereby, enabling a better control of HIV-1 infections.

Kurzfassung

Dank seiner außergewöhnlich hohen Mutations- und Replikationsrate ist das humane Immundefizienzvirus Typ 1 (HIV-1) in der Lage sich schnell an den vom Immunsystem des Wirtes oder durch die antiretrovirale Arzneimittelexposition ausgeübten Selektionsdruck anzupassen. Da weder ein Heilmittel noch ein Impfstoff verfügbar sind, ist die Viruskontrolle eine wichtige Säule im Kampf gegen die HIV-1-Pandemie. Ohne Arzneimittelexposition werden interindividuelle Unterschiede in der Viruskontrolle teilweise durch genetische Faktoren des Wirts wie das humane Leukozytenantigensystem (HLA) und virale genetische Faktoren wie die vorherrschende Korezeptornutzung des Virus beeinflusst. Eine genaue Überwachung der Viruspopulation innerhalb des Patienten, gegebenfalls Anpassungen der Behandlungsschemata sowie eine kontinuierliche Entwicklung neuer Wirkstoffkomponenten sind daher unerlässliche Maßnahmen, um dem Auftreten viraler Fluchtvarianten entgegenzuwirken. Für eine erfolgreiche Behandlung ist eine schnelle und genaue Bestimmung der Anpassung einer Variante essentiell.

Die Thesis basiert auf vier Studien, deren Ziel es ist statistische Lernverfahren zu entwickeln und anzuwenden, um (1) die Anpassung von HIV-1 an breit neutralisierende Antikörper, eine neuartige vielversprechende Therapieoption, vorherzusagen, (2) den Einsatz von Antikörper-basierte Immuntherapien für den klinischen Einsatz voranzutreiben, und (3) die virale Anpassung von HIV-1 an das HLA-System vorherzusagen, um den Wechsel der HIV-1 Korezeptornutzung besser zu verstehen.

Zusammenfassend umfasst diese Thesis mehrere statistische Lernverfahrenansätze, um HIV Anpassung vorherzusagen, wodurch eine bessere Kontrolle von HIV-1 Infektionen ermöglicht wird.

Scientific contribution

This cumulative thesis is based upon three accepted, peer-reviewed publications and a preprint manuscript under peer-reviewed submission. All four manuscripts are included in Appendix A. In the following, a list of all included publications is given with their reference. Afterwards, the personal scientific contribution of the author Anna Hake (AH) (née Anna Feldmann) is listed for each publication.

List of publications

(i) Journal articles (peer-reviewed)

Paper 1 A. **Hake** and N. Pfeifer. „Prediction of HIV-1 sensitivity to broadly neutralizing antibodies shows a trend towards resistance over time.“ In: *PLOS Computational Biology* 13.10 (Oct. 2017), pp. 1–23. DOI: 10.1371/journal.pcbi.1005789. URL: https://doi.org/10.1371/journal.pcbi.1005789

Paper 2 T. Schoofs, F. Klein, M. Braunschweig, E. F. Kreider, A. **Feldmann**, L. Nogueira, T. Oliveira, J. C. C. Lorenzi, E. H. Parrish, G. H. Learn, A. P. West, P. J. Bjorkman, S. J. Schlesinger, M. S. Seaman, J. Czartoski, M. J. McElrath, N. Pfeifer, B. H. Hahn, M. Caskey, and M. C. Nussenzweig. „HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1.“ In: *Science* 352.6288 (2016), pp. 997–1001. DOI: 10.1126/science.aaf0972. eprint: https://www.science.org/doi/pdf/10.1126/science.aaf0972. URL: https://www.science.org/doi/abs/10.1126/science.aaf0972

Paper 3 J. Scheid, J. Horwitz, Y. Bar-On, E. Kreider, C.-L. Lu, J. C. Cetrulo Lorenzi, A. **Feldmann**, M. Braunschweig, L. Nogueira, T. Oliveira, I. Shimeliovich, R. Patel, L. Burke, Y. Cohen, S. Hadrigan, A. Settler, M. Witmer-Pack, J. West, B. Juelg, and M. Caskey. „HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption.“ In: *Nature* 535 (June 2016). DOI: 10.1038/nature18929

(ii) Preprints

Paper 4 A. **Hake**, A. Germann, C. de Beer, A. Thielen, M. Däumer, W. Preiser, H. von Briesen, and N. Pfeifer. „Insights to HIV-1 coreceptor usage by estimating HLA adaptation with Bayesian generalized linear mixed models.“ In: *bioRxiv* (2022). DOI: 10.1101/2022.07.06.498925. eprint: https://www.biorxiv.org/content/early/2022/07/06/2022.07.06.498925.full.pdf. URL: https://www.biorxiv.org/content/early/2022/07/06/2022.07.06.498925

Scientific contribution

Paper 1 - bNAb resistance study

Paper 1 is a first-author publication of **AH** under the supervision of **NP** without further collaboration partners. **AH** was responsible for the data curation, formal analysis, investigation, methodology, validation, visualization, writing of the original draft, and editing and reviewing the final version of the manuscript.

Paper 2 - bNAb efficacy study I

In Paper 2, **AH** was responsible for performing the statistical analyses under the supervision of **NP**. In particular, the scientific idea of **AH** to use a noise-corrected version of the area under the dose-neutralization curve as measure for neutralization enabled the inclusion of all study participants even those not reaching the IC50 values to study the change in neutralization capability. In addition, the responsibility of **AH** was to perform a univariate (Bayesian) linear regression to correct for the influence of confounding variables on the change in neutralization and to compare the different groups with respect to change in neutralization capability. All other statistical analyses in the paper were not performed by **AH**.

Paper 3 - bNAb efficacy study II

AH performed the statistical analyses under the supervision of **NP** in Paper 3. In detail, the modeling of the rebound time as well as performing the survival analysis to compare the different treatment groups was the responsibility of **AH**, including the correction for confounding variables. In addition, the performance of hypothesis tests was done by **AH**. While **AH** contributed to the choice of the weighted log-rank test, the test was implemented by **NP**.

Paper 4 - HIV immunoadaptation study

Paper 4 is a first-author publication of **AH** under the supervision of **NP** with further collaboration partners. The collaboration partners were mainly responsible for data collection, data access, sequencing and genotyping the data, and writing the corresponding Material and Methods section in the manuscript. **AH** was responsible for the data curation and data preprocessing, formal analysis, investigation, methodology, validation, visualization, writing of the original draft, and editing and reviewing the final version of the manuscript.

Contents

List of Tables

1. Introduction

> HIV has indeed escaped from
> Pandora's microbial box.
>
> *Simon Wain-Hobson, 1993*

HIV-1 is an escape artist beyond all comparison. With $4.1\text{x}10^{-3}$ mutations per base per cell, HIV-1 has the highest encountered mutation rate in vivo in a biological entity [5]. Together with a high turnover rate [6] and the ability to recombine genetic material [7], these viral characteristics are responsible for the rapid evolution of HIV-1 and the cause for the extraordinary viral genetic diversity within and between HIV-1 patients. This extraordinary genetic diversity is the major reason why the virus is able to escape and to adapt to the selection pressure imposed by antiretroviral drug exposure or by the host immune system such that natural clearance is very rare [8, 9]. The viral variants also differ globally, clustering into different geographical subtypes, i.e., subtype B is more prevalent in Europe and North America, while subtype C HIV-1 is most prevalent in Southern Africa and India, accounting for approximately 50% of the infections [10]. The global viral genetic diversity is also the reason why there is still no vaccine available despite four decades of research.

Infection with HIV-1 still a global burden By infecting the human $CD4^+$ cells - specific immune cells carrying the cluster of differentiation 4 (CD4) molecule on the surface, HIV-1 weakens the immune system in two ways. Infected $CD4^+$ cells can be directly destroyed by the virus after several rounds of viral replication but also via a successful elimination from the immune system itself. Eventually, an untreated infection leads to the acquired immunodeficiency syndrome (AIDS), resulting in the death of the majority of the patients by opportunistic infections [11]. In 2020, 37.7 million people have been living with HIV-1 and 680 000 people have died from AIDS related illnesses, making infection with HIV-1 still a serious worldwide health issue according to UNAIDS [12]. The incidence rate of newly infected HIV-1 patients has been reduced by 31% from 2.1 million in 2010 to 1.5 million in 2020.

Tracking HIV-1 adaptation essential for viral control With neither a cure nor a vaccine within reach, viral control is one of the major pillars to end the HIV pandemic by 2030 [13]. Viral control can, in general, be established naturally by our human immune system or with the help of antiretroviral drugs. The rapid viral evolution allows HIV-1, however, to adapt and to evade the selection pressure imposed by the host immune system and by antiretroviral drug exposure. Thus, viral adaptation interferes with our attempts to control the virus. Here, adaptation is defined as the accumulation of changes in the viral genome - viral mutations - that result in a fitness advantage of the virus with respect to a specific environment. Thus, understanding and determining viral adaptation towards the human immune

system and to existing and novel antiretroviral drugs is vital for the global efforts to control the HIV-1 pandemic.

This cumulative thesis combines several computational approaches all aiming at determining viral adaptation using statistical learning techniques. The herein presented work is focused on adaptation of HIV-1 to the adaptive immune system and to a novel treatment option with bNAbs. A major challenge in modeling adaptation is the diversity of the viral proteins of interest, such as Env, but also interindividual variability of the immune system response. For a better understanding of the projects, further biological details are provided on viral adaptation to the adaptive immune system and to current antiretroviral drugs as well as the role of the envelope protein glycoprotein gp160 (Env) and the coreceptor usage for viral adaptation. In addition, some general challenges with respect to so-called trustworthy machine-learning based models are introduced. Afterwards the objectives and challenges shared across all projects of this thesis are presented, followed by a more detailed project-wise overview, and an outline of the structure of the remaining thesis.

1.1. Viral and host characteristics influencing viral adaptation

1.1.1. The role of the *env* gene

The genetic diversity of HIV-1 is also distributed unevenly across the viral genome. Among the nine genes of HIV-1, the gene *env* has the highest genetic diversity varying 15-20% within subtypes and up to 35% between subtypes [14]. A reason is the role of its encoded proteins in mediating viral entry while being the only target for the extracellular immune response. The gene *env* codes for the precursor glycoprotein gp160 (Env) that is further cleaved into the external surface protein gp120 and the transmembrane protein gp41 [15]. While gp120 is required for host cell recognition via the CD4 molecule and a coreceptor as well as for positioning, gp41 mediates membrane fusion. The proteins assemble to form non-covalent trimeric structures of gp120-gp41 heterodimers - so-called spikes - on the surface of the virion. Since the envelope of the virus is composed of a double-layered membrane of host-lipids uptaken in the budding process of the virus from the host cell, the envelope membrane is recognized as self by the immune system. Thus, the spikes are the only viral proteins on the surface and the only target for the immune system - more precisely for antibodies. Within this role, the envelope protein, foremost the gp120 protein, acquired several mechanisms to evade the immune system while remaining functionally intact: (i) high tolerance of mutations in non-functional parts to disguise the immune system, (ii) heavy glycosylation, steric occlusion and conformational shielding of functional important and conserved parts [16–18]. In contrast, the viral protein p24, which is encoded by the *gag* gene and forms the capsid of the virion, is rather conserved, not tolerating mutations that are not beneficial. While other genes are also under selection pressure and important for viral control, the genes *env* and *gag* are in the focus of this thesis.

1.1.2. The role of the coreceptor usage

Interindividual differences in viral control without treatment are mainly influenced by host factors like the individual HLA complex and by viral factors like the coreceptor usage [19, 20]. Upon binding to the the CD4 molecule, HIV-1 requires a second coreceptor for successful entry. Among many possible, only two coreceptors have clinical relevance in HIV-1 infection, namely, the C-C chemokine receptor type 5 (CCR5) and the C-X-C chemokine receptor type 4 (CXCR4) [21]. The genetic determinant of the coreceptor usage is the *env* gene, where the coreceptor binding site is located. Based on their coreceptor usage, viruses are termed *R5-capable*, if they only use the CCR5 coreceptor, or *X4-capable*, if they are able to use the CXCR4 coreceptor [22, 23]. Since the coreceptor binding is essential for viral entry, a defective CCR5 coreceptor provides a potential natural resistance to R5-tropic HIV-1 variants. The importance of the CCR5 coreceptor has been also exploited for drug development leading to the approval of the coreceptor antagonist Maraviroc, which blocks the CCR5 coreceptor binding site [24, 25]. In 50% of subtype B HIV-1 infected patients a switch from R5 to X4 usage occurs over the course of an infection. While the trigger mechanisms behind the coreceptor switch are still unknown, the clinical significance of a coreceptor switch with respect to pathogenesis are known. CCR5 coreceptor usage is usually observed early in infection and is associated with slow progression to AIDS, while CXCR4 coreceptor usage occurs at late stages of infection and is associated with rapid progression to AIDS, and $CD4^+$ cell depletion. This is in line with the knowledge that CXCR4 coreceptor usage is associated with lower glycosylation, which renders the virus more prone to antibody detection [26]. With $CD4^+$ cell depletion, the antibody production is not properly activated and thus the virus replicative efficacy is not reduced by having less glycosylation and present the vulnerable sites to the immune system. How CXCR4 coreceptor usage arises with moderate to high $CD4^+$ cells remains, however, puzzling.

1.1.3. Viral adaptation to antiretroviral drugs

Antiretroviral drugs for HIV-1, which suppress the replication of the virus to undetectable levels, are our only counteractive measure to control the virus, since it also reduces the probability of further transmission. Current drugs represent, however, not a cure, since they are not able to eliminate the virus within the patient. Instead, the different drug classes target and inhibit different steps in the replication cycle of the virus. A virus that integrates in the host deoxyribonucleic acid (DNA) but does not replicate, a so-called latent provirus, is, however, not affected by current drugs [27, 28]. Consequently, an interruption of the treatment results in a viral rebound from these so-called viral reservoirs. Thus, an infection with HIV-1 leads nowadays to a chronic disease requiring lifelong therapy. To counteract the emergence of drug resistances, combination antiretroviral therapy (cART) has been developed as an effective treatment. The treatment usually consists of a three-drug cocktail covering at least two different drug classes. While three-drug cocktails make the emergence of resistance mutation more unlikely, multi-drug resistances still occur due to the limited number of effective drug combinations, the high mutation rate of the virus, and the lifelong exposure to the virus. Consequently, a tight monitoring of the regimen is required to detect escape variants. Additionally, a constant development of new effective drugs is needed. New drugs will likely suffer from the emergence of

resistances in a similar way as established drugs and require therefore fast resistance testing approaches to reach clinical routine.

1.1.4. Viral adaptation to the adaptive immune system

Interindividual differences in viral adaptation are in particular observed to the adaptive immune system. The adaptive immune system is the second line of defense mechanism of our immune system that comes to action once the pathogen has not been eliminated by the innate immune system. T cells drive the adaptive cell-mediated immune response with their ability to distinguish self from non-self molecules on the surface of the host cells. Thus, they are able to detect and eliminate the pathogen within the host cells. In contrast, B cells are responsible for the adaptive antibody-mediated (humoral) immune response by proliferating to plasma cells and producing specific antibodies against antigens of the pathogen after activation by the T cells. Antibodies, also called immunoglobulin (Ig), are Y-shaped proteins that bind to antigens (specific molecules that trigger an immune response) in extracellular form and thereby either directly neutralize the pathogen by blocking viral entry or indirectly eliminate the pathogen by attracting other components of the immune system.

The role of the HLA system for viral adaptation

The complex interplay between the components of the adaptive immune system is based on the identification of non-self antigens on the surface of the host cells, so-called antigen presenting cell (APC), by the T-cell receptor of a naive T cell. This process is mediated by the major histocompatibility complex (MHC), which is also called the HLA system in humans. The main purpose of HLA molecules is to present peptides from within the cell on the cell surface for recognition by compatible T cells, which thereby detect foreign molecules. There are two major HLA classes. All cells apart from red blood cells present intracellular peptides via the HLA class I molecules. Upon recognition and successful binding to the specific HLA I-antigen complex, T cells with the CD8 marker, so-called cytotoxic T cells (CTL) or CD8 T cells, induce cell apoptosis by different mechanisms. HLA class II molecules are only present on professional APCs - specific immune cells that are able to uptake pathogens from extracellular fluid. The HLA II - antigen complex is recognized by specific T cells carrying the CD4 marker ($CD4^+$ T cells) and lead to the activation of B cells, followed by proliferation and differentiation of B cells into plasma cells that produce specific antibodies, and into memory B cells.

Hence, $CD4^+$ cells are a major coordinator of the immune system by mediating the information between the innate and adaptive immune system, cross-checking with the CD8 T cells, and responsible for the activation of B cells and thus antibody production. A depletion of the $CD4^+$ cells as happening in the course of HIV-1 leads to a disruption of the communication network and renders the immune system ineffective.

Viral genetic mutations that specifically hinder the binding or recognition process by the HLA molecule or the T cell are called HLA-restricted mutations and arise due to the specific individual pressure by the T cell receptors as well as the individual HLA alleles that determine which peptide fragments are presented. HLA I restricted mutations mask the infected $CD4^+$ cells from recognition by CTL cells. Thereby, less

$CD4^+$ cells are destroyed. HLA II restricted mutations mask the infected professional APCs from recognition by the $CD4^+$ T cell, such that no antibodies are produced. Hidden from the immune system, these escape mutations might enable the virus to evolve unrestricted.

1.1.5. Antibody-mediated adaptation and therapy

Antibodies are produced by the adaptive immune system against the envelope protein in the beginning of the infection, but usually they are neither broad nor potent enough to neutralize HIV-1 [29, 30]. Neutralization breadth is measured by the ability of the virus to neutralize different viral strains. The potency of an antibody is assessed by the required antibody concentration to reduce viral infectivity by 50% (IC50).

In 1993, the first generation of bNAbs was discovered [31–34], which, however, were not potent and broad enough to be clinically relevant [16, 35–37]. The development of high-throughput neutralization assays in 2005 and single-cell antibody cloning techniques in 2009 enabled the identification and production of a second generation of bNAbs from so-called *elite neutralizer* with a much higher potency and breadth [38–44]. Despite these beneficial characteristics, the bNAbs are not able to clear the infection in the corresponding patients due to the late stage of development and insufficient amount of bNAbs. [45, 46]. The new generation of bNAbs targets six different viral epitopes on the viral envelope proteins gp120 and gp41: the CD4 binding site, a V2-glycan site, a V3-glycan site, a glycan epitope on the outer domain of gp120, a membrane-proximal external region, and the interface region between gp120 and gp41 (reviewed in [47, 48]). After showing the efficacy of bNAb therapy in animal studies and humanized mice [49, 50], the first human clinical trials investigated the efficacy of a monotherapy with bNAb 3BNC117 [2] and VRC01 [51], respectively, both targeting the rather conserved CD4-binding site. The CD4bs bNAbs 3BNC117 and NIH45-46 are, however, more broad and potent than VRC01 [52, 53]. A review of all currently ongoing clinical trials with monotherapy and combination therapy with bNAbs is given in [54]. While bNAbs have the disadvantage of more complex requirements with respect to transportation, storage, and administration, they offer a new drug target for multi-drug resistant patients as well as a longer half-life time compared to existing drugs. Despite their neutralization breadth, viral escape occurs and thus, it is essential to determine if the patient harbors resistance mutations to a specific antibody prior to administration, similar as for current antiretroviral treatment options.

1.2. Challenges for genotypic prediction models for HIV-1 adaptation

Since the emergence of escape mutations have a quantifiable effect on viral adaptation, supervised learning approaches (a group of machine learning methods) are used to model the unknown relationship $f(X) = Y$ between the phenotypic observed change in adaptation (Y) based on the viral genotype (X). Based on available genotypic-phenotypic paired data, a function $\hat{f}(X) = \hat{Y}$ is learned that minimizes the error between model predictions ($\hat{Y}$) and the observed quantifiable outcome (Y).

The estimated adaptation for new samples X' is then only based on the viral genotype ($\hat{f}(X') = Y'$). Hence, apart from finding a suitable method that is able to model the relationship between X and Y, the quality and generalization ability of a machine-learning model depends on the available data that has been used for training. Whereas in the past a model-centric approach was rather deployed based on the assumption that more data and changing the model (parameters) will lead to a performance boost, there is a current trend towards a data-centric approach, where the quality of the data is seen as the key to gain better performance [55]. With the rising success of artificial intelligence (by adopting machine-learning techniques) and the increasing value of data, the protection of data and the trustworthiness of artificial intelligence (AI) models has gained importance, especially prior to the deployment of AI-based models in settings without human control or where life-dependent decisions are influenced. Consequently, several regulations and criteria have been developed recently (European High-Level Expert Group on AI [56], General Data Protection Regulation [57], European Data Act [58]). Since the consideration of these regulations and criteria increase the clinical relevance of machine-learning based models, we evaluate our models and results under these aspects in Chapter 2 though some of the regulations have been developed after the publication of our studies.

1.2.1. Trustworthy models

In their whitepaper on trustworthy use of AI, Englander et al. [59] present six audit areas that form the basis for the development of a future AI certification system in cooperation with the Germany's Federal Office for Information Security (BSI): (1) fairness, (2) reliability, (3) transparency, (4) data protection, (5) security, and (6) autonomy and control. The models and findings of this thesis are not discussed under security aspects and criteria for autonomy and control, since they are not relevant for the current state of the models. In the following, the first four criteria are further defined as in Englander et al. [59].

Fairness A key assumption in machine learning is that the training data set is a representative sample of the population of interest in the deployment setting. Unequal distribution of a variable between training and deployment might introduce a systematic error into the model (so-called bias) with potentially systematic differences in the outcome for underrepresented groups, thereby violating the right of equal treatment.

Reliability Reliability relates to the risk of a model to behave unexpected within the normal use case (accuracy and uncertainty), unexpected with unintended use (erroneous or adversarial attacks), but also if the data input is slightly changed (robustness).

Transparency In order to gain trust into the models (i) the user needs to understand the model, (ii) the model needs to be reproducible, and (iii) the model decision needs to be explainable. An often alternative used term for transparency nowadays is *explainable AI*.

Data privacy Data privacy concerns about the protection of sensitive information and re-identification of people within current regulations. A major concern

is that potential new technologies or findings will allow to reconstruct the identities based on previously published data.

1.2.2. Confounding variables

A brief definition of confounding variables is given in the following as well as the state-of-the-art methods to correct for them. Not accounting for potential confounding variables might add another form of bias into the model and potentially lead to unrobust behaviour of the model. All four projects put emphasis on correcting for potential confounding variables.

A confounding variable is defined as a third variable Z that has an effect on the independent variable X but is also a predictor for the dependent variable, Y. This confounding variable Z obscures the learned effect of X on Y by either overestimating or underestimating the estimated effect –the effect is *mixed*. Thus, confounding is a type of bias by adding systematic error to the model.

Stratification (or standardization) or nowadays multivariate analysis are common methods to adjust for potential confounders. In the latter, the idea is to test if a confounding variable Z has an effect by including it into the existing linear regression model and compare the new model with the model without the potential confounder variable (null model) using the likelihood ratio test over the maximized likelihoods [60]. An alternative approach is to compute Bayes Factors. The Bayes Factor computes the ratio of the likelihood of the Bayesian linear regression model with the confounding factors (alternative hypothesis) to the likelihood of the model without the confounding factor (null hypothesis) based on the marginal likelihoods [61, 62]. Once a confounding variable is detected, there is the possibility to remove the association between confounding variable and independent variable by weighting each sample with the inverse-probability of the occurrence of the sample with the confounding variable [63].

1.3. Project overview

1.3.1. Overall objectives

This cumulative thesis comprises four different projects that all share the objective to develop computational approaches to further our understanding of HIV-1 adaptation. On the one hand, the methods aim to advance the development of a new promising treatment option involving bNAbs and deal with antibody-mediated adaptation (P1-P3). On the other hand, the developed methods enable to investigate the relationship between viral adaptation to the T-cell-mediated immune response and the coreceptor usage that has remained unexplored so far to the best of our knowledge (P4).

In two projects (P1 and P4), we develop machine learning-based support decision tools to predict and analyze HIV-1 adaptation with respect to antibodies and the host immune system, respectively. A challenge in modeling adaptation is to generalize well across intra-host HIV variants and the global diversity of HLA variants based on a rather small data set. Therefore, we designed the models to be as flexible as possible by incorporating as little prior information as possible. Another shared

characteristic is the relevance of the coreceptor usage for antibody-mediated and cell-mediated adaptation.

The aim of the two other projects was to investigate the clinical efficacy of a therapy with a specific bNAb in a clinical trial (P2 and P3). Here, we supported the clinical trial by using statistical learning methods to statistically validate the effect of the antibody therapy while adjusting for potential confounding variables. In addition, an alternative method to measure antibody neutralization was developed since the classical approach of using the IC50 value was suitable for the underlying data set.

1.3.2. Paper 1 (P1) - bNAb resistance study

A genotypic resistance test for bNAbs is essential for bNAbs reaching the clinical routine but also for ongoing clinical trials to select patients that do not harbor initial resistances to bNAbs. It is likely that a combination of bNAbs will be required to counteract the emergence of resistance mutations. Nevertheless, it is important to first have a system for predicting resistance to single bNAbs. Therefore, we investigated whether neutralization response to bNAbs is predictable based only on the envelope sequence for bNAbs covering different epitopes with a non-linear machine learning method. In particular, we used existing paired data of the envelope protein sequence of pseudoviruses and corresponding neutralization response (IC50) for 11 different bNAbs to train a support vector machine (SVM) classifier [64, 65] to predict the neutralization response given the envelope sequence of the virus. By using the oligo kernel [66], which encodes a sequence by the occurrence of its substrings of size k, we are able to also model more complex binding sites of the bNAbs. We put emphasis on the explainability of our non-linear models, since in general the learned coefficients for the features cannot be interpreted directly as feature importances in contrast to linear models. We also focused on possible confounder effects like the coreceptor usage on the neutralization capacity of bNAbs, as they share the same genetic determinant. In addition, the aim was to use our prediction model to test whether there is a trend towards antibody resistance in the global viral evolution to extend and confirm previous findings from a small cohort with subtype B HIV-1 infected patients [67–69].

1.3.3. Paper 2 (P2) - bNAb efficacy study I

While the development of a genotypic resistance method is essential for bringing bNAbs to the clinical routine, first the efficacy and the characteristics of such a bNAb treatment needs to be investigated and proven. Antibody-mediated immunotherapy has the potential to impact the human immune system [70, 71].

In this study, we test the hypothesis that a single infusion of the bNAb 3BNC117 can enhance the neutralization activity of the immune system of the patients. Therefore, we compared the neutralization activity of patients receiving a single 3BNC117 infusion without antiretroviral treatment (Group A, n = 15), and on ART (Group B, n = 12) against a control group of viremic patients not receiving the treatment (Group C, n = 36). To study whether there is a change in neutralization activity, neutralization activity is measured with patients' antibodies (IgG) before the start of bNAb treatment at day 0 (d0), and after 24 weeks (w24), when the remaining an-

tibody concentration is negligible. Neutralization activity of the d0 and w24 IgGs is then investigated on the patients' matched viruses at d0 and week 4 (autologous response) and on a panel of HIV-1 pseudoviruses (heterologous response). Some of the patients do not reach the IC50 value that is usually taken to measure neutralization, though showing a change in neutralization between the two time points. A major contribution was the development of an alternative method to quantify neutralization by considering the area under the neutralization curve (AUC) upon which the complete analysis is based on. The benefit of using the AUC has been shown previously [72]. For a more robust measure, we correct the curves for noise and variance in the experimental assays by using the neutralization information from replicates and a control virus, and scale the neutralization curves by the maximum possible AUC. A change in neutralization activity might be confounded by other factors like difference in viral load, CD4 T cell count, or initial neutralization activity. To test whether potential confounding factors have an effect on the observed change in neutralization activity, we used a likelihood ratio test approach and a Bayes Factors approach, respectively, to study whether the inclusion of the potential confounding factor leads to an improvement of the model or not. The differences in the neutralization changes are compared across different treatment groups and control groups using appropriate significance tests.

1.3.4. Paper 3 (P3) - bNAb efficacy study II

Apart from determining the effect of bNAb therapy on neutralization activity, it is important to quantify the actual effect of the treatment - namely the ability to suppress the virus. Due to the existence of latent viral reservoirs, antiretroviral treatment interruption results in viral rebound. To study the effect of 3BNC117 infusion, a phase IIa open label clinical trial is conducted where a controlled analytical treatment interruption (ATI) in 13 HIV-1 infected individuals is performed. Prior to ATI, these individuals received a treatment with 3BNC117. The time until viral rebound in the treated group is compared to a historically untreated group, where treatment interruption was performed without antibody treatment. The challenge is to correctly model the time to viral rebound, and adjust suitable survival regression methods for potential confounder variables like age or year of infection.

1.3.5. Paper 4 (P4) - HIV immunoadaptation study

In contrast to the three previous studies, the fourth project investigates the adaptation to the cell-mediated immune response, namely adaptation to T cell pressure and its role on the coreceptor usage. Though, the HLA system and viral coreceptor usage are known to be important factors in viral control in treatment-naive patients, their interplay has not been studied so far.

As mentioned in Section 1.1.2, the CXCR4 coreceptor usage is usually observed together with a depletion in $CD4^+$ cells. How CXCR4 coreceptor usage arises with moderate to high $CD4^+$ cell counts remains, however, puzzling. A possible hypothesis is that viral adaptation to the cell-mediated immune response determined by the emergence of HLA-restricted escape mutation has a similar effect like $CD4^+$ cell depletion. The HLA-restricted escape mutations prevent the HLA presentation or the T cell recognition of the viral peptides. HLA I-restricted mutations mask the

virus from CD8 T cell recognition and thus the CD4 cells from elimination. HLA II-restricted mutations prevent the virus from detection by CD4 cells, and thus the activation of the antibody production is suspended. Hence, these escape mutations enable the virus to evolve without further pressure by the immune system, eventually leading to mutations and changes in glycosylation as required for the CXCR4 coreceptor usage.

To test this hypothesis, three components are required: (1) a tool to predict viral adaptation to the HLA I and HLA II molecules, (2) a data set that comprises treatment-naive patients with known HLA I, HLA II genotypes with matched viral proteins Env (for determining the coreceptor use) and (3) a rather conserved viral protein to learn the cell-mediated adaptation mutations.

To the best of our knowledge, there is no publicly available data set meeting our requirements. Therefore, we sequenced the viral *env* and *gag* gene from 312 treatment-naive, chronically subtype C HIV-1 infected individuals from South Africa. In addition, we genotyped the HLA I and HLA II alleles of the cohort. The sequences were provided by the Fraunhofer Institute for Biomedical Engineering, and sequenced and genotyped by the SEQ-IT GmbH & Co.KG in Kaiserslautern, Germany.

While the existing approach to predict viral adaptation to the HLA I alleles [73] has been also applied for HLA II molecules [74], there is no approach that jointly predicts HLA I and HLA II adaptation. To reduce the complexity of the viral diversity but also the diversity of the HLA system, the existing approach is based on pre-learned HLA-polymorphism candidates from a separate large genome-wide association study. Apart from the preselected HLA-polymorphism pairs, the algorithm requires further feature selection steps to find the most important HLA alleles per polymorphism. Although the algorithm corrects for potential phylogenetic relatedness of the viral strains, it requires a second model therefore, which is combined with the adaptation model. In addition to the rather complex construction, the underlying data for training the algorithm are not publicly available such that extending the approach to jointly model the HLA I and HLA II adaptation was not an option.

Therefore, we developed a novel computational approach that jointly models HLA I and HLA II adaptation while correcting for potential confounders like phylogenetic relatedness, age, sex, or ethnicity within the model. The basic idea behind our approach is similar as in Carlson et al. [73]. The adaptation of a viral sequence to the HLA system is decomposed to the ratio between the likelihood that each variant residue of the viral sequence has emerged under HLA pressure and the likelihood that it emerged without HLA pressure. The conditional probability of the occurrence of the variant site under or without pressure is modeled using Bayesian generalized linear mixed models, respectively. Using the horseshoe prior on the coefficients of the model allows to incorporate the complete HLA repertoire of the cohort without prior preselection of potential candidates. By construction, we are also able to provide sequence logos determining which variant sites contributed most to the estimated adaptation. Using our adaptation models, we are able to investigate the correlation between viral adaptation and the coreceptor usage that has remained unexplored so far.

1.4. Thesis outline

The remaining thesis is structured in the following way. In Chapter 2, the results of all four papers, which are included separately in Appendix A, are summarized and discussed together. Section 2.1 provides a high-level summary for each project as well as an overall summary. In addition a project-wise but also a cross-project-wise discussion is included in Section 2.2 discussing the realization of the objectives, the limitations of the projects, and the embedding in current research setting. The presented work is concluded in Chapter 3, where additionally a perspective is given on the requirement for clinical relevant genotypic recommendation systems in Section 3.2.

2. Results and discussion

> All models are wrong, but some are useful
>
> *George E. P. Box*

As pointed out in Chapter 1, assessing viral adaptation is of great importance to effectively control HIV-1. This chapter starts with a high-level summary of the contributions and results for each of the four included studies (see Appendix A), followed by an overview across all studies emphasizing the shared challenges and findings. Afterwards, the findings and developed models are discussed with regard to their limitations and possible extensions. In particular, aspects of clinical relevance, data quality, model choice, and choice of performance measures are taken into consideration. Additionally, the findings are inspected with respect to criteria for trustworthy machine-learning models including fairness, robustness, explainability, and data privacy aspects.

2.1. Results

The overall aim of this thesis was to provide computational methods to predict viral adaptation to immune system pressure - namely to broadly neutralizing antibodies and to the host immune system. In particular, the objective was to support the advance of broadly neutralizing antibody therapy to clinical routine and to investigate whether viral adaptation to the host immune system is associated with viral coreceptor usage. On the one hand, machine-learning based prediction tools have been developed to predict adaptation to bNAbs and the HLA system (P1, P4). On the other hand, we used statistical learning techniques to characterize the effect of bNAb therapy in clinical trials to pave the way for bNAbs from benchside to clinical routine (P2, P3).

2.1.1. Paper 1 (P1) - bNAb resistance study

A support-decision tool for determining neutralization susceptibility is essential for the usage of bNAb-mediated therapy in clinical routine, similarly as for existing antiretroviral drugs. The following section summarizes the developed models and findings from the bNAb resistance study [1] (see Appendix A.1).

For 11 different bNAbs, we observed that bNAb resistance is well predictable using only the envelope sequence with prediction performances up to 84% AUC for the bNAb 10-996. Using SVM-based models with a cost-sensitive regularization parameter enables to control for the imbalance in the class distribution of the data. We have chosen the oligo kernel [66] based on its superior performance in comparison with other kernels. In addition, we observed that the size parameter k of the oligo kernel, denoting the length of the k-mers to consider, correlated with the length of

the binding sites of the 11 different bNAbs. Moreover, by construction of the oligo kernel, we were able to consider non-linear relationships of the amino acids, yet still explain the feature importances learned by the model. Based only on the viral envelope protein, the trained models have learned parts of the binding sites of the bNAbs but also potential glycosylation patterns as important features, indicating thereby their biological relevance. A novelty is the derivation of the contribution of each amino acid of the query sequence to the classification outcome using the oligo kernel definition presented as a sequence logo. Hence, the model is transparent with respect to learned discriminant features but the reasoning behind each model decision is also explained via the sequence logos.

To analyze bNAb neutralization susceptibility over time, we used our trained SVM regression models to predict the neutralization sensitivity of roughly 34 000 HIV-1 samples from the LANL HIV sequence database (http://www.hiv.lanl.gov/). In this data set, we observed a trend towards increased bNAb resistance over time for all 11 bNAbs for the subtype B variants, and for 6 out of 11 bNAbs for the other subtypes. This supports and extends a previous finding on a small subtype B HIV-1 cohort [67, 68]. Additionally, we discovered a bias in the LANL HIV sequence database with respect to the coreceptor usage of the samples over time. The frequency of viral strains with X4-usage is decreasing over time. We further observe that the neutralization capability of the bNAbs PGT121 and PGT128 have a coreceptor bias similar to PG9 and PG16 [75], meaning that they are better in neutralizing HIV-1 variants with R5-coreceptor usage than X4-coreceptor usage. This finding further suggest that coreceptor usage is an important confounding variable and should be considered when assessing neutralization capability. Additionally, the administration of a bNAbs with an known R5-coreceptor bias within a bNAb therapy can potentially exert a biased selection pressure towards X4-usage on the virus that is associated with faster progression to AIDS.

Apart from its usage as a fast and cheap resistance genotypic-to-phenotypic prediction method for a monotherapy with a bNAb, our approach can be used as a foundation to build the best combination of bNAbs as potentially required for clinical usage. As the effect of a bNAb therapy strongly depends on the susceptibility of the patient's quasispecies to the bNAb, current ongoing clinical trials can benefit from our model by using it for the patient screening process. Thereby, it is possible to determine the patients' sensitivity to the bNAb of interest prior to their inclusion in the study instead of performing viral outgrowth cultures. The learned discriminative amino acids in our model, which confer susceptibility or resistance, can guide the selection of potential immunogens for vaccine design as well as the discovery of epitopes of new bNAbs if only the sequence is available.

2.1.2. Paper 2 (P2) - bNAb efficacy study I

Apart from its neutralizing activity, antibody-mediated therapy has the potential to engage the host immune system with the crystallizable fragment (Fc) region of the antibody that interacts with other host immune cells like the natural killer cells. The following section is based on Paper 2 [2] (see Appendix A.2). The objective of this study was to investigate whether an infusion of the bNAb 3BNC117 has an effect on the neutralization activity of the patient over a 6-month period. A major challenge was that many included study participants did not reach the half maximum

inhibitory concentration (IC50) value that is usually taken to measure neutralization capacity of the bNAbs. The personal main contribution was the derivation of a robust, noise- and variance-corrected version of the area under the neutralization dose-response curve (AUC) as alternative measure to the IC50 value. The neutralization AUC was the basis for comparing the change in neutralization in all individuals before and after treatment administration, including patients not reaching the IC50 value. Based on the neutralization AUC, we observed a statistically significant difference in neutralization change between treatment and control groups indicating that an infusion has the capability to boost the patients immune system beyond the treatment. The increase in neutralization capability was significantly less pronounced in ART-treated individuals compared to viremic individuals, both receiving the bNAb therapy. Using linear and Bayesian regression techniques, we found that the increase in neutralization activity is not confounded by factors like the initial neutralization activity at day 0 (d0), age, sex, initial viral load, or $CD4^+$ T cell levels. We also observed no correlation between the neutralizing activity at day 0 and neutralization improvement.

The findings support the hypothesis that an immunotherapy with bNAbs can boost the host immune system against HIV-1 beyond the bNAb therapy. Moreover, the findings also suggest that irrespective of the host genetics or viral quasispecies, the patients have the potential to develop broadly neutralizing antibodies. This is important for the vaccine design research that aims to trigger the development of bNAbs with immunogens.

2.1.3. Paper 3 (P3) - bNAb efficacy study II

A prerequisite for bNAb therapy reaching clinical routine is its efficacy proven in human clinical trials. In the following, the major findings of Paper 3 are presented [3] (see Appendix A.3).

With this study [3], we further contributed to the advancement of bNAbs to clinical routine by investigating whether 3BNC117 can suppress viral rebound from the latent reservoir during analytical treatment interruption (ATI) in chronically suppressed HIV-1 infected humans in a phase IIa open label clinical trial.

By comparing the time to viral rebound between the treatment group (n=13) and a historical control group (n=36) after ATI using Kaplan-Meier curves, we observed that viral rebound in the treatment group was statistically delayed by an average of 6.7 (2 infusions) and 9.9 (4 infusions) weeks compared with 2.6 weeks for historical controls. We investigated whether variables such as gender, CD4 nadir (lowest CD4 point), age, years on ART, and CD4 count prior to ATI are predictive for viral rebound to rule out potential confounding effects by these variables. Therefore, we compared a prediction model with the potential confounder against a model without the variable of interest using a likelihood ratio test. Rebound time was modeled as log-normal distribution based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC) of a goodness-of-fit test. Based on the data and our tests, we observed that *years on ART* and *age* might be confounding factors for rebound time. Thus, the change in rebound time might be masked by these variables and a proper test to compare the two treatment groups should thus account for the confounding variables. Since the ratio of viral rebounds between the groups differ over time, the proportional hazard ratio assumption by the classical Cox-

survival regression model was violated. Thus, we used a weighted log-rank test [76] to additionally adjust for the potential confounders *years on ART* and *age*. Adjusting for the confounders is achieved by computing a separate weight for each sample that is inverse-proportional to its frequency with respect to the confounder variable, thereby the potential confounding effect is canceled out. The viral rebound was also delayed in the treatment group, when adjusting for the confounding variables using a parametric survival regression approach.

Showing that a monotherapy with 3BNC117 is safe and statistically delays viral rebound is a further step towards the development of a combination immunotherapy for clinical routine. In our study, it was also observed that the therapy might impact the viral reservoir. The occurrence of resistant viral strains at rebound suggests that viral reservoirs harboring resistant variants are selected by an infusion with 3BNC117. Moreover, the viral outgrowth seems to be restricted after the therapy. The extent to which bNAb therapy has an impact on the viral reservoir needs to be evaluated in further studies.

2.1.4. Paper 4 (P4) - HIV immunoadaptation study

The following results are based on Paper 4 [4] (see Appendix A.4). In this study, the aim was to investigate whether there is a relationship between the coreceptor usage and viral adaptation to the host adaptive immune system represented by the host HLA I and HLA II alleles. Especially if viral adaptation enables a coreceptor switch to the X4 usage with still moderate to high $CD4^+$ cell counts.

Central to this project is the novel approach to jointly predict HLA I and HLA II adaptation using Bayesian generalized linear mixed models. Using the horseshoe prior for the coefficients of the model, we were able to incorporate the complete HLA repertoire in the data cohort as potential predictors without the usage of p-value-based greedy feature selection methods or prior extensive search for HLA-polymorphism candidates on external data. The horseshoe prior is a shrinkage prior shrinking most coefficients to zero and only allowing some large coefficients to escape shrinkage [77, 78]. Thus, it is suitable for sparse models that only have a few features with predictive power. Since the HLA molecules have different binding sites, only a few HLA alleles exert selection pressure at a specific site in the viral genome and might drive the emergence of an HLA-restricted escape mutation.

Though there is no ground truth for viral adaptation within the available data, we validate our model by observing that certain expectations are met: (1) adaptation of autologous viruses are higher compared to heterologous viruses, (2) viruses are more adapted to host HLA profile compared to random shuffled HLA profiles, and (3) viruses in chronic patients are more adapted compared to acutely-infected HIV-1 patients.

Moreover, we observed that the per-site models have learned known important HLA footprints, and that the most predictive sites are some known HLA-restricted single variant sites. This indicates that the adaptation score is based on biological relevant models. Moreover, it shows that our adaptation model is transparent with respect to the learned features and their contributions to the adaptation score. Due to the definition of the adaptation score, the estimated adaptation of a viral sequence to its host HLA profile is completely explainable by the probabilities of each potential single variant site to be under HLA pressure. In addition, we provide motif logos

that show the single variant sites in the query sequence that contributed to the corresponding predicted adaptation score.

By comparing the distribution of the estimated HLA adaptation with coreceptor usage, we observe that in general HIV-1 variants with X4 coreceptor usage are more adapted than variants with R5-coreceptor usage. Moreover, in variants with high R5-coreceptor usage, higher adaptation is related to a higher FPR score by the geno2pheno[coreceptor] tool.

Whether or not increasing adaptation in R5 variants is indeed an indicator for an imminent coreceptor switch has to be further determined in future studies. Viral adaptation might, however, be an additional discriminative feature for determining the correct coreceptor usage as required for current CCR5 antagonist drugs. The identified potential variant sites under HLA pressure might guide current vaccine designs to not use these sites for immunogens as they might depend on the host genetics. Larger sample sizes harbor the potential to find new HLA-polymorphism candidates for subtypes that have been neglected so far.

2.1.5. Cross-study results

The developed prediction models (P1, P4) share the characteristic that they provide a rather general framework that can be easily applied to new antibodies (P1) or additional viral proteins and subtypes (P4). While the initial model parameters are the same for each bNAb (P1) and each per-site model (P4), internal feature selection takes place by learning different settings of the oligo kernel parameters (P1) as well as by applying the horseshoe prior (P4). The feature selection is essential for both approaches to tackle the challenge of the high viral genome diversity (P1) and the diversity of the HLA repertoire in the global population (P4). We decided to use internal model-based feature selection processes to circumvent the commonly used p-value-based feature selection methods that might lead to overfitting and non-reproducible results. We also observe that adaptation is already predictable using only the viral genome as input, i.e., without incorporating additional prior knowledge such as the structural information of binding sites, glycan patterns, the bNAb sequence (P1), or known HLA-polymoprhism candidates (P4). We refrained from incorporating this kind of information for several reasons. First, the model should be usable in settings where this prior knowledge does not exist, such as for a novel bNAbs (P1) or other populations of interest (P4). Second, the model should be as simple as possible, because the models already have more parameters than samples.

In general, the results from both approaches indicate that only few sites in the viral genome confer the adaptation. Thus, a more fine-tuned approach using only the predictive features might lead to a better prediction power.

A major focus of both prediction models was to provide explainable machine-learning based model despite learning potential non-linear relationships. This was possible because both models have decomposable characteristics. The oligo kernel has an explicit decomposable feature mapping function, which enables to trace back the learned features for each oligomer at each site in the viral genome, even for larger sizes of the oligomer. The adaptation model can be decomposed into the separate models for each variant site in the viral genome. Thus, for both models we are able to show which sites in the genomes contributed the most to the predicted outcome. The

learned discriminant sites could be partly mapped to known discriminant sites from experimental validation reported in the literature, thereby indicating the potential biological relevance of the learned models. Whether the remaining unknown learned discriminant sites are under adaptation pressure needs to be validated in further knock-out experiments. Learning known discriminant adaptation sites makes the models also suitable for applications where limited prior knowledge exists about the adaptation pressure. In addition, we also provide sequence logos in both projects showing the contribution of each site in the query sequence to the final prediction outcome. This enables the user, with their expertise, to make an informed decision based on the prediction outcome.

We also observe that adaptation to antibodies as well as adaptation to the HLA system have in common that they differ depending on the coreceptor usage.

In conclusion, both prediction approaches provide a useful framework to assess the adaptation of HIV-1 to bNAbs and the host immune system. The identified sites under selection pressure are also of interest to the HIV-1 vaccine research. It is advantageous to select immunogens for the vaccine that are susceptible to bNAbs or that can induce the development of bNAbs, with low probability of mutation. Moreover, the immunogens should not be under the individual selection pressure of immune system to be globally effective. The methodological frameworks in both projects are not HIV-dependent and can easily be applied to study virus-host or virus-antibody interactions in other pathogens such as influenza or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

Across all projects we observed that there are usually several methods that can be applied leading to very similar results. We also showed that it is important to control for potential confounders.

2.2. Limitations and extensions

The here presented projects build a good foundation for clinical-relevant support-decision tools and present important findings to advance immunotherapy with bNAbs. Still many aspects have to be addressed until the models and findings make the transition from research to clinical routine.

In the following, the presented models and findings are discussed with respect to different criteria. First, the appropriateness of the choice of data, model and performance for the applied tasks is discussed. Then, the projects are further evaluated with respect to the criteria for trustworthy artificial intelligence models, in particular with respect to fairness, robustness, transparency, and data privacy. Another important factor for the utility of the findings and models is the consideration of the future trends and advancements with respect to technology, requirements, and changing use cases. While guaranteeing a safe and secure execution of a support-decision tool is essential for clinical routine, the models are not analyzed with respect to this criteria, since they are currently not in deployment.

2.2.1. Data quality

Currently, there is a paradigm shift in the machine-learning community from model-centric approaches to data-centric approaches supported by Andrew Ng, who initiated the NeurIPS Data-Centric AI Workshop in 2021 [79] and launched the Data-

Centric AI Competition [80]. Model-centric approaches assume that more data and changes in the model result in a performance boost in contrast to data-centric approaches, where the improvement of the data quality is considered the key to a better prediction performance.

A data-centric approach additionally facilitates the production of trustworthy models since criteria such as fairness, robustness, and data privacy are inherent to data quality rather than to model quality.

All projects presented in this thesis share the characteristic that either no real-world data were available, no ground truth was given, or no comparison with other research results was possible due to the novelty of the research question. In addition, small sample sizes limit the generalization of the results. Further, the models share the characteristic that they provide rather a general framework than fine-tuned models. Thus, in our case, it is very likely that more representative data will yield more robust performance. However, it is possible that a fine-tuning of the model parameters to specific tasks (bNAb (P1) or polymorphism (P4)) or incorporating more informative features (like glycosylation) might boost the prediction performance.

Sample distribution shift

Central to each machine learning model is the assumption that the training data set is a valid, representative sample of the population distribution. The estimated generalized prediction performance of the model in the development stage would otherwise not match the prediction performance in the deployment stage on real-world data. However, distribution shifts may occur, where the real-world data differs systematically from the training data. A possible shift in the data distribution from training to deployment time can occur through changes at three levels: the input data, the output data, and the relationship between input and output data.

Due to the novelty component in all four projects, the available data at development stage of the models and studies are not completely representative for the global population or the deployment setting. Thus, to increase the reliability of the results further more representative data are required. In the following, these potential issues are discussed further.

Pseudovirus neutralization panels Due to the novelty of the project at the development stage, the training data for our neutralization prediction models consisted of neutralization assays with single-round-of-infection Env-pseudoviruses instead from HIV-1 isolates from clinical patients. Thus, it is possible that the performance on real HIV-1 isolates from patients might be different than reported on pseudoviruses. There is a potential shift in the data distribution on all three levels described above.

First, the input data are not the same at deployment time, since the viral env protein in the training scenario was based on the artificially created pseudovirus. Thus, the observed sequence space can be systematically different compared to clinical data. Though the panel data are aimed to be very diverse with respect to tier or subtypes, it might not be representative for the real viral population that will be present in clinical settings.

Second, the scale of the output data might be shifted at deployment time because of a constant increase both in the potency of the bNAbs [81] and in the resolution of the neutralization assays over time [82]. Consequently, the IC50 cutoff to distinguish

between susceptible and resistant samples might shift over time. Therefore, we suggest not to use predefined cutoffs for dichotomizing the neutralization capability, but rather learn regression models with censoring for the neutralization detection limit. Models learned in that way are more robust to the dynamic development of new bNAbs. Still, it is possible that the IC50 value distribution might shift closer to 0 in the future with increasing potency of the modified bNAbs. In addition, with newer bNAbs and assays with better resolution on their way, it is advisable to adjust for sampling time as potential confounder in future models.

Third, the model learned the relationship between neutralization capacity and the viral Env protein from pseudovirus neutralization assays. Host factors like ethnicity, age, but also variables like viral load, CD4 cell count, or previous drug exposure might have an effect on the neutralization capacity of the patient in contrast to the assay environment. With many ongoing clinical trials, new data are currently generated and should be used for further studies and compared to our results.

The CATNAP database [83] has recently been established storing all bNAb related data. An automated retraining of the models using the CATNAP database would have two benefits. First, the models are updated to the latest bNAbs and learned with the maximal available amount of data. Second, a change in the data distribution is detected faster. However, by merging several datasets together, the models have to be corrected for potential batch effects like year, assay, publication, or number of replicates. In a clinical setting, it is also of interest to investigate whether the clinical data distribution changes over time to detect a shift in the data distribution or a shift in the use case application.

Clinical trial sample data There is the potential that the selection of patients is not a good representative for the global HIV-1 infected population, due to the small sample size in the initial phases of clinical trials. Therefore, the sample size and the patient selection criteria are usually expanded in the follow-up phases of the clinical trials to observe if the previous results can be confirmed. Moreover, it has to be investigated further whether other host factors or previous medication history impact the effect of the bNAb therapy. If bNAbs will be used as an additive component to existing antiretroviral treatments in the future, potential cross-effects have to be examined in more detail.

Patient cohort to study HLA adaptation To investigate the relationship between viral T-cell based adaptation and viral coreceptor usage, the *gag* and *env* genes of a cohort of treatment-naive, chronic, subtype C HIV-1 infected individuals from South Africa have been sequenced. In addition, the host's HLA I and II alleles have been genotyped, which exert pressure on the viral genome. Alterations in the *gag* gene conditional on the host HLA profile were used to learn the T-cell based adaptation of HIV-1, while the *env* gene was used to predict the coreceptor usage. A data set consisting of chronically-infected HIV-1 patients is useful for learning the acquired adaptation mutations, since there is enough exposure time of HIV-1 to the host HLA system to accumulate HLA-restricted adaptation mutations. However, only few HIV-1 patients harbored viruses with X4 coreceptor usage with intermediate to high CD4 cell count at the same time (n = 27).

We selected treatment-naive individuals to learn adaptation solely based on the HLA system without a potential masking from drug-dependent resistance mutations.

However, in real-life applications, it is of interest to assess the viral adaptation to the HLA system in treated individuals. A potential issue could be that the model is not able to distinguish if a mutation is equally likely to have emerged by drug exposure or if the mutation patterns are different in drug-exposed individuals. The same issue of potential distribution shift is given by changing from chronically infected to acutely infected HIV-1 patients, though viral adaptation is less likely to be observed in acutely infected patients. An exception are mother-child transmissions, where the host immune system is partly shared and transmitted adapted mutations might be beneficial for the virus and an important factor for the treatment decision.

Due to the small sample size in comparison to the large amount of possible global HLA profiles, larger sample sizes are required to verify if the HLA distribution underlying our adaptation model is representative both for South Africa and the global population.

While the data is appropriate to learn viral adaptation, longitudinal data would be beneficial to investigate whether increasing HLA adaptation is associated with an impending coreceptor switch. Using longitudinal data of HIV-1 patients with different $CD4^+$ cell counts enables to observe whether a change in coreceptor usage is associated with viral adaptation and differs with respect to $CD4^+$ levels. Note that the duration of infection also has an effect on the viral adaptation but also on the coreceptor switch and thus should be adjusted for.

Although, the viral protein p24 is commonly used to study adaptation, other genes are also under selection pressure by the HLA system. Since potential cross-effects might exist, it would be beneficial to jointly model the adaptation to all genes.

Our approach focused on single amino acid polymorphisms. Consequently, compensatory mutations have been neglected that can reduce or enhance the adaptation effect of escape mutations. Therefore, it might be beneficial to consider interactions between sites.

Another rather technical issue is that currently usually only HLA I alleles are genotyped without the HLA II alleles. In addition, the resolution of the genotyped alleles is often restricted to HLA supertypes, which classify the HLA alleles based on their shared binding site specificities. The HLA supertypes can be used as a feature reduction technique to reduce the vast amount of possible HLA alleles. However, it is unclear whether all HLA alleles from a supertype exert the same pressure on the virus. More frequent HLA alleles within a supertype might also dominate the potential learned effects for a supertype in such a model setting. Hence, the models are less interpretable. Nevertheless, it is possible to include the dependencies of HLA alleles belonging to the same supertype by modeling the relationship as a random component of the GLMM.

Sample size

If the sample size is small, the sample distribution might not properly represent the underlying population distribution. Consequently, not all potential relationships can be observed and unequal distribution of certain categories of the variables (such as year, sex, ethnicity, HLA alleles, subgroups, coreceptor usage, duration of HIV-1 infection) can lead to the introduction of various forms of bias.

Small sample sizes are a common problem in healthcare, where the number of potential predictors p is often larger than the number of samples n. At the same time, it is recommended to have at least 5 to 10 samples per degree of freedom in

the model for a robust generalization performance [84, 85]. Hence, feature reduction or selection methods are often applied to reduce the number of predictors. The sample size in all four projects was rather small with (P1) 115 –220 samples, (P2) 15 treated viremic and 12 treated non-viremic against 36 viremic control patients, (P3) 13 treated and 52 control patients, and (P4) 268 samples (274 samples for the HLA I model).

While the developed models have learned already known discriminant signals, the training data might not have been large enough to observe less pronounced but still discriminant features. There is potentially not enough statistical power to learn associations of less frequent mutations or HLA profiles. Potentially, this could lead to less reliable results on future data. Quantifying the uncertainty with regard to unseen viral mutations or HLA alleles would be a potential countermeasure.

Fairness

The unequal distribution of a variable within the training data set might lead to a biased outcome for samples with less frequent observations for this variable such that the expected performance is over- or underestimated for these samples. If models systematically perform worse for the underrepresented part of the variable, this can have very discriminating and harmful effects in the healthcare sector.

While the pseudovirus panel data set is collected such that it is representative for a variety of different strains with respect to subtypes and the level of difficulty to neutralize the virus (tier), the frequency of subtypes A, B, C clearly dominate over other minor subtypes. Demographic and patient data such as sex, age, immune status, year of infection and others have not been considered as the data set is based on artificially created viruses. We show that some bNAbs such as PG9, PG16, PGT121, and PGT128 have an R5-bias by neutralizing R5-capable variants better compared to X4-capable variants. Thus, it is important to consider coreceptor usage not only to get correct prediction performances for each subgroup, but also not to drive patients with R5-capable variants to a coreceptor switch.

In clinical trials, the distribution of clinical and demographic variables between the treatment and control groups are matched, if possible. Nevertheless, the patient cohort does not adequately represent the variety of ethnicities, age distribution, or clinical variables such as viral load and CD4 count in the global population due to small sample size. A descriptive statistic of all available information as well as a test for confounding variables within the data has been performed, however.

To study the adaptation of HIV-1 to the immune system in Paper 4, we narrowed down the patient cohort to subtype C, treatment-naive, HIV-1 infected individuals from South Africa. While we have used a coarse definition of ethnicity, we were not able to distinguish between the many different ethnic groups in South Africa due to inconsistent labeling. In addition, the current predictive performance is limited to subtype C HIV-1 patients. It needs to be confirmed yet, how the adaptation models deviate for different subtypes. Moreover, the adaptation model is tailored to treatment-naive patients, and might be less predictive for acutely infected HIV-1 patients.

Data resolution

A possible performance boost can be achieved by using higher data resolution, moving away from averaged information to more fine-grained, not aggregated variables. In theory, it might be beneficial to use the dose-response curve itself instead of the area under the neutralization curve or the IC50 value [86, 87]. A drawback, however, is that the dose-response curve has to be modeled again using the antibody dilution series containing only several points and few replication measurements, if any. Moreover, there are several approaches for modeling the dose-response curve with respect to the shape of the curve. On top, it is not trivial to compare two dose-response curve with respect to the model parameters or testing if the two dose-response curves differ significantly.

While using the next generation sequencing reads might allow to have a better resolution of the HIV-1 quasispecies within the patient in theory, there are several technical and clinical issues that need to be solved first (see 2.2.5).

Feature selection and engineering

Since the quality of the data is essential to learn a model with high generalization performance, data preprocessing steps like data cleaning, dimensionality reduction, feature selection, and feature engineering (deriving new features from the existing ones) play a key-role in the data-centric approach.

To reduce our feature space, we have used feature selection methods within the model compared to p-value based selection methods. The horseshoe prior, which is used as prior for the coefficients in the Bayesian GLMMs in Paper 4, shrinks most of the coefficients to zero. Thereby, only few predictors (clinical variables and HLA alleles) with strong effect contribute to the model. By using the oligo kernel in Paper 1, we transform the features (envelope protein amino acids) from the input space to a feature space, where two sequences are not compared by the sequential amino acid occurrence but by the overlap of the occurrence of k-mers across the complete sequence. Across all projects, we have observed that only few positions in the viral genome confer adaptation, which agrees with the biological understanding. Thus, the usage of feature selection methods is likely to be important for these prediction tasks.

HIV-1 uses glycosylation to mask conserved parts of its envelope protein from antibody responses. Still, some bNAbs have even a glycan-specific binding site such as the V2/V3-glycan specific antibodies. Hence, the glycan shield of the viral envelope protein is often predictive for the neutralization efficacy of an antibody. While our bNAb neutralization prediction model has learned potential binding sites and potential N-linked glycosylation sites (PNGS) from the sequence without incorporating the glycosylation amount of the virus, recent models have shown that incorporating the glycosylation information as a predictive feature into the model improves prediction performance [88–90].

2.2.2. Model choice

This subsection discusses the choice of the methods used within the four different projects. Additionally, it provides a comparison with the related work at time of

development as well as recent developments in the corresponding field. Finally, an outlook is given what kind of models might be better suited to meet the objectives.

There is a plethora of machine-learning based models, each having its advantages and disadvantages. The chosen method should arise from the study design where the requirements of the model with respect to explainability, data distribution, resources at development and deployment stage are all considered and evaluated. Nevertheless, it is beneficial to compare different methods to obtain robust findings, detect potential mistakes in the implementation, or reveal unmet model assumptions.

Modeling rare events

In our projects, we often face the challenge of modeling rare events: (1) predicting neutralization resistance, and (2) analyzing X4 coreceptor usage. The unequal distribution of these variables of interest is not due to sampling error or study design, but because these samples are also rare or less frequent in the global population. For bNAb prediction, the rareness of resistant samples arises due to positive improvements in the development of very potent bNAbs leading to few observed resistant samples. Thus, the class distribution of the outcome is imbalanced. Even in the future, the class distribution is likely to stay imbalanced, since the patient selection criteria in clinical studies lead to the inclusion of patients that are sensitive to the bNAb of interest prior to treatment. The unequal distribution of HIV isolates with X4 coreceptor usages is due to the circumstance that nowadays HIV-1 patients are earlier diagnosed and treated compared to the past. Thereby, the number of R5-capable viruses is larger compared to X4-capable variants.

Imbalanced data sets represent a challenge for machine-learning approaches, since the misclassification cost for the rare samples is often higher compared to the majority class. The consequence of classifying a virus falsely as sensitive to a bNAb, when it is resistant, is the administration of an ineffective treatment and the loss of viral control. This is more harmful compared to the misclassification of a virus as resistant when it is actually sensitive. Therefore, it is fundamental that the prediction performance measure takes the different class proportions into account in contrast to evaluation metrics like the accuracy.

We used the area under the receiver operating characteristic curve (AUC) to evaluate our neutralization prediction models in Paper 1 [1]. AUC is a ranking based performance measure that sorts all predictions in increasing order of the estimated probabilities and calculates the true positive rate (TPR) and false positive rate (FPR) for each possible occurring probability decision cutoff. Consequently, the AUC can be seen as the probability that a random positive sample (class = 1) achieves a higher probability of being ranked higher than a random negative sample (class = 0). Thus, it is an aggregate measure over all ratios of TPR and FPR. By construction, AUC is aware of the (imbalanced) class distributions, since TPR adjusts the positive predicted samples against the positive class and FPR computes the ratio of samples incorrectly predicted as positive against the negative class. Recently, there has been some controversy regarding whether or not AUC is insensitive against imbalanced data. The major criticism is that if the majority class is set to the negative class, and the model has a low precision and high recall, the AUC will not reflect the poor prediction performance on the minority class in terms of precision. How well AUC reflects the prediction quality depends on the cost for false negatives (FN) and false positives (FP). A possible solution is to consider a fixed clinically approved

false discovery rate (FDR) to choose the final probability decision cutoff between the minority and majority class. The AUC alone is not very useful to select clinically relevant models, since it also covers ratios between TPR and FPR that are not clinically relevant.

If only the minority class (assigned to the positive label) is of interest, the area under the precision-recall curve (PR-AUC) has been proposed as a more suitable evaluation metric for imbalanced settings [91–94]. We used PR-AUC as an evaluation metric for the per-site adaptation models in Paper 4 [4], however, this metric also harbors some drawbacks. Since precision-based measures are based on the baseline frequency of the minority class in the training data, they are, however, not useful for comparing the prediction performance over multiple data sets with varying frequencies. To overcome this and other issues, precision-recall-gain curves [95] have been proposed as a more suitable measure.

Apart from a suitable evaluation metric, there are some machine-learning methods that have internal class-sensitive cost parameters. To predict HIV-1 neutralization resistance, we have chosen to use a cost-sensitive SVM [64], where the regularization cost parameter C is set per class to be inverse-proportional to the class distribution. SVMs are large margin classifiers that fit a linear decision hyperplane with the largest possible margin between the two classes. Using kernel functions, non-linear separable data in the input space can be transferred into a potential hyperdimensional feature space where the data can be separated with the linear hyperplane. Since the data are often not separable, there is a cost regularization parameter C (or the inverse-proportional soft-margin hyperparameter λ) that regulates the trade-off between maximizing the margin while keeping the misclassifications low. Increasing C allows for more misclassifications and vice versa. The resulting soft-margin is usually symmetrical on both classes. Given high imbalances, this setting will lead to a model, whose margin is maximized by allowing the minority class to be misclassified as majority class. By setting the class-dependent C values to the inverse of the class distribution, the boundary is pushed from the minority class to the majority class, allowing for more misclassifications of the majority class compared to the minority class. Other machine-learning methods, such as Random Forests [96] also have the ability to use adapted class-weights [97] and thus are suitable for modeling rare events.

In addition, we used stratified (nested) cross-validation [98] to select the hyperparameters of the model and to estimate the generalization error. In a stratified cross-validation, the class distribution is kept the same over all folds.

In general, there is the possibility to use various sampling techniques to overcome the imbalances in the class distribution. Since we had rather small sample sizes in our projects, we don't use any downsampling technique such as random sampling, downsampling using cluster centroids, or downsampling using ensemble methods, which reduce the size of the majority class to the size of the minority class in different ways [99–104]. Approaches that increases the minority class artificially using either over-sampling techniques (duplicating the data) (Bagging [105], XGBoost[106]) or synthesizing data by slightly perturbing the data with noise (SMOTE [107]) are prone to overfitting and have not been used in our projects as well.

Predicting neutralization susceptibility

We have chosen SVMs for modeling neutralization susceptibility in Paper 1, due to several reasons. First, SVMs can handle settings where the sample size n is smaller than the number of predictors p. Second, they have a cost-sensitive regularization parameter allowing them to handle imbalanced data. Third, using a suitable kernel, the method allows to handle non-linear relationships. Nevertheless, we have observed in our study that a random forest approach leads to similar prediction performances for this classification task.

We have chosen the oligo kernel [66] as the kernel function in the SVM. The oligo function encodes a string sequence by the occurrence of k-mers in the sequence, modeled by a Gaussian with peaks at each occurrence of the k-mer. The second parameter of the kernel is the width of the Gaussian σ allowing for a shift of the k-mer. The usage of the oligo kernel has many advantages. First, it outperformed other kernels in an additional cross-validation for the bNAb VRC-PG04. Second, we noticed that the learned kernel parameters k for the oligomer size and σ for the width of the Gaussians coincide with the size of the binding pattern of the bNAbs, and thus are biologically even more interpretable compared to a Random Forest approach. We also observed that linear models such as linear SVM or oligo kernel with a k-mer size of 1 are suitable for single binding sites of bNAbs but not for more complex binding sites. Thus, the oligo kernel is especially useful if the binding pattern is not known beforehand.

There are many established genotypic-to-phenotypic prediction methods for detecting drug-resistances in current anti-retroviral drugs [108–111] or determining the coreceptor usage[112, 113]. At the time of model development there was no established method to predict neutralization susceptibility. At that time, the research focus was rather on predicting bNAb immunogens (that would elicit bNAbs) or bNAb epitopes (since structural information was not always available) [114–121]. Neutralization susceptibility was then predicted as a by-product using the identified immunogens or epitopes. As discussed in more detail in Paper 1, these approaches are not tailored for predicting neutralization susceptibility. The major drawback is that residues outside the epitope are not considered but might be important for bNAb success. In addition, immunogens triggering the development of bNAbs might not be deterministic for the specific binding success of the bNAb. Moreover, many approaches assumed either a purely linear relationship between the changes in the viral envelope protein and the change in neutralization susceptibility, or independence of the residues in the viral protein.

Apart from the epitope and immunogen predicting approaches, only two alternative approaches existed that were directly using the viral envelope amino acid sequence for predicting neutralization resistance [89, 122]. As discussed in Paper 1, IDEPI [122] is using a linear SVM for the prediction task and only evaluating the prediction on 2F5, which is an older bNAb. The other approach [89] uses a neural network approach despite small sample sizes, which potentially leads to overfitting. Moreover, they encoded the different amino acids as integers instead of factors, which might lead to potential unintended model interpretation, since the integers have an intrinsic order which the amino acids don't have.

After the publication of our work, more neutralization susceptibility prediction methods have been published [88, 90, 123–125].

The bNAb-ReP tool [123] uses a gradient boosting machine (GBM) to predict

neutralization resistance. While the pseudovirus panel data has been used for model training, the prediction performance of the models is also assessed on the HIV-1 clinical isolates. A drawback is that the clinical data had almost no resistant samples, since the HIV-1 isolates have been prescreened using outgrowth cultures in the clinical trials. Unfortunately, the authors have also used a binary classification approach using 50 μg/mL as a cutoff to distinguish sensitive and resistant samples. As pointed out in Section 2.2.1, this is not in accordance with the latest cutoffs used in clinical trials.

While the SuperLearner approach by Margaret et al. [124] has many advantages like (1) using a regression and classification approach, (2) defining sensitivity as IC50 $\leq 1\mu$g/mL, (3) providing confidence intervals for the predictions, (4) correcting for the geographical origin of the viral strains, and (5) considering the learned feature importances of the stacked models, there are still some drawbacks. First, their model has been tailored only to the VRC01 neutralization susceptibility. Second, they use a predefined set of amino acid features to reduce the viral genetic diversity. Third, using a stacked learner like the SuperLearner makes the predictions less interpretable.

Their follow-up model SLAPNAP [126] is again a potential SuperLearner based model (though the user can choose to select a single learner) able to compute also the neutralization sensitivity to a combination of bNAbs using a Loewe additivity model [127–130] or a Bliss-Hill model [131].

Yu et al. [88] use a Bayesian SVM approach to predict neutralization susceptibility incorporating the glycan shield information as an additional predictor into the model.

Conti and Karplus [125] predict the IC50 value for bNAbs binding to the CD4-binding site using an artificial neural network. The input features are derived from a 3D-atomistic model per antibody-virus complex, which consists of the amino acid sequence of the virus and one heavy and light chain of the antibody as well as a known crystallographic structure of a related complex. While it is an advantage that a regression model and a classification model have been built (with a cutoff of IC50 <1 μg/ml), there are several potential issues with this approach. First, the robustness of the models has to be confirmed in future studies, since no cross-validation was performed. In addition, the random sampling is unaware of the differences in the distribution between resistant and susceptible strains. Second, the models rely on the availability of a crystallographic structure. Third, the models are built only for CD4-binding site bNAbs. Fourth, it seems that the samples are not completely independent of each other, since many viral sequences are the same across the different CD4 binding site bNAbs included in the models.

Recently, a multi-task deep-learning approach was presented by Dănăilă et al. [90], which adds the antibody sequence information as predictor into the model and considers all bNAbs in the CATNAP database together. This approach has several drawbacks. First they use again a binary decision cutoff of IC50 $>$ 50 μg/mL to discriminate between the susceptible and the resistant viral strains. Second, the idea to use a multi-task approach is in general appropriate to overcome the small sample size issue, however, the same viral panels have been used for the different bNAbs, such that there is no information gain from incorporating other bNAbs on the same viral strains. Third, the author state that their model lacks interpretability. Hence, it is not applicable in a clinical setting.

Due to the closer approval of bNAbs as therapeutic agents, a clinically relevant genotypic resistance prediction tool for neutralization susceptibility is still urgently needed.

Modeling antibody neutralization

In Paper 2, we have used the area under the neutralization titration curve (AUC) instead of the IC50 value to determine the neutralization capacity of an antibody. The advantage is the inclusion of samples that do not reach the IC50 value, but still show an increase in neutralization capacity. The benefit of using the area under the titration curve has been shown before [72]. For a more robust version, we computed the background noise of the assays based on the neutralization capacity for the Murine Leukemia virus, which has been used as a control virus on each assay well plate. The variation of the assays was also assessed by computing the standard deviation between replicates. Each titration AUC was normalized by the maximal possible AUC for better comparison.

Why it might be more informative to directly model the titration curve to compare the neutralization titration profiles between different time points or treatment groups, this approach harbors many caveats. Modeling the shape of the titration curve is not trivial based on the usually small sample size from a titration experiment. The curves are commonly described by modeling a monotonic sigmoidal curve with the top plateau value, the bottom plateau value, the Hill slope (the slope between the plateaus), and the IC50 value. However, it has been observed that in HIV-1 the neutralization titration curve is often not monotonic [132]. Moreover, by definition, the titration points are not independent of each other, complicating the modeling process. Finally, it is to assess if two titration profiles differ statistically. Instead of comparing the the Hill-slope or the IC50 value, it would be desirable to compare the shape of titration profiles directly allowing for some variational shift.

Handling confounding variables

There are several ways to control for confounding variables [133]. Some can be applied during study design, such as (1) randomization of the samples, (2) restricting the data set such that no unequal distributions with respect to the confounding variable exist, or (3) grouping samples that match with respect to potential confounding variable. After data collection, confounding variables can only be adjusted for by either (1) stratification of the samples with respect to all potential confounders or (2) multivariate analysis. Stratification is often not feasible due to the small sample size of the studies. In Paper 2 and Paper 3, we have used univariate linear and Bayesian regression to test whether a variable is a potential confounder for an effect of interest.

In the linear regression setting, we compared the likelihood of a model with the potential confounder as predictor with a model without the potential confounder (null model) using a likelihood ratio test.

In the Bayes regression setting, we also have built a model with the potential confounder and a corresponding null model. Then the Bayes Factor are computed instead of performing a p-value based test.

A slightly different approach is to use so-called propensity scores [134–136]. First a linear model is built to classify the individuals into the different treatment groups

based on clinical variables and potential confounders. This model predicts for each individual the probability (propensity score) to get the treatment based on their individual characteristics. In a second step, the effect of interest is compared between the different treatment groups adjusted for the estimated propensity scores.

A disadvantage of all of these approaches is the assumptions that the confounding variable has a linear relationship with the outcome. This is a very strong assumption that often does not hold.

Once a variable is detected as confounding variable, inverse-probability weighting can be used to adjust for the unequal distribution of the predictors with respect to the confounding variable. However, caution is required if some samples have extreme weights.

A complete different approach is to use causal inference to model and control for confounding variables [137]. A major advantage is that the definition of a confounder can be modeled directly in the causal graph as the common ancestor of the predictor variable and the outcome variable. Moreover, causal inference is independent of the data per definition by modeling potential interventions. Finally, there is no linear assumption between predictors and the outcome. However, the causal model is rather suited for univariate analyses with only one potential confounder at a time.

Flexible framework vs. fine-tuned models

In Paper 1 and 4, we investigated whether bNAb resistance and HLA-restricted adaptation is predictable using the viral genotype, respectively. In the current form, both prediction approaches present rather general frameworks than specific fine-tuned models. We have decided on purpose to incorporate as little prior knowledge as possible into the model apart from the viral genotype. Thereby, our models are suitable for settings where the task-specific prior knowledge is not available. In the approach in Paper 4, we included only some general prior knowledge about potential confounders and the fact that only few HLAs have influence on a viral polymorphism. Moreover, the models already suffer from a high number of parameters compared to the sample size. Therefore, it is advisable to use as few predictors as possible. Nevertheless, we have included the complete viral envelope protein and the HLA profile, respectively, to allow for an exploratory search to find also previously unknown potential predictors. This might reduce our prediction power, though we have used internal feature selection approaches like the horseshoe prior in Paper 4. Since the prediction methods do not depend on specific prior knowledge, they can be easily used to study virus-host adaptation for other viruses like influenza or SARS-CoV-2.

In both prediction approaches, we have observed that only few amino acids in the viral genome are predictive for adaptation based on our data. Thus, if the aim is to provide prediction models with a high prediction power for a well-known bNAb for example, it might be beneficial to incorporate only the found discriminative features and incorporate further prior knowledge additionally. Note that the model might have learned only few discriminative features due to the rather small sample size. Thus, further explorative approaches with more data are needed.

2.2.3. Performance metric

A rather neglected field in machine-learning is the choice and development of performance measures that are used to optimize the models. The performance metric should be carefully chosen to not only compare different methods in the training scenario but also be clinically relevant in the deployment setting.

The choice of AUC

While we have used the AUC in Paper 1 for optimizing the model parameters and comparing different methods, we argue that AUC is not a suitable choice with respect to clinical relevance. AUC considers all ratios of TPR and FPR. This is, however, not applicable for clinical settings where a certain FPR is not allowed to be exceeded. It might be beneficial to consider either a partial AUC, where the range of the AUC of interest is chosen in agreement with the clinical stakeholders. Alternatively, a clinical relevant FDR has to be chosen. Consequently, the decision cutoff is chosen in a nested cross-validation that does not exceed the selected FDR.

Two-sample hypothesis tests

Across all four projects, we have used the Wilcoxon-signed-rank test (paired data) as well as the Wilcoxon-rank-sum test (unpaired data) to test whether there is a difference with respect to a parameter of interest within two groups [138–141]. Both tests are considered to be non-parametric, since they make no hard assumption on the shape of the data distribution of the effect of interest. If the variable of interest is, however, normally distributed and the variance in both groups is the same and has the same location parameter, a t-test will have a higher power [142]. Another potential issue with the Wilcoxon-signed-rank test (paired data) as well as the Wilcoxon-rank-sum test (unpaired data) is that the tests assume that the two populations are from the same distribution under the null hypothesis. If this assumption is not met, then there are some newer alternatives, namely the Brunner-Munzel and Fligner-Policello test [143, 144]. If the spread is very different in two groups, then the Mood's median test might be a better approach [145]. While the Wilcoxon-signed-rank test (paired data) as well as the Wilcoxon-rank-sum test (unpaired data) are widely used, there is some controversy whether a bootstrap two-sample hypothesis test is less likely to overfit for small sample sizes [146]. Here, the idea is to have a process to directly sample from the null hypothesis and therefore be able to provide better confidence intervals for the variable of interest. First the variable of interest is shifted in both groups such that the mean of the variable is the same in both groups (by removing the group mean and adding the overall mean). After having now the assumption that both groups have the same mean, bootstrap samples are drawn and confidence intervals can be computed for the differences in between the two groups under the null hypothesis.

2.2.4. Trustworthiness

Until now the models and findings have been discussed with respect to data quality issues, model choice issues, as well as the issues with performance measures. Among these categories fall also trustworthy criteria like fairness, robustness, and model

transparency (explainability). In the following, the prediction models are also discussed with respect to data privacy concerns but also reproducibility and utility. Prediction models that are used in clinical routine, should also be trustworthy with respect to security aspects, however, the current models and results are not discussed with respect to that criteria here.

Data privacy

In Paper 4, we have used the patient's HLA genotypes to study the adaptation of HIV-1 to the HLA system. Due to privacy concerns, we were not allowed to publish the HLA information along with the viral sequences. Consequently, we were not able to publish our adaptation model due to potential retrieval of the HLA information from the model. In general, there are many privacy concerns with respect to sharing human genomic data [147–149]. While the concerns are legitimate and sensitive information has to be protected, it contradicts the current efforts to make data findable, accessible, interoperable, and reproducible (FAIR) [150].

In HIV-1 patients, the risk of re-identification might have severe legal consequences [151, 152]. Apart from potential stigmatization, there are still countries which criminalize HIV infection by law [153, 154].

Many data privacy concerned methods have been recently developed and investigated for their use in the healthcare sector, such as differential privacy (DP), fully homomorphic encryption (FHE), secure multi-party computation (SMPC), or federated learning (FL) [155–157]. Nevertheless, more research needs to be performed to ensure the protection of sensitive data by these methods while upholding the utility and quality of the downstream machine-learning models and data analysis.

In a setting where the data cannot be shared due to privacy concerns, a FL approach might be useful. In a FL setting, many clients train a machine-learning model locally conducted by a centralized server. While the data never leaves the local client, learned parameters and model characteristics are propagated to the central server. The utility and potential issues with FL are currently under research [158, 159]. Since the centralized orchestration of the FL approach is seen problematic, a blockchain-based swarm-learning approach has been recently proposed to counteract this issue [160]. Another unexplored issue with federated and swarm learning is the propagation and adjustment of confounding factors.

A complete different approach is the attempt to create and use synthetic data based on the sampling data instead of sharing the sampling data. More research has to be done, however, to assess the utility of synthesized data with respect to their data quality (representing the underlying distribution, capturing the variance, and the power to generalize) but also to the privacy concerns - if re-identification or reconstruction of the original data is possible [161].

While so far privacy concerns were related to sharing human genetic material, the current SARS-CoV-2 pandemic has drawn attention to the privacy concern related to viral genomic sequences from human patients [162–166]. HIV genomic sequences might be used to reconstruct the host HLA genotype based on the HLA-restricted escape mutations in the viral sequence.

Another problem is that contextual information such as country, age, and ethnicity, and sample date is also regarded as sensitive information enabling a potential re-identification of the individual. Nevertheless, we have shown in our research that

this information harbors potential confounding effects that need to be accounted for.

In the future, newer technologies might allow to re-identify sensitive information from already published data that has not been considered sensitive before. Thus, it is advisable to further put effort in developing privacy-preserving machine-learning techniques and best practices to circumvent the privacy concerns beforehand. Upcoming regulations based on the European Data Act from February 2022 [58] will likely drive the transformation to a privacy-preserving modeling and in general to more trustworthy AI models.

Utility

Apart from criteria like fairness, robustness, transparency, data privacy or clinical relevance, prediction models need to be reproducible and show their usefulness after development.

We have performed several efforts to tackle the reproducibility aspect. We have used the Snakemake Workflow Manager [167], which facilitates the maintaining of the complete model pipeline (from data preprocessing to model training to model deployment). The Anaconda Software Distribution [168] allows to control the dependencies of different programs and packages in the running environment of the model. Thereby, it facilitates the usage of the model by other users. We also used code versioning and code repositories where possible using github/gitlab or SVN repositories. Manual steps during data preprocessing like retrieval of reference sequences, sequence alignments, or sequence subtyping using web tools have been automatized as well. While we cannot provide the models for Paper 4, due to privacy concerns, we provide a minimal data set that allows to reproduce all findings, figures, and tables of the manuscript. In addition, we provide the coding pipeline as far as it is in accordance with the data privacy policies. For both genotypic prediction tools, we also provide the top discriminant features within the manuscript.

In terms of utility more steps need to be performed, however, to create clinically relevant prediction models. For the bNAb prediction model, it is advisable to create an automatic update of the models using the CATNAP database. Thereby, any occurring data shifts can be identified. For settings where the code or the pipeline cannot be shared, either the prediction models must be hosted on a publicly available server or a federated-learning approach must be enabled to provide the reuse of the data.

2.2.5. Future trends

Neutralizing antiretroviral drug components

Besides bNAbs, single-chain variable fragments (scFvs) [169–172] antibodies are currently investigated as novel antiretroviral drug components targeting the same epitopes as bNAbs. ScFvs are composed of only the variable regions of the heavy and light chains of antibodies, thus containing the complementary antigen-binding site (Fab) of antibodies. The smaller size of scFvs compared to IgGs facilitates the absorption of the scFvs in mucosal tissue, where most HIV infections occur. In addition, scFvs might also prevent viral cell-to-cell transmission [173]. A drawback

is that by lacking the Fc region of the heavy chains, the scFvs do not activate further immune cells and have a shorter half-life compared to IgGs. In addition, it is possible to design the scFv as bi- or trispecific [174–178] targeting multiple binding sites of distinct bNAbs with only one molecule. Since we have not used the bNAb structure nor the bNAb sequence in our models, our neutralization prediction tool (P1) [1] can be easily extended to other bNAb alternatives like scFvs.

Usage of NGS data

While we sequenced the viral *env* and *gag* gene using next generation sequencing (NGS) methods for studying the adaptation of HIV-1 to the immune system (P4), we have used the simulated consensus Sanger sequence for further analysis. NGS are cost-effective for massive parallel sequencing, and in contrast to Sanger sequencing, are able to detect low-frequency fragments, which is important to assess the minority variants within the quasispecies of a patient for drug resistance testing. The frequency-cutoff at which the detected drug resistant variant are of clinical relevance need to be determined yet [179–181]. In addition, the cutoff might differ for different drug-classes, drugs, or distinct mutations. Moreover, output from NGS methods are hard to compare since standardized protocols across different platforms are missing and unified quality assurances need to be determined. Once the standardization and threshold issues are tackled, it is likely that NGS-based drug testing will become a standard procedure. Until then, Sanger sequencing is still seen as gold standard. Hence, the generation of simulated Sanger consensus sequences based on NGS data is currently a common procedure.

Despite the current described issues with directly using NGS data, there are still ways to include some additional information from the reads to model the quasispecies within HIV-1 patients. A possibility is to compute the genetic diversity based on the reads and include this information as a new feature in the model [182–184]. Another option is to use the frequency counts of the nucleotides over the reads as additional information over the actual sequence distribution. This is, for example, used in the geno2pheno[ngs-freq] tool [185] that allows to predict drug-resistances based on the frequency files. Still these methods require the selection of a prevalence cutoff.

3. Conclusion and perspective

Section 3.1 concludes the thesis by providing a summary of the made contributions across all four studies. In Section 3.2, a personal perspective is given on how current genotypic-based adaptation prediction models need to be improved based on the discussion in Section 2.2.

3.1. Conclusion

The four studies included into this thesis contribute on the one hand to our understanding of HIV-1 adaptation to bNAbs and to the host immune system. On the other hand, they introduce fast genotypic-based prediction models to detect HIV-1 adaptation, thereby increasing our means to control HIV-1 infections. Three studies present methods and findings that contribute to the advancement of bNAb therapy from benchside to bedside (P1 - P3). The last study (P4) introduces a novel method to predict HIV-1 adaptation to the HLA system and links HIV-1 adaptation to the HLA system to the coreceptor usage.

In the first study [1] (see Section 2.1.1), we introduce well-performing SVM-based models that predict HIV neutralization susceptibility to 11 bNAbs based only on the viral envelope protein. Due to the rapid evolution of the virus within the patient, this rapid resistance testing is essential for the selection of an effective personalized bNAb treatment in the future. While our models predict the neutralization only to a single bNAb at a time, they can be used as a foundation for a support-decision tool to predict a bNAb combination therapy using an additive model or the Bliss-Hill model approach. By providing not only the learned discriminant features, but also the contribution of the query sequence composition to the classification outcome, we provide well explained machine-learning based models as required for a clinical application. On top, our findings reveal that neutralization capacity is influenced by the coreceptor usage. More precisely, we observe that the neutralization capacity of the bNAbs PGT121 and PGT128 has an R5-bias as already discovered for bNAbs PG9 and PG16 [75]. The general framework to predict neutralization susceptibility is not only useful for the clinical routine, but can also support current clinical trials in their patient selection process. As current vaccine efforts have the goal to elicit bNAbs, the learned discriminative features might guide the selection of suitable immunogens.

While the support decision framework is essential for the administration of the bNAb therapy, first the characteristics and efficacy of bNAb-based therapy have to be proven. Therefore, we use statistical learning methods to support two clinical trials.

In the second study (P2) [2] (see Section 2.1.2) the usage of a noise-corrected version of the area under the neutralization titration curve (AUC) enabled the inclusion of patients with low neutralization capacity not reaching the IC50 value, which is usually taken as neutralization measure. Using the developed measure, we

observe that a monotherapy with 3BNC117 statistically increases the neutralization capacity of the patient compared to a control group. In addition, we ensure that the observed increase in neutralization capacity is irrespective of other potential confounding factors such as initial neutralization capacity, age, sex and other clinical variables by performing a linear and Bayesian regression analysis. While further studies are required to validate the impact of a bNAb therapy on the viral reservoirs, the findings advance our understanding of the mode of action of a potential bNAb therapy.

In the third study (P3) [3] (see Section 2.1.3), we use several survival regression techniques to model viral rebound after analytical treatment interruption. Thereby, we observe that a monotherapy of 3BNC117 statistically delays viral rebound in contrast to a historical control group even after adjustment for the identified confounding factors *age* and *years on ART* with a weighted log-rank approach [76]. This finding is the prerequisite of the inclusion of 3BNC117 in a potential combination therapy that is currently further investigated [186].

In the last study (P4) [4] (see Section 2.1.1), we introduce a novel prediction method that jointly models the adaptation of HIV-1 to the HLA I and HLA II profile of the patient. Using our models, we observe that HIV-1 adaptation to the HLA system differs depending on the coreceptor usage. The relationship of HIV-1 adaptation and the coreceptor usage has been neglected so far. The machine-learning based adaptation model is transparent with respect to the learned polymorphims in the viral p24 protein that are more likely under HLA pressure. In addition, we provide a sequence-based logo that explains for each query sequence which amino acids are likely to be under HLA pressure and which are not according to the model. As a correct determination of the coreceptor usage is essential for the administration of CCR5 coreceptor antagonists, an estimation of the viral adaptation to the HLA system might be an additional useful predictor for coreceptor usage or even the coreceptor switch.

Due to the general and flexible model definition of both genotypic-based prediction methods (P1 and P4), they are easily extendable to other related tasks, i.e., to predict neutralization susceptibility to (1) other bNAb alternatives such as scFvs, (2) to predict neutralization to combinations of bNAbs, (3) to predict the combined HLA adaptation for further viral proteins such as Nef and Pol, or (4) to apply the approach to other less-prevalent subtypes. While the studies have been concerned with the adaptation of HIV-1 to the adaptive host immune system responses, both approaches can be used to study any virus-host adaptation, i.e. for viruses like influenza or SARS-CoV-2. The major strength is however their usage in settings where little prior information is available such as neutralization prediction for newly derived bNAbs, or the adaptation to newly emerging viruses such as in the current ongoing SARS-CoV-2 pandemic, where no prior knowledge exists about HLA footprints nor large cohorts to identify them.

3.2. Perspective

In this thesis, we have presented several novel and existing methods that advance our understanding and assessment of HIV-1 adaptation towards the adaptive immune system components. Due to small sample sizes across all our studies, further studies are required to confirm our findings on larger data sets. Moreover, further biologi-

cal knock-out experiments are required to validate whether the identified unknown discriminant features truly impact HIV-1 adaptation. The clinical relevance for the two novel support-decision frameworks (P1 and P4) need to be further improved by considering the aspects of trustworthy artificial intelligence (AI) models and the requirements of clinical practice.

Currently none of the existing neutralization prediction methods is deployed in a clinical setting or for clinical trials. To prevent a distribution shift between training and deployment setting, the support-decision tool needs to be trained and evaluated not only on pseudovirus panel data but also on HIV-1 clinical isolates. Then, it is also of interest to analyze if host factors such as the immune status also play a role on the neutralization capacity of the bNAbs. The evaluation metric to optimize and evaluate the models needs to be not only class-sensitive but the misclassification costs should match the clinical requirements. Further, the usage of fixed cutoffs or thresholds should be avoided. Apart from choosing a machine-learning method that is explainable and produces explainable outcomes, the support-decision tool should provide confidence intervals or other means to quantify the expected uncertainty with regard to new unseen samples, or certain value ranges of the features, or the estimated neutralization susceptibility. Alignment-free approaches might also be beneficial, since currently insertions in the query sequence that are not present in the training data are discarded for the prediction in the alignment process of the query sequence to the training sequences. Recent studies have shown that an estimate of the glycosylation of the envelope protein improves performance and should be incorporated as a predictor into the model [88–90]. In order to achieve fair and robust models, it is important to carefully explore the training data and perform a confounder analysis on potential factors such as coreceptor usage, country, and subtype. If possible, the analysis should be extended by host factors such as sex, age, ethnicity, and clinical variables, i.e., CD4 count, CD8 count, coinfections, infection year, and years on ART. Another source of bias lies in the sample generation process, i.e., sample year, sample data, sample location etc. Neutralization susceptibility will also likely be decreasing in the sample collection over time due to higher resolution of the assays but also more potent bNAbs or bNAb combinations. Thus, time is a potential confounder in databases that has to be considered. Due to the highly overparametrized model setting, it is important that the choice of features are discussed with the clinicians with respect to clinical relevance and the modelers with respect to data quality. To maintain its utility, it would be beneficial if the tool is retrained automatically on the CATNAP database. Thereby, not only are the models trained on the largest available portion of data, but also distribution shifts or change of use cases can be detected. While there are several machine-learning methods that have shown to be appropriate for this model task (SVMs, Bayesian SVMs, Random Forests, GBM), it is of interest to investigate, where these methods differ with respect to the misclassified samples. The benefit of a generative model is that samples from the learned generative model can be compared to the training data and to the expected data distribution at deployment stage. Finally, a data-privacy preserving approach would facilitate the exchange of clinical data that is needed to achieve more accurate and robust prediction models. While the approach by Yu et al. [88] is not meeting all of the mentioned requirements, it is currently the most promising approach in my personal opinion.

Our HIV immunoadaptation study (P4) had several drawbacks. First, there is no

ground truth for the adaptation score. Apart from a required in-vitro experimental validation of the estimated adapatation score, it is possible to perform further in-silico validatation experiments. Existing tools to predict the HLA binding of peptides can be used to validate whether the peptides of a virus with high predicted adaptation are also less likely to be bound by the HLA molecule [187, 188]. Second, the data is a cross-sectional snapshot of the viral population and the immune system. Thus, it does not allow to observe whether increasing adaptation is an indicator for a coreceptor switch. Therefore, a longitudinal observation of the patients would be beneficial. Third, the duration of the infection might be a confounder for adaptation and should be measured and corrected for. Fourth, the distribution of CD4 counts and coreceptor usage should be equally balanced to study whether viral adaptation in patients with high to intermediate CD4 counts play a major role for the coreceptor switch. Note that having longitudinal data does not ensure to capture a coreceptor switch. Still the longitudinal data is useful to further validate our adaptation score. It can be used to investigate whether a potential change of adaptation over time is in accordance with emerging HLA-restricted polymorphisms. Fifth, we have not incorporated dependencies such as the linkage disequilibrium between some HLA alleles as well as the shared binding sites of several HLA alleles that are captured in the HLA supertype information. Finally, a data-privacy preserving approach is recommended to improve the utility of our framework.

The current data-centric trend as well as the upcoming laws with regard to data protection and trustworthiness of machine-learning models will likely increase the awareness of potential risks in machine-learning models, especially for the healthcare sector. Hopefully, this will push the research community to focus even more on study design, data quality, and potential uncertainties within the training data but also the real-world application requirements. Data-privacy preserving techniques might facilitate the usage of existing data among different institutions and clinics increasing the sample size of the studies and reducing the waste of resources. The resulting transparency of the model risks might also facilitate the assessment and acceptance by clinical practitioners leading to a faster deployment of machine-learning based support-decision tools in clinical practice.

A. Scientific papers

A.1. Paper 1 - bNAb resistance study

Copyright clearance

PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Prediction of HIV-1 sensitivity to broadly neutralizing antibodies shows a trend towards resistance over time

Anna Hake[1]*, **Nico Pfeifer**[1,2,3]*

1 Department Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany, **2** Methods in Medical Informatics, Department of Computer Science, University of Tübingen, Germany, **3** Medical Faculty, University of Tübingen, Tübingen, Germany

* anna.hake@mpi-inf.mpg.de (AH); nico.pfeifer@mpi-inf.mpg.de (NP)

OPEN ACCESS

Citation: Hake A, Pfeifer N (2017) Prediction of HIV-1 sensitivity to broadly neutralizing antibodies shows a trend towards resistance over time. PLoS Comput Biol 13(10): e1005789. https://doi.org/10.1371/journal.pcbi.1005789

Editor: Alice Carolyn McHardy, Helmholtz-Zentrum fur Infektionsforschung GmbH, GERMANY

Received: March 20, 2017

Accepted: September 22, 2017

Published: October 24, 2017

Data Availability Statement: Most relevant data are within the paper and its Supporting Information files. At https://github.com/annahake/g2p-bnab, we additionally provide the computed kernels for the final models as well as the resampling instance for the 10 runs of stratified 5-fold cross-validation.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Treatment with broadly neutralizing antibodies (bNAbs) has proven effective against HIV-1 infections in humanized mice, non-human primates, and humans. Due to the high mutation rate of HIV-1, resistance testing of the patient's viral strains to the bNAbs is still inevitable. So far, bNAb resistance can only be tested in expensive and time-consuming neutralization experiments. Here, we introduce well-performing computational models that predict the neutralization response of HIV-1 to bNAbs given only the envelope sequence of the virus. Using non-linear support vector machines based on a string kernel, the models learnt even the important binding sites of bNAbs with more complex epitopes, i.e., the CD4 binding site targeting bNAbs, proving thereby the biological relevance of the models. To increase the interpretability of the models, we additionally provide a new kind of motif logo for each query sequence, visualizing those residues of the test sequence that influenced the prediction outcome the most. Moreover, we predicted the neutralization sensitivity of around 34,000 HIV-1 samples from different time points to a broad range of bNAbs, enabling the first analysis of HIV resistance to bNAbs on a global scale. The analysis showed for many of the bNAbs a trend towards antibody resistance over time, which had previously only been discovered for a small non-representative subset of the global HIV-1 population.

Author summary

Several sequence-based approaches exist to predict the epitope of broadly neutralizing antibodies (bNAbs) against HIV based on the correlation between variation in the viral sequence and neutralization response to the antibody. Though the potential epitope sites can be used to predict the neutralization response, the methods are not optimized for the task, using additional structural information, additional preselection steps to identify the epitope sites, and assuming independence and/or only linear relationship between the potential sites and the neutralization response. To model also the neutralization response to bNAbs with more complex binding sites, including for example several non-consecutive residues or accompanying conformational changes, we used non-linear, multivariate

machine learning techniques. Though we used only the viral sequence information, the models learnt the corresponding binding sites of the bNAbs. In general only few residues were learnt to be responsible for a change in neutralization response, which can additionally reduce the sequencing cost for application in clinical routine. We propose our tailored models to aid the patient selection process for current clinical trials for bNAb immunotherapy, but also as a basis to predict the best combinations of bNAbs, which will be required for routine clinical practice in the future.

Introduction

With around 36.7 million people living with HIV in 2015 and an incidence rate of around 2.1 million each year [1], infections with HIV continue to be a major global health issue. However, despite more than three decades of research, there is neither a vaccine against nor a cure available for infection with HIV-1. HIV-1 infected patients are usually treated with a highly active antiretroviral therapy (ART). ART suppresses the replication of the active virus, but it is not capable of eliminating viral reservoirs and thus clearing the infection. To reduce the emergence of drug-resistant viruses, ART usually consists of a combination of three or more drugs from at least two different drug classes. In total, there are six different drug classes, which differ in their mode of interference with the HIV-1 life cycle, resulting in more than 20 available antiretroviral drugs. A change of the drug regimen is still often required, due to emerging drug resistances or side-effects. Since lifelong treatment is inevitable, for some patients no efficient drug regimens might be left eventually. Hence, there is still a high demand for drugs with new targets [2].

A currently investigated treatment option is the passive transfer of a combination of broadly neutralizing antibodies (bNAbs) to HIV-1 patients. Upon the advent of new single-cell antibody cloning techniques [3–5] and followed structure-based rational design approaches [6], an abundance of these new bNAbs has been isolated and their higher neutralization potency and breadth have been shown in several studies [6–10]. The potency of an antibody is defined as the antibody concentration needed to inhibit HIV-1 infectivity by 50% (IC50) or 80% (IC80), while the neutralization breadth of an antibody is measured by the ability of the antibody to neutralize viruses from different subtypes. The latter characteristic is very important in the case of HIV-1 due to its high molecular diversity within a patient but also within a population.

The sole target of these neutralizing antibodies is the viral envelope glycoprotein, the so-called envelope spike, on the surface of the virus. The surface of the virus itself is made of host-lipids and is therefore undetectable by the immune system. Each spike consists of a trimeric heterodimer of two viral envelope glycoproteins, gp120 and gp41, which are cleaved from the envelope glycoprotein, gp160. While gp41 mediates host cell fusion, gp120 is essential for cell entry [11]. By successful binding of a neutralizing antibody to a spike, a chain reaction is initiated by the host immune system that eventually leads to the elimination of the virus.

So far, there are five known sites on the envelope glycoprotein, which are targeted by a variety of bNAbs (given in brackets): on gp120 the CD4 binding site (e.g., VRC01, VRC-PG04, 3BNC117, NIH45-46) [9, 12–14], the V1/V2 region (e.g., PG9 and PG16) [7, 8, 15–17], and the V3 loop (e.g., PGT128, PGT121, 10-996, 10-1074) [8, 10, 18–21]; the membrane proximal external region (MPER) on gp41 (e.g., 10E8) [22–25]; and a newly identified site comprising parts of gp41 and gp120 (e.g., 35O22) [26]. Since the specific binding sites of bNAbs, so-called epitopes, on the envelope protein are not similarly accessed by any available drug, a therapy

with bNAbs would offer a new effective treatment option for patients with resistance to all current therapies or might boost existing therapy combinations with few active drugs [27]. The efficacy of a treatment with a combination of these broad and potent neutralizing antibodies has been first shown in HIV-1 infected humanized mice [28, 29] and non-human primates [30]. Tolerance and safety of the bNAbs VRC01 [31] and 3BNC117 [32] have been shown in phase 1 clinical trials in HIV-1 infected humans, where for 3BNC117 also the effective suppression of viremia could be observed. In addition, recent studies have shown that antiretroviral therapy with only one bNAb (3BNC117) is able to enhance the host immune response against HIV-1 [33] and leads to a significant delay of viral rebound after treatment interruption [34]. In contrast to ART, which usually requires a daily intake of the drugs, bNAbs have a longer half-life time, being able to control the viral load for more than 28 days in humans after administration [32]. High genetic variation of the viral envelope glycoproteins together with a glycan shielding of more conserved regions on the envelope often allow the virus to escape immune recognition [35]. Thus, for treatment success, neutralization resistances of the patient's viral strains to the given bNAbs must be detected beforehand. Up to now, the neutralization sensitivity of a virus to an antibody can only be determined in time-consuming and expensive neutralization assays.

To ensure a routine clinical practice, these tests have to be more rapid and cost-effective. This can be achieved, for example, by developing a genetic resistance test, coupled with a resistance prediction method similar to current decision support for ART treatment against HIV [36]. Since the envelope spike is the sole target of bNAbs, it is sufficient to consider the changes in the genetic composition of the viral envelope glycoproteins associated with changes in neutralization sensitivity of the virus.

So far, the neutralization together with the genetic information has been mainly used to determine potential epitopes of bNAbs or to identify immunogens to elicit bNAbs. The aim of neutralization-based epitope prediction models is to learn potential epitopes or patches of the bNAb in the amino acid sequence of the envelope protein. There are approaches using only the neutralization information [37–41] or including structural information [41, 42]. Changes in the amino acid composition of the epitopes are assumed to be associated with a change in neutralization sensitivity and thus can be learned from neutralization activity information. As a consequence, the model learns potential sites instead of predicting neutralization sensitivity. Nevertheless, some of the models, or more precisely the learnt sites, have been used to predict the neutralization activity for validation purpose. Unfortunately, the performance might be overoptimistic if the same data is used for learning the sites and the prediction task [42]. Another application is the identification of immunogens to elicit bNAbs. Therefore, Gnanakaran et al [43] compared the viral sequences of HIV-1 infected individuals with and without a broad and potent antibody response, hypothesizing that shared features among the viral sequences in individuals eliciting bNAbs might be potential immunogens. Shared features have been learnt using conditional mutual information together with an ensemble learning technique using classification trees. Similar to the above approaches, the identified features have been validated by predicting the neutralization sensitivity. An overview of a variety of computational approaches for epitope vaccine design is given by He et al. [44].

Recently, an artificial neural network approach has been proposed to directly model the IC50 value based on the envelope sequence information [45]. For this, the amino acids were mapped to integers. However, the authors modeled each position in the sequences as a continuous variable instead of a categorical one, which leads to a different interpretation of changes between different amino acids. In addition, only the performance of the older bNAb 2F5 was provided. IDEPI [46] is a very generic framework that, among other features, models the neutralization sensitivity of the virus to bNAbs using a linear support vector machine (SVM) and

the envelope sequence of the virus. The above presented models have several shortcomings. First, potential epitopes can be poor immunogens. Second, sites outside the epitope can have an influence on the binding success of a bNAb as well, and thus also have an influence on the neutralization sensitivity. Structural information and other prior information about the binding sites might not be available for newly identified bNAbs. Most methods assume a linear relationship between changes in the amino acid composition and neutralization sensitivity on the one hand [37, 46] and the independence of the epitope sites on the other [37]. This assumption might not hold for bNAbs targeting a more complex binding site. Another important point involves the handling of amino acid positions in the variable regions of the envelope protein. Though the variable regions are hard to align, they are also the regions where resistance mutations are likely to appear and thus these sites should not be dropped from the analysis [43].

In this study, we present prediction models for 11 different bNAbs (VRC01, VRC-PG04, 3BNC117, NIH45-46, PG9, PG16, PGT121, PGT128, 10-996, 10-1074, and 35O22) that learnt discriminant *signals* (amino acids or patterns of amino acids) in the genetic sequence of the envelope glycoprotein gp160 (*envelope sequence*), which influence the neutralization sensitivity to the particular antibody. To learn the neutralization susceptibility of HIV-1 strains to bNAbs, we trained our prediction models on data from three previously published neutralization assays [10, 26, 47]. Depending on the neutralization assay, IC50 titers for 115 to 220 HIV-1 isolates were available for each of the bNAbs. Following neutralization assay protocols, we used an IC50 value above 50 μg/mL as a threshold to determine neutralization resistance of a virus to a particular antibody. Based on the available IC50 titers for the HIV-1 isolates, the corresponding envelope sequences, and the threshold, we built binary classifiers with non-linear support vector machines (SVM) and string kernels to distinguish between HIV-1 resistance and susceptibility to a bNAb. As non-linear prediction models are often seen as black boxes, we trace back what each classifier learnt from the data and show that many of the learnt discriminant signals are known to play an important role for the binding success of the antibody. For a better interpretation of the classification decision (resistant or susceptible), we provide a new way to produce motif logos that illustrate which and up to what extent amino acids in the tested sequence contributed to the particular classification result. Though we use the complete envelope sequence information, we show that only a few signals are important for the classification outcome and that models based only on these signals achieve comparable prediction power.

To study the evolution of HIV-1 resistance to bNAbs, we additionally built regression models using support vector regression that directly predict the IC50 value from the envelope sequence of the virus. With these models we analyzed the neutralization sensitivity of HIV-1 to the considered 11 bNAbs for around 34,000 HIV-1 samples of different subtypes over a time period of more than 30 years from the Los Alamos HIV sequence database [48]. Thereby, we could not only confirm previous, experimental results, showing that there is a trend towards bNAb HIV-1 resistance over time in the subtype B population of HIV-1 on a much larger and more diverse data set, but for the first time, the trend could also be observed for the global HIV-1 population—a scale-up that would be very expensive in an experimental setting.

A preliminary version of this study [49] has been published as a preprint.

Results and discussion

Prediction performance

Accurate prediction of bNAb resistance from the genetic sequence of the envelope protein of HIV-1. We used support vector machine (SVM) models to build our prediction

models. A crucial step in building SVM models is the choice of the kernel that encodes the similarity structure in the input data. Upon performance comparison between different kernels (see S1 Table), the oligo kernel was selected for all bNAbs to predict the neutralization susceptibility to each bNAb for new viral strains. The idea of the oligo kernel is to define the similarity between two sequences x and x' of same length L by the similarity of the co-occurrences of their substrings (oligomers) of length l with $1 \leq l \leq L$ within a certain distance (controlled by the width parameter σ^2). Fig 1 shows the prediction performance of each of the 11 classifiers measured as the area under the ROC curve (AUC). The prediction performance was assessed in 10 runs of a stratified 5-fold nested cross-validation in the kernel comparison step. All 11 classifiers are better than a random classifier (dashed line) and have good performances, up to 0.84 AUC for the V3 loop targeting bNAbs. The prediction performances of the regression models are provided in S1 Fig. To determine the best parameter setting for each bNAb prediction model, we performed an additional 5-fold cross-validation.

Comparison to other machine learning approaches. Due to the large number of features (the length of the envelope amino acid sequence) compared to the small number of samples, we chose SVMs to build our models, which are known to generalize well for these kind of prediction problems. Additionally, we compared our final SVM models (based on the oligo kernel) to a selection of other machine learning approaches: random forests, SVM using a linear kernel, a neural network, and a logistic regression with lasso regularization (see Methods for details). Overall, only the random forest approach and our model performed well for all 11 bNAbs while not being significantly different performance wise. Similar to the other investigated kernels, the linear kernel had worse performance for the VRC-PG04 bNAb compared to the oligo kernel or the random forest approach. The prediction performances are presented in S6 Fig and S5 Table. There are a plethora of machine learning approaches that could be used to tackle the here discussed classification task. Thus, we do not claim that there cannot be a better method than SVMs based on the oligo kernel. From our analysis, it seems that for bNAbs that need a single specific amino acid for a successful binding such as the V3-loop or V1/V2-loop targeting bNAbs, simpler models will perform equally well as the oligo kernel approach. Depending on the learnt hyperparameters, the oligo kernel however can also capture more complex l-mers, an advantage if the binding site pattern of the bNAb is not known beforehand.

Model reliability and user features

Learnt hyperparameter of prediction models agree with binding patterns of bNAbs. Table 1 presents the final parameters settings for the classifiers for the bNAbs PG9, PG16, 10-669, 10-1074, PGT121, VRC01, and VRC-PG04 fitted by a stratified 5-fold cross-validation. For the PGT121 and VRC-PG04 classifier an l-mer of length 6 led to the best performance whereas the l-mer length for the other antibodies was comparatively small (2-mers for VRC01 and single positions for the remaining antibodies). The length differences of the l-mers for different epitope classes supports the knowledge gained from experimental findings. For the N-glycan dependent antibodies, a single glycan site is the most important residue for successful binding. The N332-linked (V3 loop directed) antibodies PGT121, 10-1074, and 10-996 need in the first instance an asparagine at position 332 for successful binding [16]. The N160-linked antibodies PG9 and PG16 bind in a hammerhead-like way to the virus, building contacts with two glycans (160 and 156 or 171) [15]. For the CD4 binding site (CD4bs), which forms a cavity, it is only known that it is sterically not easy to bind to for antibodies [50]. Longer l-mers led to the best prediction results for the CD4bs classifiers, which is likely due to the fact that the CD4bs-directed bNAbs target a larger epitope compared to the other bNAbs.

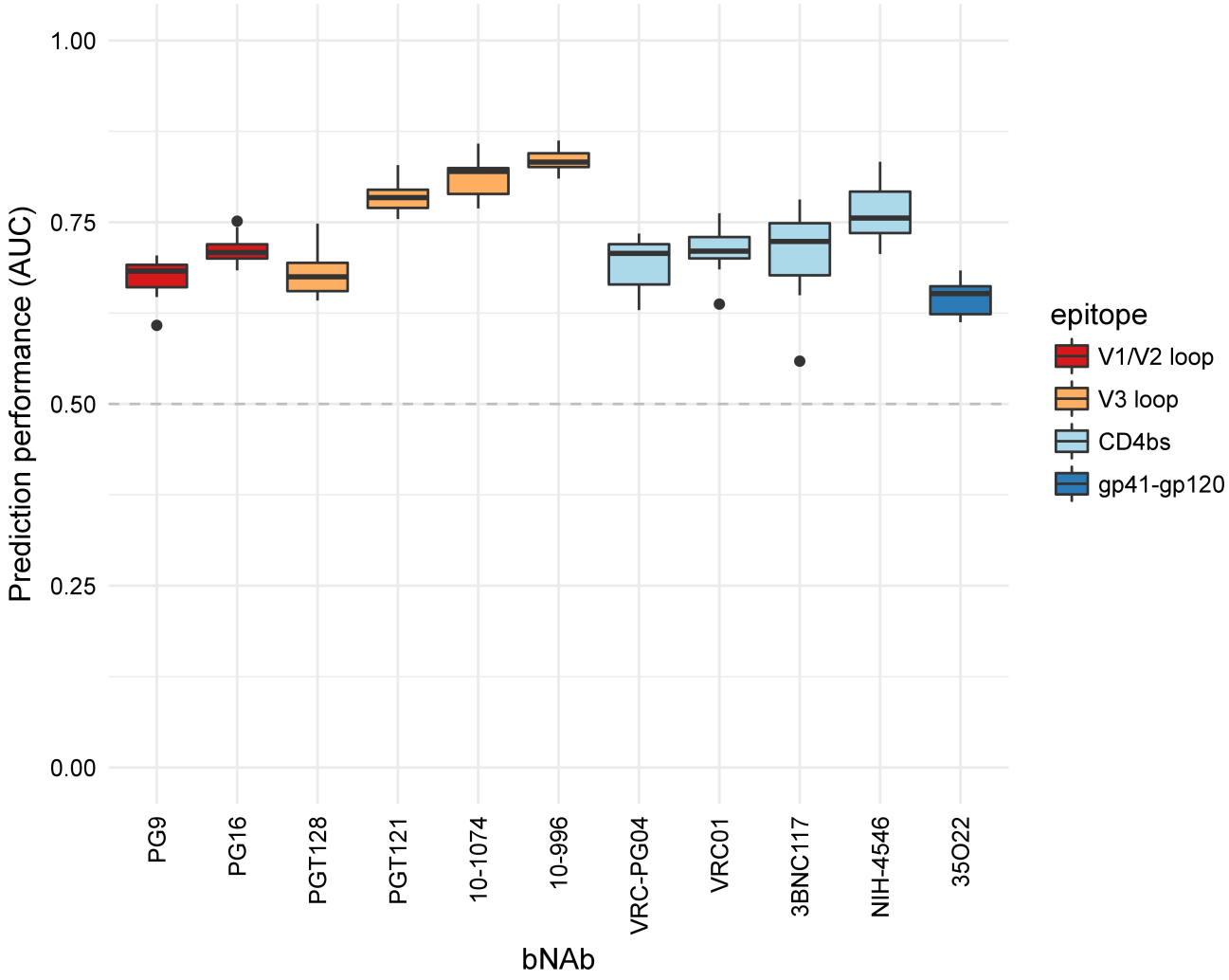


Fig 1. Prediction performance of the classifiers. The AUC performances for each bNAb classifier using the oligo kernel were determined by 10 runs of a stratified 5-fold nested cross-validation. On the x-axis, the different classifiers are presented, named according to the bNAb they are trained on. The colors of the boxes refer to the epitope category of the corresponding bNAb. The prediction performance of a random classifier is depicted by the gray dashed line.

https://doi.org/10.1371/journal.pcbi.1005789.g001

Table 1. Final parameter settings for the oligo kernel classifiers for each bNAb.

Epitope	bNAb	l	width
V1/V2 Loop	PG9	1	1
	PG16	1	0.4
V3 Loop	PGT121	6	1.6
	10-996	1	2.6
	10-1074	1	1.6
CD4bs	VRC01	2	3.6
	VRC-PG04	6	20

The parameter *l* denotes the size of the *l*-mer and the parameter *width* (σ^2) the allowed positional uncertainty of the kernel.

https://doi.org/10.1371/journal.pcbi.1005789.t001

Classifiers learnt important binding sites. In general, the learnt signals of a non-linear kernel-based SVM classifier can be traced back, if the kernel incorporates positional information such as the weighted degree kernel with shifts (WDKS) [51] or the oligo kernel [52]. By construction of the oligo kernel (see Methods), it is possible to retrieve the learnt weight of all occurring oligomers at each position in the sequence to the classifier.

Considering the 15% strongest learnt signals for each classifier, we found that several amino acids (*residues*) of the envelope protein were learnt by the classifiers to influence neutralization resistance or susceptibility, which are also supported by literature [39, 53]. In Table 2 we present the learnt signals of the classifiers exemplarily for the bNAbs PG9, PG16, 10-669, 10-1074, PGT121, VRC01, and VRC-PG04 that are supported by previous studies.

Most of the found discriminant signals for the N-glycan dependent antibodies, that is, for the V1/V2 loop and V3 loop directed antibodies, contain the amino acids asparagine (N), serine (S) and threonine (T). These amino acids are also part of the pattern N-X-[S or T], which defines potential N-glycosylation sites [54]. The classifiers for the CD4bs antibodies identified known required residues for CD4-binding as reported in [53]. The fact that all classifiers learnt some known discriminant position, further support the reliability of the prediction models in addition to the provided prediction performances. Additionally to the already known epitope sites, we found further discriminant residues whose role needs to be validated in knock-out experiments and might be interesting for follow-up structural studies (see S7 Table for a complete list of 1% discriminant signals).

Motif logo improves classifier interpretability. To improve the interpretability of the classification decision, we show how to produce for each classification of a test envelope sequence a motif logo—a representation of the test sequence—that displays those residues in the test sequence that contributed the most to the classification result. Using the available kernel feature representation of the oligo kernel, it is possible to retrieve the contribution of each residue of the test sequence to the classification. As the envelope glycoprotein consists of around 800 amino acids, visualizing the contribution of all amino acids to the classification would not be very informative. Instead, since the prediction performance of classifiers based only on the strongest p% signals with $p \in \{1, 3, 5, 7, 10, 15, 20, 25\}$ performed not significantly worse than the classifiers based on the complete envelope sequences (see S4 Table), we present only the contribution of the strongest signals in the motif logo.

For demonstration purposes, we retrieved several HIV-1 envelope sequences from the Los Alamos HIV sequence database [48] serving as test input for the classifiers. In Fig 2 we present the motif logo for the test sequence with the GenBank ID HM469973, which was classified by the PG9 classifier as susceptible, using the strongest 5% learnt discriminant signals to the

Table 2. Learnt discriminant signals by each bNAb classifier that are supported by literature.

bNAb	susceptible	resistant
PG9	N160, N301, S393, S613, K168, K169, K171	N624, D187
PG16	N136, N141, N160, N186, N234, N289, N356 S393, K169, K171, D167, T138	N230
VRC01	N186, N276, N279, N280, G459, K232	
VRC-PG04	N186, N276, N279, N280, G459, K232, R456, D368	
10-996	N332, S334	N334
10-1074	N332, S334	N334, T388, T818
PGT121	QAHCN328-332, R332	

Signals among the 15% strongest learnt signals for each classifier were considered.

https://doi.org/10.1371/journal.pcbi.1005789.t002

Fig 2. Motif logo for the test sequence HM469973 using the PG9 classifier. For the motif logo the contribution of 5% of the strongest discriminant signals to the classification is considered. The height of the letters depends on the proportional contribution to the classification. Amino acids of the test sequence that influence the classification outcome towards neutralization susceptibility are displayed in capital letters and blue color; lowercase letters and orange color if they contribute to neutralization resistance. For better interpretability, the corresponding positions of the amino acids in the envelope sequence of the HIV strain HXB2 are shown on the x-axis.

https://doi.org/10.1371/journal.pcbi.1005789.g002

classification outcome of the test sequence. The asparagine (N) at position 160, which is known to be decisive for a successful binding of the PG9 bNAb, as well as the lysine (K) at position 157 have the highest contribution to the classification result, more precisely to susceptibility. In general, most of the 5% strongest signals influence the classification result towards susceptibility.

Trend towards bNAb resistance over time

In order to investigate whether neutralization sensitivity of HIV-1 to bNAbs has changed over time, we additionally built support vector regression models to directly predict the (logarithmized) IC50 value for the 11 considered bNAb. For subtype B variants, a continuous trend towards resistance has been already confirmed in certain cohorts (around 40 samples) of the French and Dutch HIV-1 population [55–57]. Since evolving resistance to antibody neutralization in the HIV-1 species would have major implications on the antibody selection for current vaccine development, it is important to know whether such a drift towards resistance also exists in the global HIV-1 population for all subtypes. In contrast to an experimental setting, where the large number of viral strains and the accruing costs make neutralization assays for the comprehensive global population hardly possible, our prediction models can be easily used to examine this question based on the vast amount of available sequence data.

To model the global HIV-1 population over time, we used all available envelope sequences from the Los Alamos HIV sequence database (around 34,000 after data processing, see Methods and S9 Table for accession numbers) comprising viral isolates from all major subtypes over a time interval from 1981 to 2013. We divided the given time interval into the following six time periods to account for changes in HIV-1 treatment strategies: 1981-1986 before ART, 1987-1991 ART monotherapy, 1992-1995 ART combination therapy (cART), 1996-1999 cART with protease inhibitors, 2000-2005 cART with Lopinavir/Ritonavir, and 2006-2013 cART with Maraviroc/Raltegravir. With this partitioning of the data, we additionally covered the considered time intervals in the previously performed experimental studies [55–57]. An overview of the different subtypes and country distribution per time period are displayed in S2 Fig.

In order to identify a drift towards resistance, we performed a permutation test for umbrella alternatives [58] on the predicted (logarithmized) IC50 values grouped by the six time periods. The umbrella test [59] is a more general test than the Jonckheere-Terpstra test [60, 61]. Instead of testing for a monotonic trend, it tests for a peak in one of the time periods—a trend, monotonically increasing before and decreasing after the peak. The permutation test of umbrella alternatives [62] provides in additional partial p-values for each group, which enables a better analysis of the trend. Here, we define a trend towards resistance, if the peak is in the last time periods (see Methods for details). In contrast to the experimental studies [55–57], our data set is much larger, covers longer time periods, and is more heterogeneous. Thus, we expected to see more variation in our groups and therefore decided to use the umbrella test as a more general test in our case. However, we additionally provide the statistics for the Jonckheere-Terpstra test in S6 Table, which can be seen as a more conservative test.

When considering only the subtype B variants of the around 34,000 viral isolates (17,392), we observed a statistically significant increase of the predicted (logarithmized) IC50 values over the six time periods to each of the 11 bNAbs ($P \leq 0.001$ using the umbrella test and a significance threshold $t = \alpha/\#tests = 0.05/22 = 0.0023$ with Bonferroni correction for multiple testing). Thus, we could confirm the trend towards bNAb resistance [55–57] on a larger and more diverse data set. The predicted (logarithmized) IC50 values for the subtype B samples for all 11 bNAbs are provided in S3 Fig. Note that in order to avoid misleading data visualization, we present all the predicted values for all 11 bNAbs on the same y-scale, though the bNAbs differ in their neutralization strength. Though we find the last time periods as part of a significant trend in the data for PG9, PG16 and PG128, the partial p-values indicate rather a plateau distribution than a clear trend towards resistance in the last time periods (see S6 Table).

In addition, we predicted and analyzed the neutralization sensitivity of the non-B subtype samples (16,546) to the 11 bNAbs. A statistically significant trend towards resistance was observed for all considered bNAbs, but PG9, PG16, PG121,PGT128 and NIH-4546. In Fig 3 we show exemplarily the predicted (logarithmized) values for the bNAbs (A) 3BNC117 (CD4bs), (B) PGT121 (V3 loop), (C) 35O22 (gp41/gp120), and (D) PG16 (V1/V2 loop); see S4 Fig for all bNAbs and non-B subtype samples. While for the bNAb PGT128 there was no significant peak at all, the trend towards resistance to the bNAb PGT121 was not significant after Bonferroni correction for multiple testing. For PG9, PG16 and NIH-4546, we detected a significant peak in the data, but not in the last time period, which we however required to determine a trend towards resistance (see Methods for details). The peak for NIH-4546 was slightly shifted (in the fifth time period), whereas for PG9 and PG16 a significant peak was already detected in the first time period, that is, the HIV variants tend to become more susceptible in the last time period. Since there are no experimental data on HIV-1 resistance development trends to bNAbs for the non-B subtype population, we decided to first rule out the possibility of a confounder that might lead to the contrasting trend for PG9 and PG16. Pfeifer et al. [63] discovered that there is a statistically significant bias in the neutralization susceptibility of HIV-1 variants to PG9 and PG16 depending on the coreceptor usage of the virus. For successful entry of the virus into the host cell, the glycoprotein gp120 has not only to bind to the CD4-receptor on the host cell, but also to a second chemokine receptor on the host cell that acts as co-factor (coreceptor). The coreceptors mainly used by HIV-1 are CCR5 and CXCR4. Depending on the coreceptor usage, the virus strain is referred to as R5- or X4-tropic, or dual-tropic if the virus can bind to both of these coreceptors, and X4-capable, if they are either dual-tropic or X4-tropic [64]. X4-capable viruses have been shown to be more resistant to PG9 and PG16 [63]. This means that PG9 and PG16 have an R5-bias, that is, they are better in neutralizing R5-tropic viruses than X4-capable viruses. By determining the coreceptor usage for all considered viral samples with the most widely used tool for genetic tropism

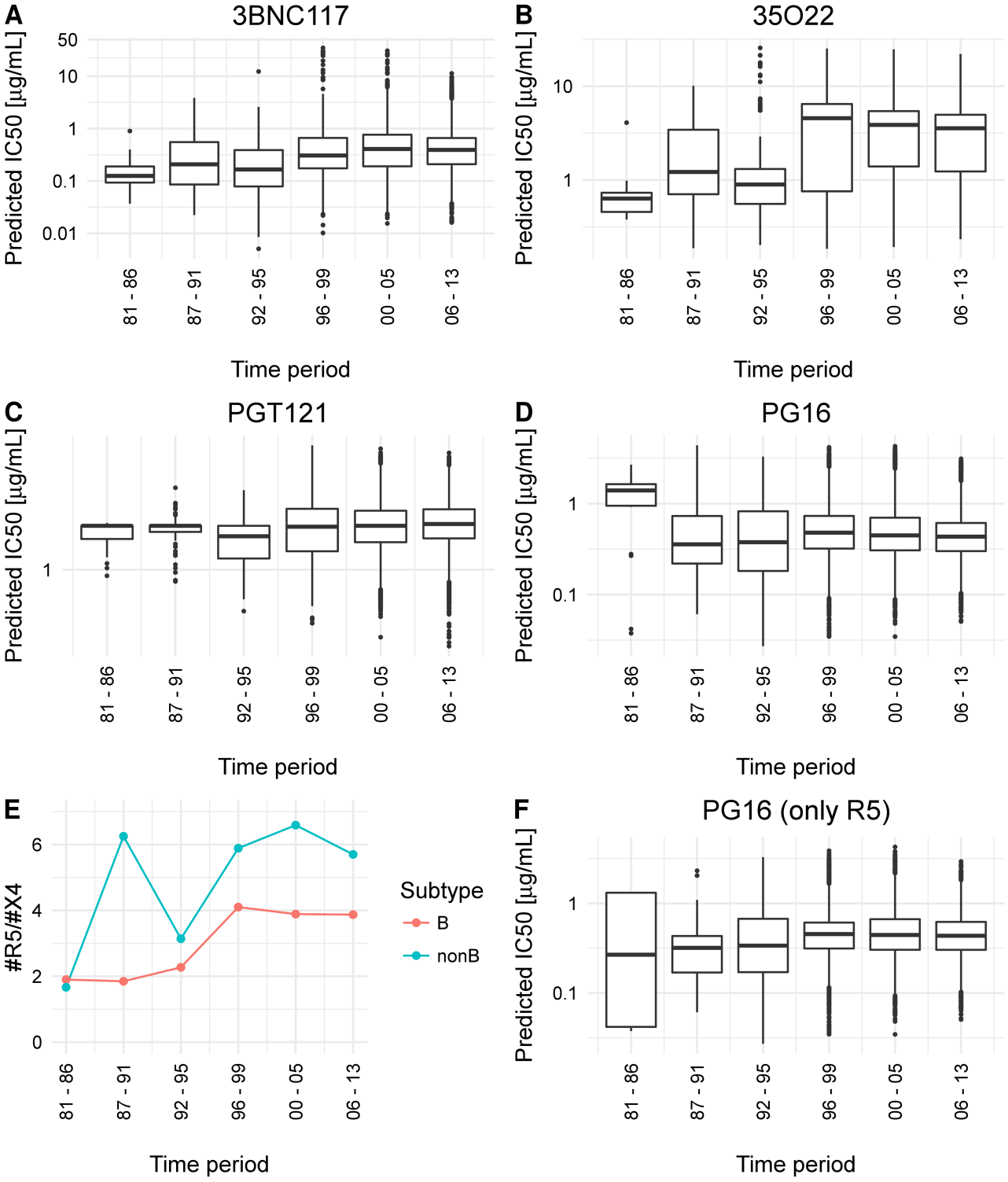


Fig 3. Characteristics of the non-B subtype HIV variants over time. Predicted neutralization sensitivity of non-B subtype HIV-1 variants to bNAbs over time. **A** - **D**: the predicted neutralization sensitivity of HIV-1 samples of the non-B subtype to the bNAbs 3BNC117, 35O22, PGT121, and PG16. **E**: the ratio of R5-tropic and X4-capable non-B subtype HIV-1 variants in the Los Alamos HIV sequence database over the six time periods. **F**: predicted neutralization sensitivity of R5-tropic variants of the non-B subtype to PG16.

https://doi.org/10.1371/journal.pcbi.1005789.g003

testing, geno2pheno[coreceptor] [65], we detected a stronger increasing ratio of R5- to X4-capable viruses over the time periods for the non-B than for the subtype B samples (see Fig 3E and S2 Table). Thus, we might see an increase in neutralization susceptibility to PG9 and PG16 due to the relative increase of R5-tropic variants in the later time periods, since R5-tropic variants are more susceptible to PG9 and PG16. With an analysis, analogous to Pfeifer et al. [63], we observed an R5-bias of the bNAb PGT128 ($P = 0.00568$ using a two-sided Fisher's exact test, see also S3 Table). Fig 4 shows the relative number of resistant and susceptible HIV strains to PGT128 in comparison to PG9, PG16, VRC-PG04 and VRC01. Data for VRC01, VRC-PG04, PG9 and PG16 was taken from Pfeifer et al. [63]. We additionally analyzed the association between coreceptor usage and neutralization sensitivity for all considered 11 bNAbs. As can be seen in S5 Fig, we could not detect other bNAbs with an R5-bias. For the bNAb PG16, a resistance trend was only detected for the R5-tropic variants (see Fig 3F). Note that sequences from the beginning of the HIV epidemic (first two time periods) were probably from patients having AIDS and not at early stage of HIV infection as nowadays. Since at early stage of clinical HIV infection usually R5-tropic viruses are predominant [66, 67], this might also explain the decrease of X4-capable variants in the database over time. The first time period contains also less samples than later time periods, which might influence the trend.

We could detect a trend towards resistance for all 11 bNAbs regarding the subtype B HIV-variants (10/11 if Jonckheere-Terpstra test is used). For the non-B subtype population, we observed the trend for only 6 of the 11 bNAbs (5 of 11 if Jonckheere-Terpstra test is used). A summary of the findings and the corresponding p-values of both statistical tests can be found in S6 Table.

Conclusion

In this study, we showed that neutralization sensitivity of new HIV-1 variants to broadly neutralizing antibodies (bNAbs) is predictable using neutralization information from existing neutralization assays. The credibility of the models were underlined by the finding that the prediction models learnt important binding sites for the bNAbs implicitly, without explicitly getting this type of information in the learning process. Hence, additional information such as structural binding site information is unlikely to boost the performance significantly. We increased the interpretability of the models, by offering the user more information on the prediction outcome in form of a motif logo where the logo displays the contribution of the pivotal residues of the test sequence to the prediction. In general, our method could be applied as a recommendation tool for bNAbs therapy, but it could already be used in planning clinical trials concerning bNAbs therapy to screen patients before those therapies are approved for clinical use.

It is unquestioned, that an effective bNAb therapy will consist of a combination of bNAbs targeting distinct epitopes on the envelope spike to prevent the emergence of antibody resistance. To determine which and how many bNAbs to choose, several studies analyzed systematically combinations of different bNAbs [68, 69] experimentally but also predicted the neutralization sensitivity using additive models. However, these prediction models need the neutralization sensitivity of the virus to the single bNAbs in the combination as input. Our learnt classifiers could be extended similarly to additive models that predict if or how effective a combination of bNAbs is requiring only the envelope sequence of the virus.

Despite the good performance and biological relevance of our classifiers, the current models are not suited for a direct application in clinical settings. In the clinical setting, it is more tolerable to misclassify a sensitive HIV variant to a bNAb than misclassifying a resistant HIV-1 variant. While the area under the ROC curve was helpful in determining, if the classification task

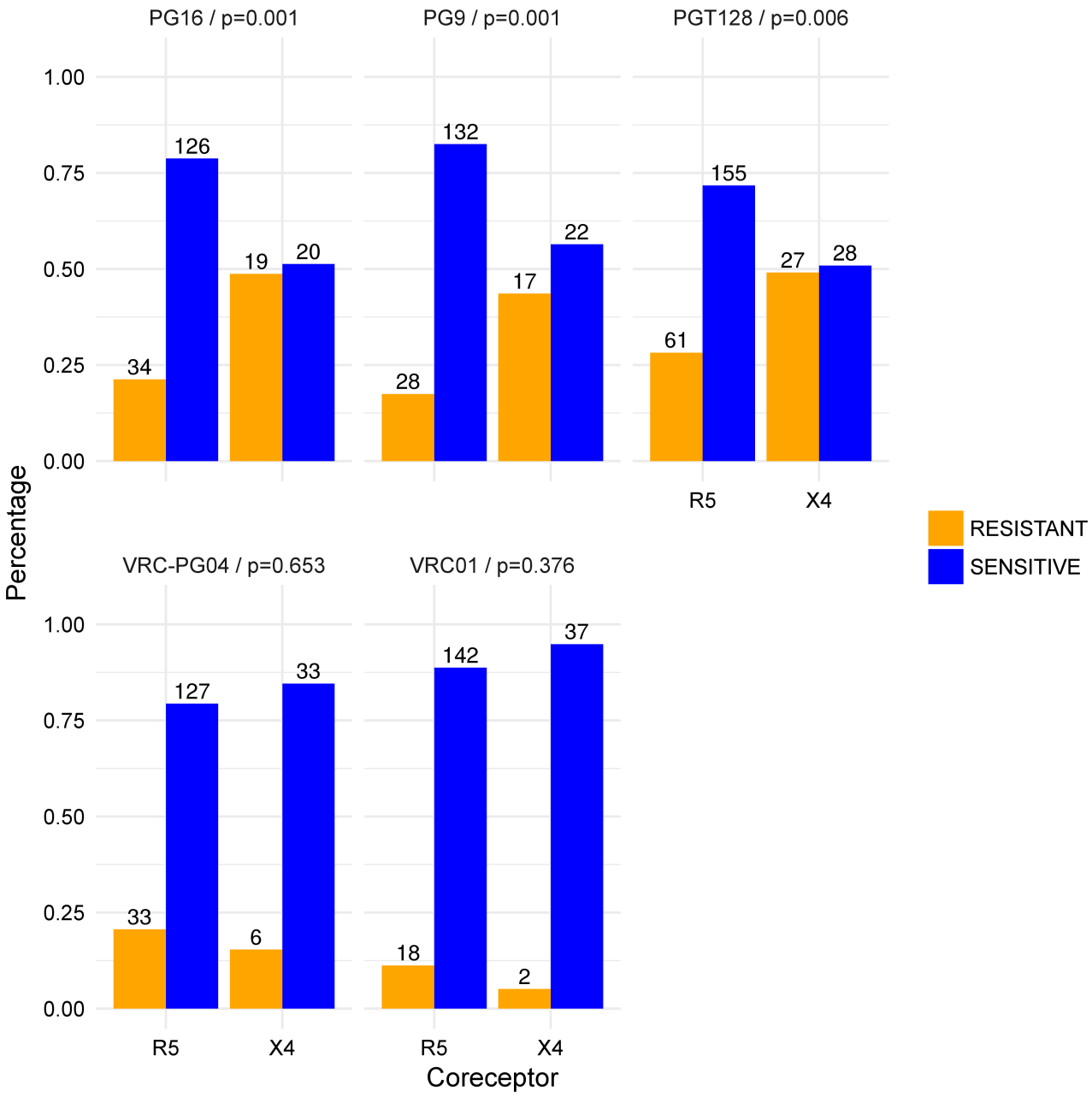


Fig 4. Association between coreceptor usage and neutralization sensitivity. Relative number of resistant (orange) and susceptible (blue) strains with regard to their coreceptor usage for the bNAbs PGT128 and VRC-PG04, as well as for VRC01, PG9, and PG16. Statistical significance was assessed with a two-sided Fisher's exact test. Data for VRC01, VRC-PG04, PG9 and PG16 was taken from Pfeifer et al. [63].

https://doi.org/10.1371/journal.pcbi.1005789.g004

can be accomplished with our proposed methods and for comparison reasons, it is not the best approach to design the models for the final application setting due to the low average specificity for some of the bNAbs (see S5 Table). In order to apply our models in the clinical setting, clinical data has to be analyzed instead of pseudovirus panel data. In addition, an appropriate false discovery rate has to be agreed on with the clinicians, for which the final models can be optimized for. This holds for any method used for this classification task. Apart from their

potential use as recommendation tool, computational prediction models can in general be used to analyze the change in the neutralization sensitivity of HIV-1 over time. We could confirm previous results suggesting a trend towards antibody resistance in the subtype B population [55–57]. Moreover, we scaled up the analysis to the global HIV-1 population, showing that there is a general drift towards antibody resistance in the world-wide HIV-1 population for most of the bNAbs. These findings are relevant for the selection of suitable vaccine candidates; a combination of bNAbs is however still very potent in neutralizing HIV-1 [56].

Materials and methods

Neutralization assay

We used the IC50 titers of 11 different antibodies (PG9, PG16, 35O22, VRC01, VRC-PG04, 3BNC117, NIH45-46, PGT128, PGT121, 10-996 and 10-1074) for 115 to 220 HIV-1 isolates from three different neutralization assays [10, 26, 47] (see S8 Table). For the bNAbs PG9, PG16, PGT121 and VRC01 neutralization information was available from two neutralization assays. Although the overlap of tested HIV-1 samples was quite high as well as the correlation of the corresponding IC50 titers, we did not merge the information from the two assays for these bNAbs. We represented each HIV-1 isolate by the amino acid sequence of its envelope glycoprotein from the Los Alamos HIV sequence database [48]. We excluded HIV-1 isolates for which no GenbankID was available, or the envelope sequence was shorter than 800 amino acids.

Data preparation for prediction models

Since the feature vector of each sample has to be of the same length for most of the kernels, we aligned the amino acid sequences with the HIValign tool from the Los Alamos HIV sequence database [48]. For the polynomial and Gaussian RBF kernel the amino acid sequences have to be transformed to a real-valued input. We used one-hot encoding to represent the sequence information for the polynomial kernel, i.e., each amino acid a_i, $i \in \{1, \ldots 20\}$ is transformed into a 20-dimensional vector, where only the i-th entry is 1, and the others are 0. For the Gaussian RBF kernel, we encoded the sequence information using physico-chemical properties (RBF1 [70] and RBF2 [71]).

In the classification task, the IC50 titers were converted to -1 if the IC50 value was above 50 μg/mL (resistant), and otherwise to +1 (susceptible) similar to Doria-Rose et al. [47]. Due to their distribution, the IC50 values for the regression task were logarithmized.

Kernel comparison and parameter settings

To test if l-mer string kernels (such as the oligo kernel [52] or the weighted degree kernel with shifts (WDKS) [51]) perform better than conventional kernels (such as the polynomial or the Gaussian RBF kernel), we compared the performances of prediction models based on each of these kernels. The comparison was conducted by 10 runs of a 5-fold nested cross-validation using AUC and Pearson Correlation Coefficient as performance measure for the classification and regression task, respectively. The tested parameter range for each kernel is listed in S1 Table. To determine the best parameter setting for each bNAb prediction model, we performed an additional 5-fold cross-validation. Since in the nested cross-validation mainly small values of the width parameter ($2\sigma^2$) led to high prediction performances, we further sampled the range between 0 and 3 for this parameter.

Method comparison and parameter settings

We compared the final SVM classifiers based on the oligo kernel with random forests, an SVM using a linear kernel, a neural network, and a logistic regression with lasso regularization (lasso). For the random forest, the neural network, and the lasso approach, the amino acid sequences were mapped to their index in the amino acid alphabet. For the linear kernel, the sequences have been encoded using the one-hot encoding approach, i.e., each amino acid a_i, $i \in \{1, \ldots 20\}$ is transformed into a 20-dimensional vector, where only the i-th entry is 1, and the others are 0. While the random forest approach can handle internally categorical variables with more than two factors, we created dummy features for each alignment position for the neural network and the lasso approach. We used the R package randomForest [72] setting the number of variables randomly sampled as candidates at each split (mtry) to the square root of the number of features in the model and the numbers of tree to grow (ntree) to 500. For the neural network, we used the R package neuralnet [73], we used one layer and set the number of hidden layers to the square root of the number of features. To build the logistic regression models, we used the R package glmnet [74] where we used lasso as regularization ($\alpha = 1$) and tuned the lambda parameter in an internal cross-validation. For the linear kernel, we used the R package kernlab [75] setting the kernel to vanilladot using the default cost parameter C. The performance was assessed on 10 runs of stratified 5-fold cross-validation. We did not compare the performance over a nested cross-validation iterating over different hyperparameters for the models, due to the infinite range of possibilities. We used the R package mlr [76] to compare all the methods.

Retrieving discriminant signals from the oligo kernel

A kernel $k(x, x')$ can be considered as a similarity function between instances x and x'. The oligo kernel computes $k(x, x')$ between two sequences x and x' of same length L by comparing the co-occurrences of their substrings (*oligomer*) of length l with $1 \leq l \leq L$ within a particular distance (width parameter σ^2). Therefore, the occurrence of a particular l-mer in a sequence x (denoted as x_ω) is encoded by the so-called oligo function μ

$$\mu_\omega(t) = \sum_{p \in x_\omega} \exp\left(-\frac{1}{2\sigma^2}(t-p)^2\right) \quad (1)$$

with the continuous position variable $t \in [1, L]$ and σ^2 controlling the positional uncertainty. As described in [52], the corresponding learnt weight of the classifier for each oligomer ωat each position t can be retrieved by

$$|w_\omega(t)| = |\sum_{i=1}^{N} \alpha_i y_i \mu_\omega^i(t)|, \quad (2)$$

where $i \in \{1, \ldots, N\}$ denotes the i-th training sample with $\alpha_i \geq 0$ and $y_i \in \{-1, 1\}$ being the learnt weight and classification label of the i-th sample, and with $\mu_\omega^i(t)$ being the oligo function of l-mer ω of the training sequence i at position t.

Considering the weights of each oligomer for the test sequence, there exists only one oligomer ω containing the actual residue as starting point whose contribution is calculated as

$$S_\omega^*(t) = \sum_{i=1}^{N} \alpha_i y_i < \mu_\omega^i(t), \mu_\omega^*(t) >, \quad (3)$$

with μ_ω^* being the oligo function of l-mer ω of the test sequence. For l-mers > 1 the computed

contribution is assigned to all amino acids of the oligomer. To visualize the motif logos we used Weblogo 3.0 [77].

Data preparation and analysis of the sequence data for the neutralization sensitivity trend analysis

We used all available envelope sequences from the Los Alamos HIV sequence database [48] (37,137), except the sequences that the prediction models were built on and those that were too short, resulting in 35,524 envelope sequences. For 1586 sequences no date was given, and thus these sequences were excluded as well, resulting in 33,938 considered viral envelope sequences. Before predicting the IC50 value for each test sequence, the sequences were aligned to the data sets using profile-to-profile MUSCLE alignment [78] with the Ugene tool [79]. Instead of predicting the IC50 value, the regression models were trained to predict the logarithmized IC50 value.

To identify a drift towards resistance, we performed a permutation test for umbrella alternatives [58] on the predicted (logarithmized) IC50 values grouped by the six time periods. We applied the umbrella test according to Basso et al. [58, 62] with the provided R code. The umbrella test is a generalization of the Jonckheere-Terpstra test, testing for a peak instead of a monotone trend. A significant peak in the last time period was considered as indicator for an increasing trend in IC50 values and thus, a trend towards bNAb resistance.

Coreceptor prediction

To predict the coreceptor usage, we used the well established prediction tool geno2pheno[coreceptor] [65]. The prediction tool uses a linear support vector machine to predict whether a sequence is from a X4-capable or an R5-tropic virus, only based on the V3 loop sequence of the viral envelope sequence. For each V3 sequence, geno2pheno[coreceptor] provides the false-positive rate (FPR), which is a measure for the confidence of the prediction. geno2pheno[coreceptor] reports the minimal FPR threshold at which the sequence would be classified as X4-capable. For the manuscript, we used an FPR cutoff of 10% to determine X4-capable ($\leq$ 10%) and R5-tropic viruses ($>$ 10%) as recommended by the European Consensus Group on clinical management of HIV-1 tropism testing [80]. Since there are also reasons for other cutoff choices, we additionally provide the results for the FPR cutoffs according to the German and Austrian treatment guidelines ($\leq$ 5%: X4-capable; $\geq$ 15%: R5-tropic) in the Supporting Information.

Coreceptor usage distribution in the Los Alamos HIV sequence database

We used the prediction tool geno2pheno[coreceptor] [65] to determine the coreceptor usage of the 33,938 viral isolates from the Los Alamos HIV sequence database [48]. According to the prediction tool, we excluded in total 545 sequences due to warnings regarding the alignment quality and due to warnings regarding the V3 loop quality (alignment score $\geq$ 95th percentile).

Association between coreceptor usage and neutralization by PGT128

For this analysis, we used all available sequences from the CATNAP tool [81], retrieved on 2016-08-10. Since we used a neutralization sensitivity cutoff of 50 μg/mL to determine resistance and susceptibility, all sequences, whose neutralization sensitivity were only given as a cutoff less than 50 μg/mL were excluded. In addition, we excluded two sequences due to poor

V3 alignment quality. Coreceptor usage was determined using the prediction tool geno2-pheno[coreceptor] [65]. To test whether the sensitivity to an antibody is significantly different with regard to coreceptor usage, we performed a two-sided Fisher's exact test for the two-by-two contingency tables with resistant and susceptible as the row label and X4-capable/R5-tropic as the column label using significance level = 0.05 with the null hypothesis that there is no difference.

Implementation details

The prediction and analysis of the neutralization sensitivity were implemented mainly in R [82], version 3.2.1 (2015-06-18) and the R package kernlab [75]. The oligo kernels were computed using a customized version of the Shogun-Toolbox [83] (version 2.0.0). To visualize the motif logos we used Weblogo 3.0 [77].

Data availability

In S8 Table we provide the virus names that we considered for the prediction models as well as the study ID of the neutralization assay. With this, the corresponding neutralization data can be retrieved from CATNAP [81]. For the trend analysis, we provide the accession numbers of each considered HIV-1 variant in the Los Alamos HIV sequence database [48] in S9 Table.

At https://github.com/annahake/g2p-bnab, we additionally provide the computed kernels for the final models as well as the resampling instance for the 10 runs of stratified 5-fold cross-validation. As mentioned in the Conclusion, the final models are so far not adapted for clinical usage.

Supporting information

S1 Fig. Prediction performance of the regression models. The prediction performance of the 11 SVM regression models based on the oligo kernel was measured by the Pearson correlation coefficient, displayed on the y-axis. The regression models are named according to the bNAb they are trained on, shown on the x-axis. The colors of the boxes refer to the epitope category of the corresponding bNAb. The gray dashed line denotes no linear relationship. The prediction performance was assessed in 10 runs of 5-fold nested cross-validation. Most regression models show good performances (average Pearson correlation coefficient ≥ 0.3).
(TIFF)

S2 Fig. Subtype and geographical distribution of HIV-1 variants over the six time periods. For each time period, we display the number of samples from the three most frequent countries as well as the sum of samples from the remaining countries (OTHER). The country distribution is shown for the subtype B (**A**) and the subtype non-B HIV-1 variants (**B**). In **C** we display the number of samples in each time period for the five most frequent subtypes, and additionally the number of samples for the non-B subtypes (dashed line).
(TIFF)

S3 Fig. Neutralization sensitivity analysis for the subtype B HIV-1 variants. Predicted neutralization sensitivity of HIV-1 variants (subtype B) from the Los Alamos HIV sequence database to all 11 bNAbs. Neutralization sensitivity (logarithmized IC50 values) was predicted using our SVM regression models based on the oligo kernel. The HIV-1 variants are grouped in six, consecutive, time periods, displayed on the x-axis. A trend towards bNAb resistance was reported if the neutralization sensitivity increased over time with a significant peak in the last time period. The significance was determined using a permutation test for umbrella alternatives and a significance threshold t = α/# total tests = 0.05/22 = 0.0023 with Bonferroni

correction for multiple testing.
(TIFF)

S4 Fig. Neutralization sensitivity analysis for the non-B subtype HIV-1 variants. Predicted neutralization sensitivity of HIV-1 variants (subtype non-B) from the Los Alamos HIV sequence database to all 11 bNAbs. Neutralization sensitivity (logarithmized IC50 values) was predicted using our SVM regression models based on the oligo kernel. The HIV-1 variants are grouped in six, consecutive, time periods, displayed on the x-axis. A trend towards bNAb resistance was reported if the neutralization sensitivity increased over time with a significant peak in the last time period. The significance was determined using a permutation test for umbrella alternatives and a significance threshold t = α/# total tests = 0.05/22 = 0.0023 with Bonferroni correction for multiple testing.
(TIFF)

S5 Fig. Association between coreceptor usage and neutralization sensitivity. For all considered 11 bNAbs, we display the relative number of resistant (orange) and susceptible (blue) strains with respect to their predicted coreceptor usage (R5-tropic or X4-capable). Statistical significance was assessed with a two-sided Fisher's exact test.
(TIFF)

S6 Fig. Prediction performance comparison for different machine learning approaches. For each bNAb classifier, the prediction performance measured by the area under the ROC curve (AUC) is displayed for our SVM models using the oligo kernel, an SVM model using the linear kernel, a logistic regression model with lasso regularization, a random forest model, and a neural network model.
(TIFF)

S1 Table. Performance comparison of different kernels and the investigated parameter range. In order to select a kernel for the SVM models, the performance of the polynomial kernel, radial basis function kernel (RBF), weighted degree with shifts kernel (WDKS) and the oligo kernel (Oligo) were compared in 10 runs of a 5-fold nested cross-validation. The cost parameter C of the SVM was sampled in the range from 10E-6 to 10E6 by powers of 10. The two RBF kernels differ in the physico-chemical encoding of the amino acid sequences (see Materials). The parameters of each kernel as well as the sampled range for each parameter are presented in the first sheet. The second sheet contains the prediction performance of each kernel measured by the Area under the ROC curve (AUC) in 10 runs of a 5-fold nested cross-validation exemplarily for all 11 bNAbs. All kernels performed equally well for all bNAbs, apart from VRC-PG04, for which the oligo kernel performed better. Therefore, the oligo kernel was taken to build the prediction models.
(XLSX)

S2 Table. Ratio of R5-tropic to X4-capable viruses in the LANL database. The observed percentage of X4-capable and R5-tropic HIV-1 variants in the Los Alamos HIV sequence database over the six considered time-periods. The coreceptor usage was determined using the well-established prediction tool geno2pheno[coreceptor] using an FPR-cutoff of 10% as recommended by the European Consensus Group on clinical management of HIV-1 tropism testing, and the FPR-cutoff recommended by then German and Austrian treatment guidelines ($\leq$ 5%: X4-capable; $\geq$ 15%: R5-tropic).
(XLSX)

S3 Table. Association between coreceptor usage and neutralization by PGT128. HIV-1 variants from a neutralization assay against PGT128 and their coreceptor usage are presented in

this contingency table. The coreceptor usage was determined using the well-established prediction tool geno2pheno[coreceptor] using an FPR-cutoff of 10% as recommended by the European Consensus Group on clinical management of HIV-1 tropism testing, and the FPR-cutoff recommended by then German and Austrian treatment guidelines ($\leq$ 5%: X4-capable; $\geq$ 15%: R5-tropic). To test whether the sensitivity to PGT128 is significantly different with regard to coreceptor usage for each FPR-cutoff, we performed a two-sided Fisher's exact test using significance level = 0.05.
(XLSX)

S4 Table. Performance comparison of the full and reduced prediction models. Performance comparison (AUC) of the full classification models and classification models that were built using the strongest p% beforehand learnt discriminant signals. To compare the performances of the reduced to the full models, we divided our data set into three partitions. Partition A (40% of the data) was used to build a full model (parameters were fitted using a 5-fold cross-validation), from which the p% strongest discriminant signals were extracted. Partition B (40% of the data) was used to build prediction models only based on the p% of the discriminant signals. On the same data set, we also build the full prediction model. The performance of full and reduced models were tested on the third partition C (20% of the data). For each prediction model, we tested the null hypothesis that the full and reduced model have on average the same performance using a paired, two-sided, Wilcoxon test and a significance threshold t = α/#reduced models = 0.05/8 = 0.00625 with Bonferroni correction for multiple testing. For each prediction model, we could not reject the null hypothesis for most of the reduced models. Only five comparisons show a significant p-value (marked in red), which is however very close to the Bonferroni correction threshold.
(XLSX)

S5 Table. Prediction performance comparison for different machine learning approaches. For each bNAb classifier, the prediction performance measured by the averaged area under the ROC curve (AUC) is displayed for our SVM models using the oligo kernel, an SVM model using the linear kernel, a logistic regression model with lasso regularization, a random forest model, and a neural network model. The prediction performance was assessed using 10 runs of a stratified 5-fold cross-validation. In addition, we provide the mean sensitivity and mean specificity for each model.
(XLSX)

S6 Table. Statistical analysis of the neutralization sensitivity over time. To investigate whether there is a trend towards resistance over time, we performed a Jonckheere-Terpstra test as well a permutation test for umbrella alternatives. The first sheet contains an overview of bNAbs and subtypes for which we were able to detect a trend towards resistance. The second sheet provides the p-values from both tests.
(XLSX)

S7 Table. 1% learnt discriminant signals for each bNAb classifier. For each bNAb we display the 1% strongest learnt signals (oligomers), their position compared to the HXB2 reference and the learnt weight. Negative weights denote associations towards resistance, while positive weights denote associations towards sensitivity. The weights for different classifiers are not comparable. Note that a gap signal implies an absence of a certain amino acid at a certain position. Especially if several amino acids are associated with sensitivity or resistance at a certain position, the gap character might be strongly associated with the other direction (resistance or sensitivity), even stronger than the individual amino acid signals.
(XLSX)

S8 Table. Neutralization assay data used for the predictions. We built prediction models for each bNAb separately based on existing neutralization assay data. Here, we list for all 11 bNAbs, the considered viruses (virus names) together with the corresponding study ID in CATNAP [81].
(XLSX)

S9 Table. Data for the trend analysis. The accession numbers of the HIV-1 envelope amino acid sequences from the Los Alamos HIV sequence database [48] that we used to analyze neutralization sensitivity over time together with the predicted coreceptor usage and our predicted IC50 values.
(XLSX)

Author Contributions

Conceptualization: Anna Hake, Nico Pfeifer.

Data curation: Anna Hake, Nico Pfeifer.

Formal analysis: Anna Hake.

Investigation: Anna Hake.

Methodology: Anna Hake, Nico Pfeifer.

Project administration: Anna Hake, Nico Pfeifer.

Supervision: Nico Pfeifer.

Validation: Anna Hake, Nico Pfeifer.

Visualization: Anna Hake.

Writing – original draft: Anna Hake.

Writing – review & editing: Anna Hake, Nico Pfeifer.

References

1. UN Joint Programme on HIV/AIDS (UNAIDS). Global AIDS Update—2016; Accessed: 2017-01-04. Available from: http://www.refworld.org/docid/574e8d394.html.
2. Arts EJ, Hazuda DJ. HIV-1 Antiretroviral Drug Therapy. Cold Spring Harbor Perspectives in Medicine. 2012; 2(4):a007161–a007161. https://doi.org/10.1101/cshperspect.a007161 PMID: 22474613
3. McCoy LE, Weiss RA. Neutralizing antibodies to HIV-1 induced by immunization. The Journal of Experimental Medicine. 2013; 210(2):209–223. https://doi.org/10.1084/jem.20121827 PMID: 23401570
4. Scheid JF, Mouquet H, Feldhahn N, Seaman MS, Velinzon K, Pietzsch J, et al. Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. Nature. 2009; 458 (7238):636–640. https://doi.org/10.1038/nature07930 PMID: 19287373
5. Moir S, Malaspina A, Fauci AS. Prospects for an HIV vaccine: leading B cells down the right path. Nature Structural & Molecular Biology. 2011; 18(12):1317–1321. https://doi.org/10.1038/nsmb.2194
6. Diskin R, Scheid JF, Marcovecchio PM, West AP, Klein F, Gao H, et al. Increasing the Potency and Breadth of an HIV Antibody by Using Structure-Based Rational Design. Science. 2011; 334 (6060):1289–1293. https://doi.org/10.1126/science.1213782 PMID: 22033520
7. Walker LM, Phogat SK, Chan-Hui PY, Wagner D, Phung P, Goss JL, et al. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science. 2009; 326(5950):285–289. https://doi.org/10.1126/science.1178746 PMID: 19729618
8. Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, Julien JP, et al. Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature. 2011; 477(7365):466–470. https://doi.org/10.1038/nature10373 PMID: 21849977

9. Wu X, Yang ZY, Li Y, Hogerkorp CM, Schief WR, Seaman MS, et al. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science. 2010; 329(5993):856–861. https://doi.org/10.1126/science.1187659 PMID: 20616233

10. Mouquet H, Scharf L, Euler Z, Liu Y, Eden C, Scheid JF, et al. Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. Proc Natl Acad Sci U S A. 2012; 109(47):E3268–E3277. https://doi.org/10.1073/pnas.1217207109 PMID: 23115339

11. Wyatt R, Sodroski J. The HIV-1 Envelope Glycoproteins: Fusogens, Antigens, and Immunogens. Science. 1998; 280(5371):1884–1888. https://doi.org/10.1126/science.280.5371.1884 PMID: 9632381

12. Zhou T, Georgiev I, Wu X, Yang ZY, Dai K, Finzi A, et al. Structural Basis for Broad and Potent Neutralization of HIV-1 by Antibody VRC01. Science. 2010; 329(5993):811–817. https://doi.org/10.1126/science.1192819 PMID: 20616231

13. Falkowska E, Ramos A, Feng Y, Zhou T, Moquin S, Walker LM, et al. PGV04, an HIV-1 gp120 CD4 binding site antibody, is broad and potent in neutralization but does not induce conformational changes characteristic of CD4. Journal of Virology. 2012; 86(8):4394–403. https://doi.org/10.1128/JVI.06973-11 PMID: 22345481

14. Scheid JF, Mouquet H, Ueberheide B, Diskin R, Klein F, Oliveira TYK, et al. Sequence and Structural Convergence of Broad and Potent HIV Antibodies That Mimic CD4 Binding. Science. 2011; 333 (6049):1633–1637. https://doi.org/10.1126/science.1207227 PMID: 21764753

15. McLellan JS, Pancera M, Carrico C, Gorman J, Julien JP, Khayat R, et al. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature. 2011; 480(7377):336–343. https://doi.org/10.1038/nature10696 PMID: 22113616

16. Julien JP, Cupo A, Sok D, Stanfield RL, Lyumkis D, Deller MC, et al. Crystal structure of a soluble cleaved HIV-1 envelope trimer. Science. 2013; 342(6165):1477–83. https://doi.org/10.1126/science.1245625 PMID: 24179159

17. Pancera M, Shahzad-Ul-Hussan S, Doria-Rose NA, McLellan JS, Bailer RT, Dai K, et al. Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. Nature Structural & Molecular Biology. 2013; 20(7):804–813. https://doi.org/10.1038/nsmb.2600

18. Pejchal R, Doores KJ, Walker LM, Khayat R, Huang PS, Wang SK, et al. A Potent and Broad Neutralizing Antibody Recognizes and Penetrates the HIV Glycan Shield. Science. 2011; 334(6059):1097–1103. https://doi.org/10.1126/science.1213256 PMID: 21998254

19. Julien JP, Sok D, Khayat R, Lee JH, Doores KJ, Walker LM, et al. Broadly Neutralizing Antibody PGT121 Allosterically Modulates CD4 Binding via Recognition of the HIV-1 gp120 V3 Base and Multiple Surrounding Glycans. PLoS Pathog. 2013; 9(5):1–15. https://doi.org/10.1371/journal.ppat.1003342

20. Kong L, Lee JH, Doores KJ, Murin CD, Julien JP, McBride R, et al. Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. Nature Structural & Molecular Biology. 2013; 20(7):796–803. https://doi.org/10.1038/nsmb.2594

21. Sok D, Doores KJ, Briney B, Le KM, Saye-Francisco KL, Ramos A, et al. Promiscuous Glycan Site Recognition by Antibodies to the High-Mannose Patch of gp120 Broadens Neutralization of HIV. Science Translational Medicine. 2014; 6(236):236ra63 https://doi.org/10.1126/scitranslmed.3008104 PMID: 24828077

22. Muster T, Steindl F, Purtscher M, Trkola A, Klima A, Himmler G, et al. A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. Journal of Virology. 1993; 67(11):6642–6647. PMID: 7692082

23. Burton D, Pyati J, Koduri R, Sharp S, Thornton G, Parren P, et al. Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. Science. 1994; 266(5187):1024–1027. https://doi.org/10.1126/science.7973652 PMID: 7973652

24. Zwick MB, Labrijn AF, Wang M, Spenlehauer C, Saphire EO, Binley JM, et al. Broadly Neutralizing Antibodies Targeted to the Membrane-Proximal External Region of Human Immunodeficiency Virus Type 1 Glycoprotein gp41. Journal of Virology. 2001; 75(22):10892–10905. https://doi.org/10.1128/JVI.75.22.10892-10905.2001 PMID: 11602729

25. Huang J, Ofek G, Laub L, Louder MK, Doria-Rose NA, Longo NS, et al. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. Nature. 2012; 491(7424):406–412. https://doi.org/10.1038/nature11544 PMID: 23151583

26. Huang J, Kang BH, Pancera M, Lee JH, Tong T, Feng Y, et al. Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface. Nature. 2014;advance on. https://doi.org/10.1038/nature13601

27. Horwitz JA, Halper-Stromberg A, Mouquet H, Gitlin AD, Tretiakova A, Eisenreich TR, et al. HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. Proc Natl Acad Sci U S A. 2013; 110(41):16538–43. https://doi.org/10.1073/pnas.1315295110 PMID: 24043801

28. Klein F, Halper-Stromberg A, Horwitz JA, Gruell H, Scheid JF, Bournazos S, et al. HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. Nature. 2012; 492(7427):118—122. https://doi.org/10.1038/nature11604 PMID: 23103874

29. Diskin R, Klein F, Horwitz JA, Halper-Stromberg A, Sather DN, Marcovecchio PM, et al. Restricting HIV-1 pathways for escape using rationally designed anti-HIV-1 antibodies. The Journal of Experimental Medicine. 2013; 210(6):1235–1249. https://doi.org/10.1084/jem.20130221 PMID: 23712429

30. Barouch DH, Whitney JB, Moldt B, Klein F, Oliveira TY, Liu J, et al. Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys. Nature. 2013; 503 (7475):224—228. https://doi.org/10.1038/nature12744 PMID: 24172905

31. Ledgerwood JE, Coates EE, Yamshchikov G, Saunders JG, Holman L, Enama ME, et al. Safety, pharmacokinetics and neutralization of the broadly neutralizing HIV-1 human monoclonal antibody VRC01 in healthy adults. Clinical & Experimental Immunology. 2015; 182(3):289–301. https://doi.org/10.1111/cei.12692

32. Caskey M, Klein F, Lorenzi JCC, Seaman MS, West AP Jr, Buckley N, et al. Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. Nature. 2015; 522:487–491. https://doi.org/10.1038/nature14411 PMID: 25855300

33. Schoofs T, Klein F, Braunschweig M, Kreider EF, Feldmann A, Nogueira L, et al. HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1. Science. 2016; 352 (6288):997–1001. https://doi.org/10.1126/science.aaf0972 PMID: 27199429

34. Scheid JF, Horwitz JA, Bar-On Y, Kreider EF, Lu CL, Lorenzi JCC, et al. HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption. Nature. 2016; 535(7613):556–560. https://doi.org/10.1038/nature18929 PMID: 27338952

35. Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM. The challenge of HIV-1 subtype diversity. N Engl J Med. 2008; 358(15):1590–602. https://doi.org/10.1056/NEJMra0706737 PMID: 18403767

36. Lengauer T, Pfeifer N, Kaiser R. Personalized HIV therapy to control drug resistance. Drug Discovery Today: Technologies. 2014; 11:57—64. https://doi.org/10.1016/j.ddtec.2014.02.004 PMID: 24847654

37. West AP, Scharf L, Horwitz J, Klein F, Nussenzweig MC, Bjorkman PJ. Computational analysis of anti-HIV-1 antibody neutralization panel data to identify potential functional epitope residues. Proc Natl Acad Sci U S A. 2013; 110(26):10598–603. https://doi.org/10.1073/pnas.1309215110 PMID: 23754383

38. Ferguson AL, Falkowska E, Walker LM, Seaman MS, Burton DR, Chakraborty AK. Computational Prediction of Broadly Neutralizing HIV-1 Antibody Epitopes from Neutralization Activity Data. PLOS One. 2013; 8(12):e80562. https://doi.org/10.1371/journal.pone.0080562 PMID: 24312481

39. Lacerda M, Moore PL, Ngandu NK, Seaman M, Gray ES, Murrell B, et al. Identification of broadly neutralizing antibody epitopes in the HIV-1 envelope glycoprotein using evolutionary models. Virology Journal. 2013; 10(1):347. https://doi.org/10.1186/1743-422X-10-347 PMID: 24295501

40. Cai Y, Karaca-Griffin S, Chen J, Tian S, Fredette N, Linton CE, et al. Antigenicity-defined conformations of an extremely neutralization-resistant HIV-1 envelope spike. Proceedings of the National Academy of Sciences. 2017; 114(17):4477–4482. https://doi.org/10.1073/pnas.1700634114

41. Chuang GY, Acharya P, Schmidt SD, Yang Y, Louder MK, Zhou T, et al. Residue-Level Prediction of HIV-1 Antibody Epitopes Based on Neutralization of Diverse Viral Strains. Journal of Virology. 2013; 87 (18):10047–10058. https://doi.org/10.1128/JVI.00984-13 PMID: 23843642

42. Evans MC, Phung P, Paquet AC, Parikh A, Petropoulos CJ, Wrin T, et al. Predicting HIV-1 broadly neutralizing antibody epitope networks using neutralization titers and a novel computational method. BMC Bioinformatics. 2014; 15(1):77. https://doi.org/10.1186/1471-2105-15-77 PMID: 24646213

43. Gnanakaran S, Daniels MG, Bhattacharya T, Lapedes AS, Sethi A, Li M, et al. Genetic Signatures in the Envelope Glycoproteins of HIV-1 that Associate with Broadly Neutralizing Antibodies. PLOS Computational Biology. 2010; 6(10):1–26. https://doi.org/10.1371/journal.pcbi.1000955

44. He L, Zhu J. Computational tools for epitope vaccine design and evaluation. Current Opinion in Virology. 2015; 11:103—112. https://doi.org/10.1016/j.coviro.2015.03.013 PMID: 25837467

45. Buiu C, Putz MV, Avram S. Learning the Relationship between the Primary Structure of HIV Envelope Glycoproteins and Neutralization Activity of Particular Antibodies by Using Artificial Neural Networks. International Journal of Molecular Sciences. 2016; 17(10). https://doi.org/10.3390/ijms17101710 PMID: 27727189

46. Hepler NL, Scheffler K, Weaver S, Murrell B, Richman DD, Burton DR, et al. IDEPI: Rapid Prediction of HIV-1 Antibody Epitopes and Other Phenotypic Features from Sequence Data Using a Flexible Machine Learning Platform. PLOS Computational Biology. 2014; 10(9):1–10. https://doi.org/10.1371/journal.pcbi.1003842

47. Doria-Rose NA, Louder MK, Yang Z, O'Dell S, Nason M, Schmidt SD, et al. HIV-1 Neutralization Coverage Is Improved by Combining Monoclonal Antibodies That Target Independent Epitopes. Journal of Virology. 2012; 86(6):3393–3397. https://doi.org/10.1128/JVI.06745-11 PMID: 22258252

48. Los Alamos National Laboratory HIV database; Accessed: 2015-01-30. http://www.hiv.lanl.gov.

49. Feldmann A, Pfeifer N. From predicting to analyzing HIV-1 resistance to broadly neutralizing antibodies. PeerJ PrePrints. 2015; 3:e1304v1.

50. West AP, Scharf L, Scheid JF, Klein F, Bjorkman PJ, Nussenzweig MC. Structural Insights on the Role of Antibodies in HIV-1 Vaccine and Therapy. Cell. 2014; 156(4):633–648. https://doi.org/10.1016/j.cell.2014.01.052 PMID: 24529371

51. Rätsch G, Sonnenburg S, Schölkopf B. RASE: recognition of alternatively spliced exons in C.elegans. Bioinformatics. 2005; 21 Suppl 1:i369–77. https://doi.org/10.1093/bioinformatics/bti1053 PMID: 15961480

52. Meinicke P, Tech M, Morgenstern B, Merkl R. Oligo kernels for datamining on biological sequences: a case study on prokaryotic translation initiation sites. BMC Bioinformatics. 2004; 5:169. https://doi.org/10.1186/1471-2105-5-169 PMID: 15511290

53. West AP, Diskin R, Nussenzweig MC, Bjorkman PJ. Structural basis for germ-line gene usage of a potent class of antibodies targeting the CD4-binding site of HIV-1 gp120. Proc Natl Acad Sci U S A. 2012; 109(30):E2083–90. https://doi.org/10.1073/pnas.1208984109 PMID: 22745174

54. Marshall RD. The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. Biochem Soc Symp. 1974; 40(40):17–26.

55. Bunnik EM, Pisas L, van Nuenen AC, Schuitemaker H. Autologous Neutralizing Humoral Immunity and Evolution of the Viral Envelope in the Course of Subtype B Human Immunodeficiency Virus Type 1 Infection. Journal of Virology. 2008; 82(16):7932–7941. https://doi.org/10.1128/JVI.00757-08 PMID: 18524815

56. Bouvin-Pley M, Morgand M, Meyer L, Goujard C, Moreau A, Mouquet H, et al. Drift of the HIV-1 Envelope Glycoprotein gp120 toward Increased Neutralization Resistance over the Course of the Epidemic: a Comprehensive Study Using the Most Potent and Broadly Neutralizing Monoclonal Antibodies. Journal of Virology. 2014; 88(23):13910–13917. https://doi.org/10.1128/JVI.02083-14 PMID: 25231299

57. Bouvin-Pley M, Morgand M, Moreau A, Jestin P, Simonnet C, Tran L, et al. Evidence for a Continuous Drift of the HIV-1 Species towards Higher Resistance to Neutralizing Antibodies over the Course of the Epidemic. PLoS Pathog. 2013; 9(7):e1003477. https://doi.org/10.1371/journal.ppat.1003477 PMID: 23853594

58. Basso D, Salmaso L. A permutation test for umbrella alternatives. Statistics and Computing. 2011; 21 (1):45–54. https://doi.org/10.1007/s11222-009-9145-8

59. Mack GA, Wolfe DA. K-Sample Rank Tests for Umbrella Alternatives. Journal of the American Statistical Association. 1981; 76(373):175–181. https://doi.org/10.2307/2287064

60. Jonckheere AR. A Distribution-Free k-Sample Test Against Ordered Alternatives. Biometrika. 1954; 41 (1/2):133–145. https://doi.org/10.1093/biomet/41.1-2.133

61. Terpstra TJ. The asymptotic normality and consistency of kendall's test against trend, when ties are present in one ranking. Indagationes Mathematicae. 1952; 14(3):327–333. https://doi.org/10.1016/S1385-7258(52)50043-X

62. Basso D, Pesarin F, Salmaso L, Solari A. Permutation Tests for Stochastic Ordering and ANOVA: Theory and Applications with R. 1st ed. Springer Publishing Company, Incorporated; 2009.

63. Pfeifer N, Walter H, Lengauer T. Association between HIV-1 coreceptor usage and resistance to broadly neutralizing antibodies. Journal of Acquired Immune Deficiency Syndromes. 2014; 67(2):107–112. https://doi.org/10.1097/QAI.0000000000000283 PMID: 25072615

64. Berger EA, Murphy PM, Farber JM. CHEMOKINE RECEPTORS AS HIV-1 CORECEPTORS: Roles in Viral Entry, Tropism, and Disease. Annual Review of Immunology. 1999; 17(1):657–700. https://doi.org/10.1146/annurev.immunol.17.1.657 PMID: 10358771

65. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. Nature Biotechnology. 2007; 25(12):1407—1410. https://doi.org/10.1038/nbt1371 PMID: 18066037

66. Zhu T, Mo H, Wang N, Nam D, Cao Y, Koup R, et al. Genotypic and phenotypic characterization of HIV-1 patients with primary infection. Science. 1993; 261(5125):1179–1181. https://doi.org/10.1126/science.8356453 PMID: 8356453

67. Schuitemaker H, Koot M, Kootstra NA, Dercksen MW, de Goede RE, van Steenwijk RP, et al. Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population. Journal of Virology. 1992; 66(3):1354–1360. PMID: 1738194

68. Kong R, Louder MK, Wagh K, Bailer RT, deCamp A, Greene K, et al. Improving Neutralization Potency and Breadth by Combining Broadly Reactive HIV-1 Antibodies Targeting Major Neutralization Epitopes. Journal of Virology. 2015; 89(5):2659–2671. https://doi.org/10.1128/JVI.03136-14 PMID: 25520506

69. Wagh K, Bhattacharya T, Williamson C, Robles A, Bayne M, Garrity J, et al. Optimal Combinations of Broadly Neutralizing Antibodies for Prevention and Treatment of HIV-1 Clade C Infection. PLOS Pathogens. 2016; 12(3):e1005520. https://doi.org/10.1371/journal.ppat.1005520 PMID: 27028935

70. Atchley WR, Zhao J, Fernandes AD, Drüke T. Solving the protein sequence metric problem. Proc Natl Acad Sci U S A. 2005; 102(18):6395–6400. https://doi.org/10.1073/pnas.0408677102 PMID: 15851683

71. Venkatarajan MS, Braun W. New quantitative descriptors of amino acids based on multidimensional scaling of a large number of physical-chemical properties. J Mol Model. 2001; 7(12):445–453. https://doi.org/10.1007/s00894-001-0058-5

72. Liaw A, Wiener M. Classification and Regression by randomForest. R News. 2002; 2(3):18–22.

73. Fritsch S, Guenther F. neuralnet: Training of Neural Networks; 2016. Available from: https://CRAN.R-project.org/package=neuralnet.

74. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software. 2010; 33(1):1–22. https://doi.org/10.18637/jss.v033.i01 PMID: 20808728

75. Karatzoglou A, Smola A, Hornik K, Zeileis A. kernlab—An S4 Package for Kernel Methods in R. Journal of Statistical Software. 2004; 11(9):1–20. https://doi.org/10.18637/jss.v011.i09

76. Bischl B, Lang M, Kotthoff L, Schiffner J, Richter J, Studerus E, et al. mlr: Machine Learning in R. Journal of Machine Learning Research. 2016; 17(170):1–5.

77. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A Sequence Logo Generator. Genome Research. 2004; 14(6):1188–1190. https://doi.org/10.1101/gr.849004 PMID: 15173120

78. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research. 2004; 32(5):1792–1797. https://doi.org/10.1093/nar/gkh340 PMID: 15034147

79. Okonechnikov K, Golosova O, Fursov M, the UGENE team. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics. 2012; 28(8):1166–1167. https://doi.org/10.1093/bioinformatics/bts091 PMID: 22368248

80. Vandekerckhove L, Wensing A, Kaiser R, Brun-Vézinet F, Clotet B, Luca AD, et al. European guidelines on the clinical management of HIV-1 tropism testing. The Lancet Infectious Diseases. 2011; 11(5):394—407. https://doi.org/10.1016/S1473-3099(10)70319-4 PMID: 21429803

81. Yoon H, Macke J, West AP, Foley B, Bjorkman PJ, Korber B, et al. CATNAP: a tool to compile, analyze and tally neutralizing antibody panels. Nucleic Acids Research. 2015; 43(W1):W213–9. https://doi.org/10.1093/nar/gkv404 PMID: 26044712

82. R Core Team. R: A Language and Environment for Statistical Computing; 2013. Available from: http://www.R-project.org/.

83. Sonnenburg S, Rätsch G, Henschel S, Widmer C, Behr J, Zien A, et al. The SHOGUN Machine Learning Toolbox. J Mach Learn Res. 2010; 11:1799–1802.

A.2. Paper 2 - bNAb efficacy study I

Copyright clearance

The herein included manuscript is the pre-copyedited, author-produced version of the paper "HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1", which was accepted for publication in *Science* following peer review. The version of record is published in *Science 352.6288, pp. 997–1001* and is available online at link https://www.science.org/doi/abs/10.1126/science.aaf0.

Title: HIV-1 Therapy with Monoclonal Antibody 3BNC117 Elicits Host Immune Responses against HIV-1

Authors: Till Schoofs[1]*, Florian Klein[1,2,3]*, Malte Braunschweig[1,4]*, Edward F. Kreider[5], Anna Feldmann[6], Lilian Nogueira[1], Thiago Oliveira[1], Julio C. C. Lorenzi[1], Erica H. Parrish[5], Gerald H. Learn[5], Anthony P. West, Jr[7], Pamela J. Bjorkman[7], Sarah J. Schlesinger[1], Michael S. Seaman[8], Julie Czartoski[9], M. Juliana McElrath[9], Nico Pfeifer[6], Beatrice H. Hahn[5], Marina Caskey[1], and Michel C. Nussenzweig[1, 10]

Affiliations:

[1]Laboratory of Molecular Immunology, and [10]Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10065, USA

[2]Laboratory of Experimental Immunology, Center for Molecular Medicine Cologne (CMMC), University of Cologne, 50931 Cologne, Germany

[3]Department I of Internal Medicine, Center of Integrated Oncology Cologne-Bonn, University Hospital Cologne, 50937 Cologne, Germany

[4]Albert Ludwigs University of Freiburg, Freiburg, Germany

[5]Departments of Medicine and Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[6]Department of Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Campus E14, 66123 Saarbrücken, Germany

[7]Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA

[8]Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, USA

[9]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

*These authors contributed equally

Correspondence should be addressed to: nussen@rockefeller.edu

Abstract:

3BNC117 is a broad and potent anti-HIV-1 neutralizing antibody that targets the CD4 binding site on the viral envelope spike. When administered passively, this antibody can prevent infection in animal models and suppress viremia in HIV-1-infected individuals. Here we report that HIV-1 immunotherapy with a single injection of 3BNC117 impacts host antibody responses in viremic subjects. In comparison to untreated controls that showed little change in their neutralizing activity over a six-month period, 3BNC117 infusion significantly improved neutralizing responses to heterologous tier 2 viruses in nearly all study participants. We conclude that 3BNC117-mediated immunotherapy enhances host humoral immunity to HIV-1.

Main Text:

Development of serum neutralization breadth during HIV-1 infection typically occurs several years after infection and is seen as a continuum with ~50% of HIV-1-infected individuals developing some level of broad neutralization and a small fraction of individuals acquiring serum neutralizing activity of extraordinary breadth and potency (*1-4*). Antibody cloning experiments revealed that this activity is due to one or more potent broadly neutralizing antibodies (bNAbs) that target one or more epitopes on the viral spike protein, gp160 (*1, 5-10*).

bNAbs show exceptional breadth and potency *in vitro*, and can protect against or suppress active infection in humanized mice (*11-13*) and macaques (*14, 15*). Moreover, in a phase I clinical trial, a single injection of 3BNC117, a CD4-binding-site specific bNAb (*6*) was safe and effective in suppressing HIV-1 viremia by an average of 1.48 logs (*16*).

In addition to direct effects on target cells and pathogens, antibody-mediated immunotherapies have the potential to engage the host immune system and induce both innate and adaptive immune responses (*17*). In particular the Fc domains of antibodies interact with receptors on innate cells such as natural killer (NK) cells and phagocytes to enhance the clearance of viral particles and the killing of infected cells (*18*). To test the hypothesis that bNAb-mediated immunotherapy can enhance immunity to HIV-1 in humans, we examined the serologic responses to the virus in individuals who received 3BNC117.

A single 3BNC117 infusion was administered to HIV-1-infected individuals at doses of 1, 3, 10, or 30 mg/kg (Fig. 1A, Table S1A) (*16, 19*). To determine whether 3BNC117 therapy is associated with changes in viral sensitivity and serologic responses to autologous viruses, we cultured HIV-1 from peripheral blood mononuclear cells (PBMCs) of 9 viremic individuals before (d0) and 4 weeks (wks) after 3BNC117 infusion (*16*). On d0, all but one of the cultured viruses were sensitive to 3BNC117 with IC_{50} values ranging from 0.09 - 8.8 µg/ml (Fig. 1B and (*16*)). At wk 4, we found increased resistance to 3BNC117 in most individuals indicating selection for viral escape variants (Fig. 1B and (*16*)).

When the same viral isolates were tested for sensitivity to the matched individual's immunoglobulins (IgG) obtained before (d0) and 24 wks after 3BNC117 infusion (Fig. 1A), we found increased neutralizing activity in the wk 24 IgG against both d0 and wk 4 autologous viruses (p=0.0078, Fig. 1C, Table S2). Thus, while 3BNC117 infusion selected for 3BNC117-resistant HIV-1 variants, neutralizing antibody responses continued to develop against autologous viruses (*20*).

To test for changes in heterologous neutralizing activity following 3BNC117 treatment, we assayed patients' d0 and wk 24 IgG against a panel of tier 1 (n=1) and tier 2 (n=12) HIV-1 pseudoviruses that included globally circulating HIV-1 strains (*21*) (Fig. 2, Table S1, S3, S4). Neutralizing activity was compared between the two time points by measuring the area under the neutralization curve for subjects' isolated IgG against each

virus (AUC) (Table S4B). 15 subjects that received 3BNC117 were not on anti-retroviral therapy (ART) and had starting viral loads from 640 - 53,470 copies/ml (Table S1A). Control IgGs were obtained from 36 viremic individuals who did not receive 3BNC117 and had starting viral loads ranging from 150 – 303,200 copies/ml (Fig. 2, Table S1B).

During a 6-month observation period, control individuals' neutralizing activity showed no consistent improvement in either breadth or potency (Fig. 2A and B, S1A, S2, Table S4, S5A) (*4, 22*). In contrast, all but one of the 15 viremic individuals infused with 3BNC117 showed increased breadth and/or potency against the pseudovirus panel at wk 24 ($p=7.1 \times 10^{-7}$, Fig. 2A, S1B, S2, Table S4, S5, S6). The absolute change in neutralizing activity varied between viruses and individuals, ranging from small effects to dramatic increases as observed in patient 2A3 for viral strain Q769.d22 (Fig. 2C, Table S4, S5, S6). Significant differences were also evident between 3BNC117-treated and control groups regardless whether sera from all individuals were considered in aggregate, or examined against individual viruses ($p=1.9 \times 10^{-9}$, Fig. 2B, D).

In addition to viremic subjects, we examined 12 individuals that received 3BNC117 while on ART, with no detectable or low-level viremia (<20 - 100 copies/ml). In comparison to viremic subjects, the increase in heterologous neutralizing activity was significantly less pronounced in ART-treated individuals ($p=0.037$, Fig. 2A, B, and D, S1B, S2, Table S3, S4, S5).

The observed improvement in neutralizing activity could not be explained by confounding factors such as differences in initial viral load or $CD4^+$ T cell levels (Fig. S3, Table S1, S7). Moreover, we found no correlation between d0 neutralizing activity and neutralization improvement (Fig. S4). A comparison of the pattern of neutralization increase with 3BNC117's neutralization profile ruled out that remaining antibody was responsible for the effect (Fig. S5, Table S8). We conclude that 3BNC117 enhances host immunity to heterologous tier 2 HIV-1 viruses irrespective of initial neutralization breadth and potency.

To examine the effects of 3BNC117 immunotherapy on the plasma viral population of treated individuals, we performed single genome sequencing (SGS) of over 1,000 plasma-derived gp160 *env* genes (gp160) before (d0) and 4 (6), 12, or 24 wks after infusion (Fig. 3A, B and S6-S10, Table S9). With the exception of two individuals who were sexual partners, all other volunteers had epidemiologically unrelated infections (Fig. 3A). On d0, *env* sequences from subjects 2A1, 2A3, and 2C4 comprised multiple lineages, which was reflected in a multimodal distribution of pairwise diversity measurements from these individuals (Fig. 3B, S6). Analysis of *env* sequences from subsequent time points revealed significant shifts in both nucleotide (6 out of 9 individuals, Fig. 3B) and amino acid sequence diversity (7 out of 9 individuals, Fig. S6). Consistent with the observation that *env* diversity is associated with neutralization breadth (*23-25*), there was a strong correlation between the initial level of neutralizing activity and the initial diversity of the circulating viral swarm ($R^2 = 0.92$, Fig. 3C).

We next evaluated viral sequence evolution in each of the 3BNC117-treated subjects over time. Shifts in the viral quasispecies were evident regardless of initial 3BNC117 neutralization sensitivity and bNAb dose (Fig. 4, S7). However, the nature of these shifts differed depending on the subject (Fig. 4, S7-S9). For example, in subject 2A1, 15/27 d0 sequences fell into a single clade marked "group A" (Fig. 4A, S8). Four weeks following 3BNC117 infusion group A viruses contracted (2/25 sequences) and group C viruses expanded (16/25). At wk 24, the viral quasispecies was primarily comprised of group B and D viruses (Fig. 4, S8). This pattern of "clade shifting" was also seen in subjects 2A3 and 2C4 (Fig. S7). Subjects with lower initial *env* diversities, such as 2E1, did not harbor distinct viral sublineages at d0 (Fig. 3, 4A), but continued to accrue mutations some of which became fixed during the 24-week follow-up (*e.g.* changes in V1/V2 in 2E1, Fig. S9).

To assess viral sequence changes following 3BNC117 infusion, we generated longitudinal logo plots depicting 3BNC117 contact residues (*26*, *27*) for each subject (Fig. 4B, S7, S10). While viruses from all nine subjects exhibited mutations within 3BNC117 contact residues relative to the d0 consensus sequence, their number and position varied considerably as exemplified by subjects 2A1 and 2E1 (Fig. 4B, Fig. S7, S10). Using LASSIE (Longitudinal Antigenic Sequences and Sites from Intrahost Evolution) (*28*), we scanned the entire *env* protein sequence for sites selected within the 24 wk time frame (selection cutoff ≥80%) (Table S10). While selected sites were identified in all subjects, no consistent mutational pattern was observed (Table S10). These data suggest that 3BNC117 immunotherapy is associated with shifts in circulating

quasispecies and a number of different *env* mutations, some of which persist even after the infused antibody levels drop below detection.

To better understand the virus host-interactions that led to the development of enhanced heterologous neutralizing breadth, we performed neutralization assays on 63 pseudoviruses expressing the gp160s found in the circulation on d0, wk 4, 12 and 24 from 5 individuals (Fig. 4, S7, Table S11). The pseudoviruses were tested for sensitivity to the corresponding individual's IgG obtained on d0 and wk 24. In all cases, we were able to identify d0 or wk 4 viruses that exhibited greater neutralization sensitivity to wk 24 IgG compared to d0 IgG (Fig. 4, S7, Table S11). For example, all tested 2A1 and 2E1 viruses were 3BNC117 sensitive and exhibited a wk 24/d0 fold change of ~1.7 and ~4.8 in IgG IC_{50} respectively (Fig. 4). On the other hand, all tested 2C4 viruses were 3BNC117-resistant (mean IC_{50}: >20 μg/ml), yet they were ~6.5-fold more sensitive to wk 24 IgG versus d0 IgG (Fig. S7). In conclusion, viremic individuals receiving 3BNC117 produced antibodies to autologous viruses that were both sensitive and resistant to 3BNC117.

While exceptional broadly neutralizing antibodies to HIV-1 develop only sporadically in a fraction of infected individuals, most HIV-1 infected individuals develop some level of neutralization breadth (*1-4*). Here we show that 3BNC117 immunotherapy accelerates this process. This boost in heterologous breadth occurs irrespective of demographic, virologic, or dosage factors and was associated with both transient and lasting changes to the viral quasi-species. Of note, neutralization improvements observed were modest in

most individuals, potentially owing to the transient nature of therapy with a single antibody as well as the short timeframe of observation.

Although the effect of 3BNC117 on neutralizing responses to heterologous HIV-1 viruses may seem surprising, anti-HIV-1 antibodies have been associated with enhanced immunity in infants born to HIV-1-infected mothers that have circulating anti-HIV-1 antibodies and macaques treated with monoclonal antibodies or neutralizing serum (*29-31*).

How passively administered antibodies to HIV-1 accelerate the emergence of bNAbs is not completely understood. One possibility is that 3BNC117 infusion selected for viral variants with altered antigenic properties, which in turn stimulated new B cell lineages (*23-25, 32-34*). A second possibility is that immune complexes formed by 3BNC117 and circulating viruses act as potent immunogens, a phenomenon that is believed to be responsible for the enhanced $CD8^+$ T cell immunity to tumor antigens in individuals receiving monoclonal antibody based immunotherapy (*35-37*).

Irrespective of the mechanism(s), the enhanced antibody response found in individuals receiving 3BNC117 therapy indicates that immunotherapy boosts host immunity to HIV-1. Moreover, the finding that antibody responses to heterologous tier 2 viruses develop in nearly all 3BNC117-treated individuals suggests that host genetics or a specific viral envelope sequence do not limit the development of neutralizing antibodies to HIV-1.

References and Notes:

1. F. Klein *et al.*, Antibodies in HIV-1 vaccine development and therapy. *Science* **341**, 1199-1204 (2013).
2. A. P. West, Jr. *et al.*, Structural insights on the role of antibodies in HIV-1 vaccine and therapy. *Cell* **156**, 633-648 (2014).
3. I. Mikell *et al.*, Characteristics of the earliest cross-neutralizing antibody response to HIV-1. *PLoS pathogens* **7**, e1001251 (2011).
4. P. Hraber *et al.*, Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection. *AIDS* **28**, 163-169 (2014).
5. X. Wu *et al.*, Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329**, 856-861 (2010).
6. J. F. Scheid *et al.*, Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333**, 1633-1637 (2011).
7. L. M. Walker *et al.*, Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477**, 466-470 (2011).
8. F. Klein *et al.*, Broad neutralization by a combination of antibodies recognizing the CD4 binding site and a new conformational epitope on the HIV-1 envelope protein. *J Exp Med* **209**, 1469-1479 (2012).
9. M. Bonsignori *et al.*, Two Distinct Broadly Neutralizing Antibody Specificities of Different Clonal Lineages in a Single HIV-1-infected Donor: Implications for Vaccine Design. *Journal of virology*, (2012).
10. J. F. Scheid *et al.*, Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458**, 636-640 (2009).
11. F. Klein *et al.*, HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature* **492**, 118-122 (2012).
12. J. A. Horwitz *et al.*, HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 16538-16543 (2013).
13. J. Pietzsch *et al.*, A mouse model for HIV-1 entry. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 15859-15864 (2012).
14. D. H. Barouch *et al.*, Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys. *Nature* **503**, 224-228 (2013).
15. M. Shingai *et al.*, Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. *Nature* **503**, 277-280 (2013).
16. M. Caskey *et al.*, Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature* **522**, 487-491 (2015).
17. S. Bournazos, J. V. Ravetch, Fcgamma receptor pathways during active and passive immunization. *Immunol Rev* **268**, 88-103 (2015).
18. C.-L. Lu, Enhanced clearance of HIV-1-infected cells by anti-HIV-1 broadly neutralizing antibodies in vivo. *Science*, (2016).
19. Materials and methods are available as supplementary materials on Science Online.

20. X. Wei *et al.*, Antibody neutralization and escape by HIV-1. *Nature* **422**, 307-312 (2003).
21. A. deCamp *et al.*, Global panel of HIV-1 Env reference strains for standardized assessments of vaccine-elicited neutralizing antibodies. *Journal of virology* **88**, 2489-2507 (2014).
22. S. G. Deeks *et al.*, Neutralizing antibody responses against autologous and heterologous viruses in acute versus chronic human immunodeficiency virus (HIV) infection: evidence for a constraint on the ability of HIV to completely evade neutralizing antibody responses. *Journal of virology* **80**, 6155-6164 (2006).
23. H. X. Liao *et al.*, Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* **496**, 469-476 (2013).
24. N. A. Doria-Rose *et al.*, Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* **509**, 55-62 (2014).
25. P. L. Moore, C. Williamson, L. Morris, Virological features associated with the development of broadly neutralizing antibodies to HIV-1. *Trends Microbiol* **23**, 204-211 (2015).
26. T. Zhou *et al.*, Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329**, 811-817 (2010).
27. F. Klein *et al.*, Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell* **153**, 126-138 (2013).
28. P. Hraber *et al.*, Longitudinal Antigenic Sequences and Sites from Intra-Host Evolution (LASSIE) Identifies Immune-Selected HIV Variants. *Viruses* **7**, 5443-5475 (2015).
29. L. Goo, V. Chohan, R. Nduati, J. Overbaugh, Early development of broadly neutralizing antibodies in HIV-1-infected infants. *Nat Med* **20**, 655-658 (2014).
30. N. L. Haigwood *et al.*, Passive immunotherapy in simian immunodeficiency virus-infected macaques accelerates the development of neutralizing antibodies. *Journal of virology* **78**, 5983-5995 (2004).
31. C. T. Ng *et al.*, Passive neutralizing antibody controls SHIV viremia and enhances B cell responses in infant macaques. *Nat Med* **16**, 1117-1119 (2010).
32. P. L. Moore *et al.*, Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. *Nat Med* **18**, 1688-1692 (2012).
33. F. Gao *et al.*, Cooperation of B cell lineages in induction of HIV-1-broadly neutralizing antibodies. *Cell* **158**, 481-491 (2014).
34. J. N. Bhiman *et al.*, Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies. *Nat Med* **21**, 1332-1336 (2015).
35. T. T. Wang *et al.*, Anti-HA Glycoforms Drive B Cell Affinity Selection and Determine Influenza Vaccine Efficacy. *Cell* **162**, 160-169 (2015).
36. S. Bournazos, D. J. DiLillo, J. V. Ravetch, The role of Fc-FcgammaR interactions in IgG-mediated microbial neutralization. *J Exp Med* **212**, 1361-1369 (2015).

37. D. J. DiLillo, J. V. Ravetch, Differential Fc-Receptor Engagement Drives an Anti-tumor Vaccinal Effect. *Cell* **161**, 1035-1045 (2015).
38. S. J. Ratcliffe, J. Shults, GEEQBOX: A MATLAB toolbox for generalized estimating equations and quasi-least squares. *J Stat Softw* **25**, 1-14 (2008).
39. P. B. Gilbert, A. J. Rossini, R. Shankarappa, Two-sample tests for comparing intra-individual genetic sequence diversity between populations. *Biometrics* **61**, 106-117 (2005).
40. E. E. Giorgi, T. Bhattacharya, A note on two-sample tests for comparing intra-individual genetic sequence diversity between populations. *Biometrics* **68**, 1323-1326; author reply 1326 (2012).
41. W. Deng *et al.*, DIVEIN: a web server to analyze phylogenies, sequence divergence, diversity, and informative sites. *Biotechniques* **48**, 405-408 (2010).
42. A. B. van 't Wout, H. Schuitemaker, N. A. Kootstra, Isolation and propagation of HIV-1 on peripheral blood mononuclear cells. *Nature protocols* **3**, 363-370 (2008).
43. M. S. Seaman *et al.*, Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. *Journal of virology* **84**, 1439-1452 (2009).
44. M. Li *et al.*, Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *Journal of virology* **79**, 10108-10125 (2005).
45. A. P. West, Jr. *et al.*, Computational analysis of anti-HIV-1 antibody neutralization panel data to identify potential functional epitope residues. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 10598-10603 (2013).
46. S. Kryazhimskiy, D. P. Rice, E. R. Jerison, M. M. Desai, Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* **344**, 1519-1522 (2014).
47. M. A. Larkin *et al.*, Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
48. M. Kearse *et al.*, Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647-1649 (2012).
49. D. Darriba, G. L. Taboada, R. Doallo, D. Posada, jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* **9**, 772 (2012).
50. S. Guindon *et al.*, New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology* **59**, 307-321 (2010).
51. D. C. Nickle *et al.*, HIV-specific probabilistic models of protein evolution. *PloS one* **2**, e503 (2007).
52. J. L. Kirchherr *et al.*, High throughput functional analysis of HIV-1 env genes without cloning. *J Virol Methods* **143**, 104-111 (2007).

Acknowledgments:

We thank all study participants who devoted time to our research. We thank the Rockefeller University Hospital Clinical Research Support Office, the nursing staff for patient care and recruitment, the clinical study group of the Infectious Disease Division at the University Hospital Cologne, and all members of the Nussenzweig lab for helpful discussions. We thank M. Schechter and C. Baro for technical assistance, P. Fast and H. Park for Clinical monitoring, E. Gotschlich and B. Coller for input on study design and P. Hraber for helping with LASSIE analyses. The data reported in this study are tabulated in the main paper and in the supplementary material. Envelope single genome sequencing data can be downloaded from GenBank (Accession numbers KX027737 – KX028736). This work was supported in part by the Bill and Melinda Gates Foundation Collaboration for AIDS Vaccine Discovery (CAVD) Grants OPP1032144 (M.S.S.), OPP1092074 and OPP1124068 (M.C.N), the Robertson Foundation to M.C.N., NIH Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID) 1UM1 AI100663-01 (M.C.N) and 1UM1 AI00645 (B.H.H.), the University of Pennsylvania Center for AIDS Research (CFAR) Single Genome Amplification Service Center P30 AI045008 (B.H.H.), NIH grants UM1AI068618 (MJM), UM1AI069481 (MJM), F30 AI112426 (E.F.K), and HIVRAD P01 AI100148 (P.J.B.). T.S. is supported by a Deutsche Forschungsgemeinschaft postdoctoral fellowship (SCHO 1612/1-1). F.K. is supported by the Heisenberg-Program of the DFG (KL 2389/2-1), the European Research Council (ERC-StG639961), and the German Center for Infection Research (DZIF), partner site Bonn-Cologne, Cologne, Germany. M.B. is supported by the German National Academic Foundation. J.C.L. is supported by an award from CNPq "Ciencia sem Fronteiras" Brazil

(248676/2013-0). M.C.N. is a Howard Hughes Medical Investigator and an inventor on U.S Patent Application No. 14/118,496 filed by Rockefeller University related to 3BNC117.

Figure 1: *Virus sensitivity to 3BNC117 and autologous antibody responses.* **A.** Graph displays kinetics of 3BNC117 antibody decay in HIV-1-infected individuals as determined by a validated anti-idiotype ELISA (*16*). Shown are mean values of patients infused in each respective dose group. Each patient sample was measured in duplicates. Gray shaded area indicates lower level of accuracy of the assay (2 μg/ml). Red arrows indicate the timepoints of IgG purification. **B.** Autologous virus sensitivity to 3BNC117 before (day 0, grey) and 4 wks (black) after 3BNC117 infusion. Y-axis shows IC_{50}s for 3BNC117 on viral culture supernatants from PBMCs determined by TZM.bl assay. Neutralization assays performed in duplicates. **C.** Graph shows the AUC of the neutralization curves of purified IgGs obtained from sera on day 0 (orange) or wk 24 (green) against day 0 (left panel) or wk 4 (right panel) autologous viruses. Neutralization assays performed in duplicates. p-values determined by Wilcoxon signed-rank test.

Figure 2: *Heterologous antibody responses*. **A.** Graph shows the difference in overall AUC (mean AUC change) per individual in TZM.bl assays against 13 heterologous viruses (see **2D**) for day 0 vs. wk 24 IgG obtained from 36 untreated viremic controls (mean sampling interval: 26.8 wks), 15 viremic individuals infused with 3BNC117 (mean sampling interval: 24.1 wks), and 12 ART-treated individuals receiving 3BNC117 infusion (mean sampling interval: 24.0 wks) (*16*). Neutralization assays performed in

duplicates. p-values determined by unpaired Wilcoxon test (rank-sum test). **B.** Graph shows the aggregated differences in AUC between d0 and wk 24 IgG assayed by TZM.bl for all viruses and all individuals. Each dot represents a single AUC difference for a single virus from one individual displayed in **A**. Colored bars represent mean of all AUCs. Whiskers show standard deviation. p-values determined using generalized estimating equations (*38*). **C**. Graph shows 3BNC117 antibody levels (ELISA, white) and TZM.bl neutralization titer (green) against tier 2 strain Q769.d22 in subject 2A3 **D.** Mean AUCs of IgGs of all individuals at d0 (grey) and wk 24 (color of respective group) for each HIV-1 pseudovirus tested. Changes in neutralization of viremic control individuals without 3BNC117 infusion are shown in yellow (left). Change in neutralization of 3BNC117-treated individuals shown in dark (off ART, middle) and light blue (on ART, right)**.** p-values determined using unpaired Wilcoxon test (rank-sum test). Red stars indicate significant p-values after Bonferroni-correction (threshold $p < 0.0038$).

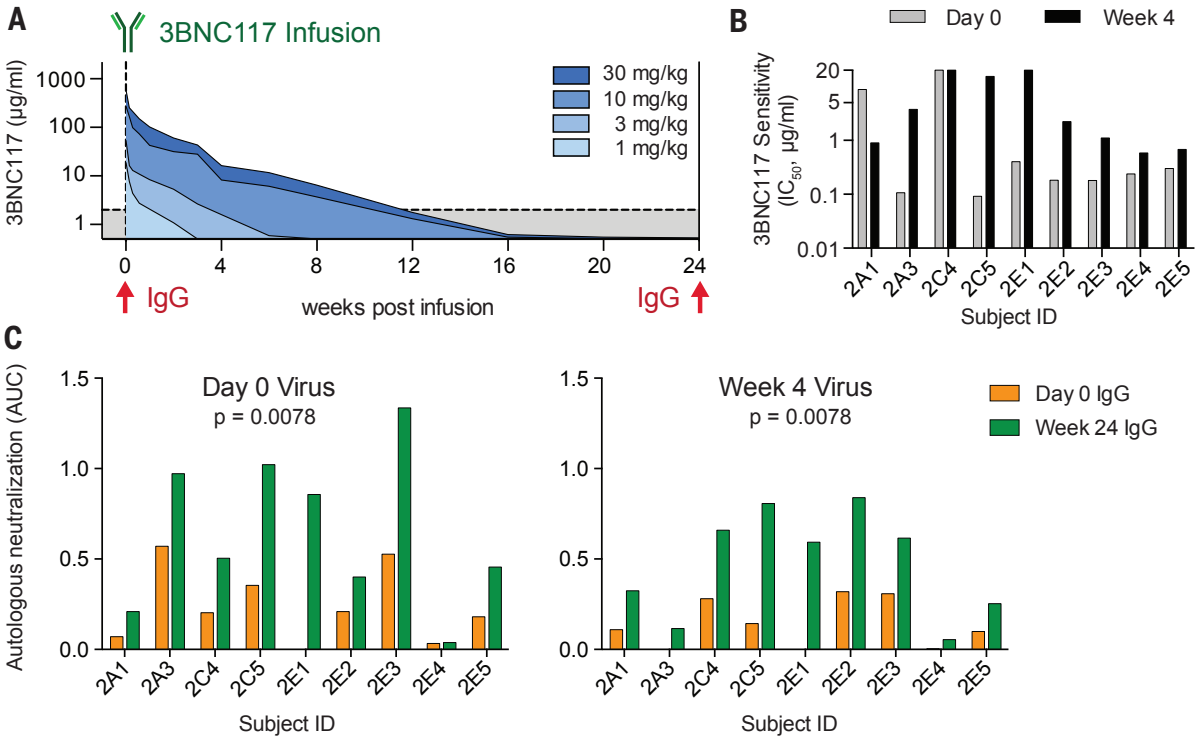
Figure 3: *HIV-1 quasispecies diversity before and after 3BNC117 infusion*. **A.** Maximum likelihood phylogenetic tree of single genome-derived *env* gene sequences from d0 plasma, before therapy with 3BNC117 (Table S9). Asterisks indicate bootstrap values of 100%. Individual viral sequences are color coded as indicated. **B.** Scatter plots depicting pairwise nucleotide sequence diversity of plasma *env* sequences on d0, and wk 4 (2E5, wk 6), 12 and 24 after infusion. Each dot represents the pairwise genetic difference between two sequences at a given timepoint. Colored bars indicate median diversity, while black bars indicates the interquartile range. p-values were determined using a two-sample U-statistic based Z-test (*39-41*). **C.** Graph shows the relationship between d0

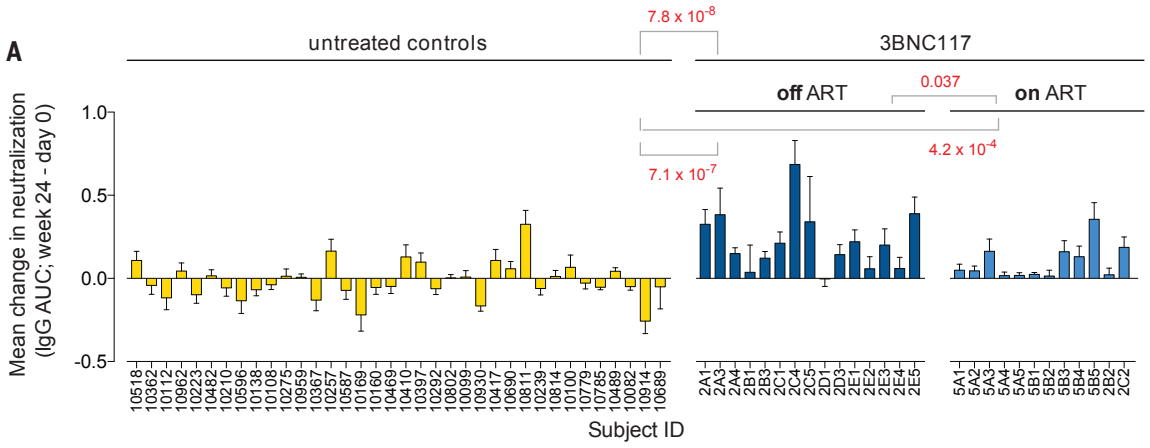
mean heterologous neutralizing AUC against a panel of tier 1 (n=1) and tier 2 (n=12) viruses (x-axis) and the median pairwise nucleotide diversity for each patient (y-axis).

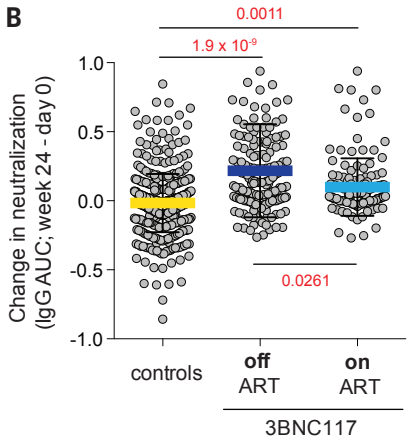
Figure 4: *Antibody responses to the evolving viral quasispecies*. **A.** Maximum-likelihood phylogenetic trees of single genome-derived *env* gene sequences from subjects 2A1 and 2E1 sampled on d0 and wk 4, 12, and 24 after 3BNC117 infusion (left). Clades with bootstrap support ≥ 70% are indicated by a black star and are arbitrarily named groups A-D in case of subject 2A1. Bar graphs (middle) indicate the timepoints from which sequences in the tree are derived. Heat maps (right) show the 3BNC117 IC_{50}, d0 IgG IC_{50} and wk 24 IgG IC_{50} values against autologous pseudoviruses using *env* sequences as indicated by colored stars. Neutralization assays performed in duplicates. **B.** Sequence logo plots illustrating longitudinal amino acid changes in and around known 3BNC117-contact residues (*26, 27*) in subject 2A1 and 2E1. Letters indicate deviations from the d0 consensus shown at the top, whereas white spaces indicate agreement with the d0 consensus. Colors indicate basic (dark blue) and acidic (red) residues and a turquoise "O" is used instead of "N" to indicate a potential N-glycosylation site. Logo plots were generated using LASSIE (*28*). + symbols indicate 3BNC117 contact residues confirmed by two crystal structures (*26*, *27*).

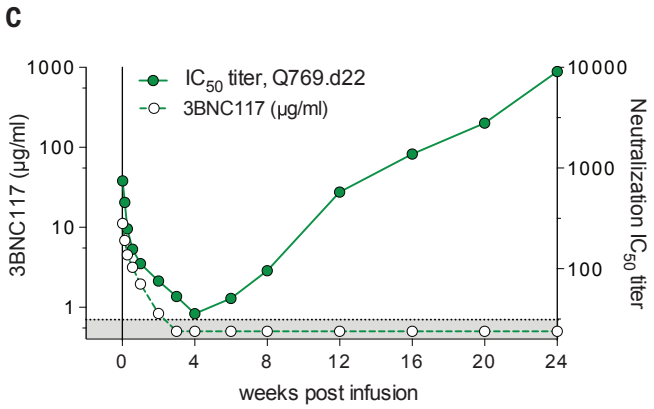
Supplementary Materials

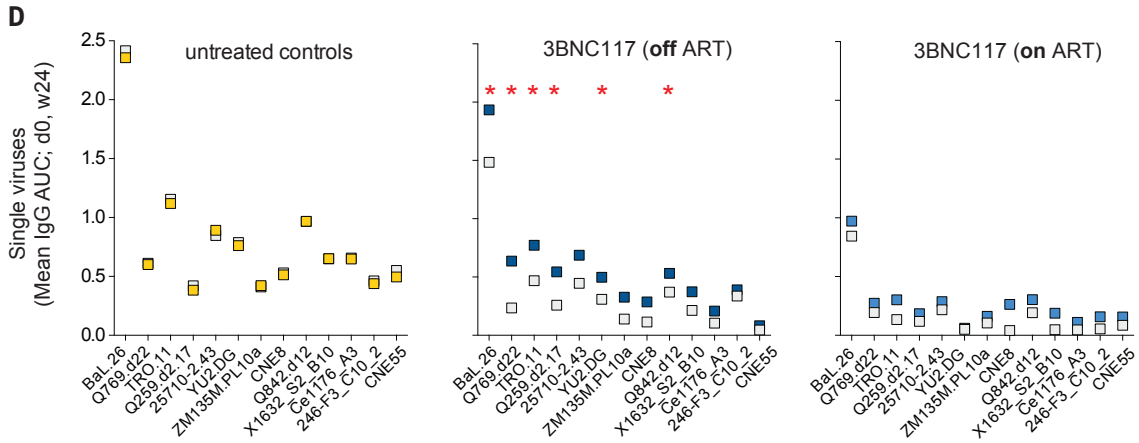
Supplementary Tables

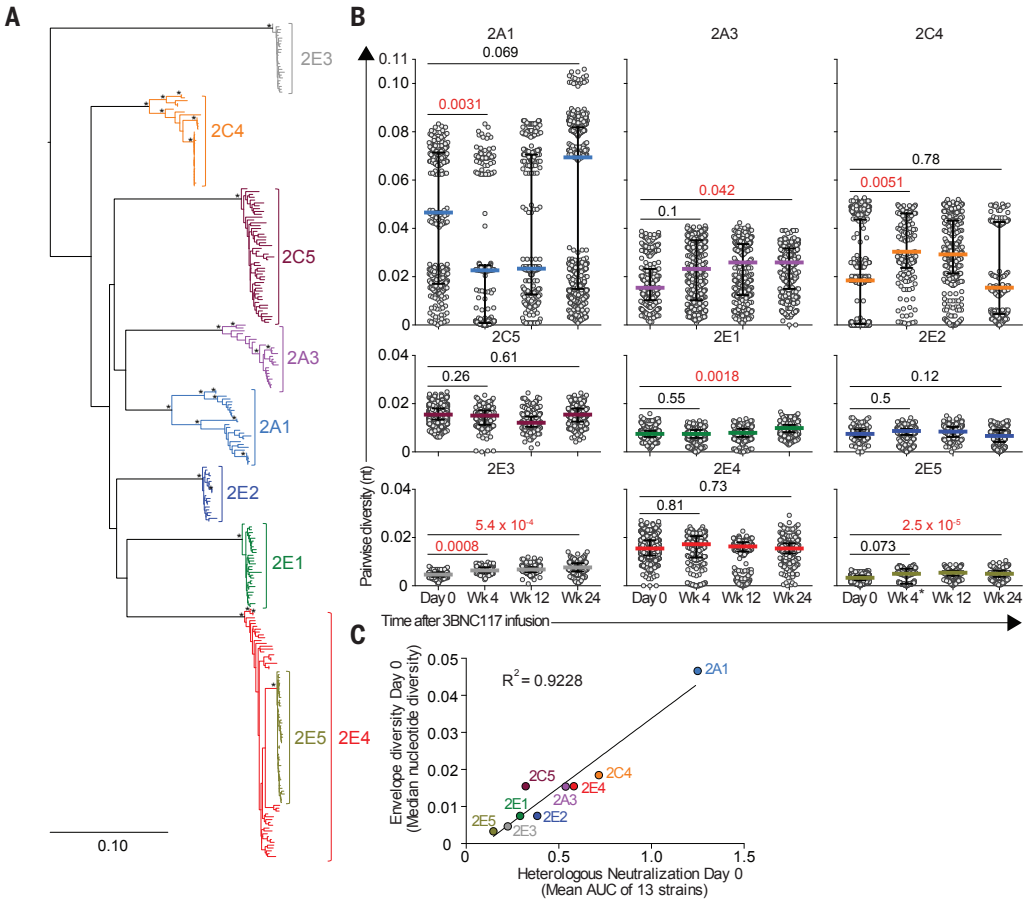

A
3BNC117 Infusion
3BNC117 (µg/ml)
1000
100
10
1
30 mg/kg
10 mg/kg
3 mg/kg
1 mg/kg
0
4
8
12
16
20
24
IgG
weeks post infusion
IgG
B
Day 0
Week 4
3BNC117 Sensitivity
(IC50, µg/ml)
20
5
1
0.1
0.01
2A1
2A3
2C4
2C5
2E1
2E2
2E3
2E4
2E5
Subject ID
C
Autologous neutralization (AUC)
Day 0 Virus
p = 0.0078
Week 4 Virus
p = 0.0078
Day 0 IgG
Week 24 IgG
1.5
1.0
0.5
0.0
Subject ID

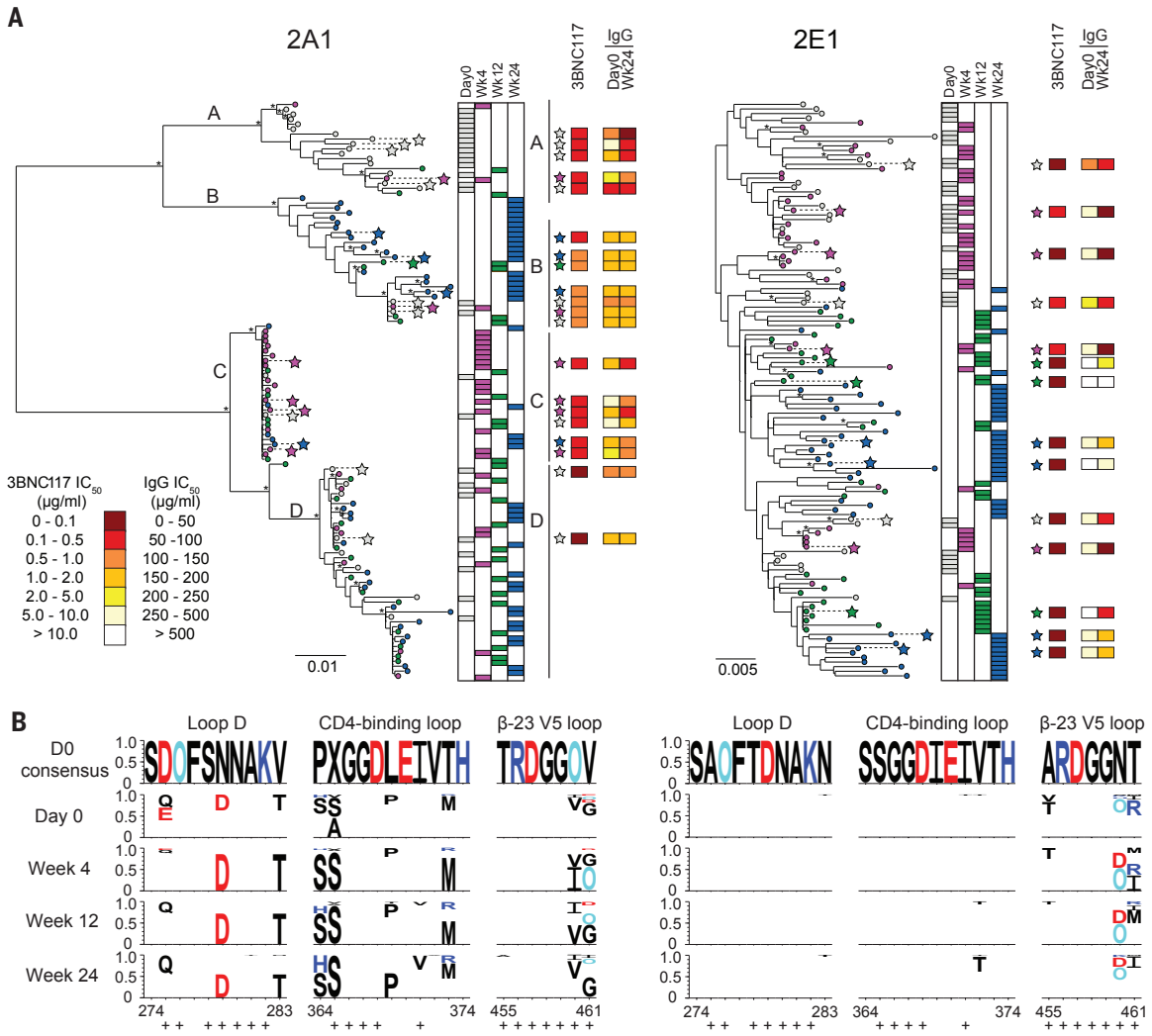

A
untreated controls
3BNC117
7.8 x 10-8
off ART
on ART
0.037
4.2 x 10-4
7.1 x 10-7
Mean change in neutralization (IgG AUC; week 24 - day 0)
1.0
0.5
0.0
-0.5
10518 10362 10112 10962 10223 10482 10210 10596 10138 10108 10275 10959 10367 10257 10587 10169 10160 10469 10410 10397 10292 10802 10099 10930 10417 10690 10811 10239 10814 10100 10779 10785 10489 10082 10914 10689
2A1 2A3 2A4 2B1 2B3 2C1 2C4 2C5 2D1 2D3 2E1 2E2 2E3 2E4 2E5
5A1 5A2 5A3 5A4 5A5 5B1 5B2 5B3 5B4 5B5 2B2 2C2
Subject ID


B
0.0011
1.9 x 10-9
0.0261
Change in neutralization (IgG AUC; week 24 - day 0)
1.0
0.5
0.0
-0.5
-1.0
controls
off ART
on ART
3BNC117


C
IC50 titer, Q769.d22
3BNC117 (µg/ml)
3BNC117 (µg/ml)
1000
100
10
1
Neutralization IC50 titer
10000
1000
100
0
4
8
12
16
20
24
weeks post infusion


D
Single viruses (Mean IgG AUC; d0, w24)
2.5
2.0
1.5
1.0
0.5
0.0
untreated controls
3BNC117 (off ART)
3BNC117 (on ART)
BaL.26 Q769.d22 TRO.11 Q259.d2.17 25710-2.43 YU2.DG ZM135M.PL10a CNE8 Q842.d12 X1632_S2_B10 Ce1176_A3 246-F3_C10_2 CNE55


A
2E3
2C4
2C5
2A3
2A1
2E2
2E1
2E5
2E4
0.10
B
2A1
2A3
2C4
0.069
0.0031
0.042
0.1
0.78
0.0051
2C5
2E1
2E2
0.61
0.26
0.0018
0.55
0.12
0.5
2E3
2E4
2E5
0.73
0.81
5.4 x 10-4
0.0008
2.5 x 10-5
0.073
Pairwise diversity (nt)
Day 0
Wk 4
Wk 12
Wk 24
Wk 4*
Time after 3BNC117 infusion
C
R2 = 0.9228
Envelope diversity Day 0
(Median nucleotide diversity)
Heterologous Neutralization Day 0
(Mean AUC of 13 strains)


A
2A1
2E1
Day0
Wk4
Wk12
Wk24
3BNC117
IgG
Day0 Wk24
A
B
C
D
3BNC117 IC50 (μg/ml)
0 - 0.1
0.1 - 0.5
0.5 - 1.0
1.0 - 2.0
2.0 - 5.0
5.0 - 10.0
> 10.0
IgG IC50 (μg/ml)
0 - 50
50 -100
100 - 150
150 - 200
200 - 250
250 - 500
> 500
0.01
0.005
B
Loop D
CD4-binding loop
β-23 V5 loop
D0 consensus
Day 0
Week 4
Week 12
Week 24
274
283
364
374
455
461

Supplementary Materials for

HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1

Till Schoofs, Florian Klein, Malte Braunschweig, Edward F. Kreider, Anna Feldmann, Lilian Nogueira, Thiago Oliveira, Julio C. C. Lorenzi, Erica H. Parrish, Gerald H. Learn, Anthony P. West, Jr, Pamela J. Bjorkman, Sarah J. Schlesinger, Michael S. Seaman, Julie Czartoski, M. Juliana McElrath, Nico Pfeifer, Beatrice H. Hahn, Marina Caskey and Michel C. Nussenzweig

correspondence to: nussen@rockefeller.edu

This PDF file includes:

Materials and Methods

Study participants

Study participants that received 3BNC117 were all part of the 3BNC117 Phase-I open-label clinical trial NCT02018510 (*16*). The protocol was approved by the Federal Drug Administration in the USA, the Paul Ehrlich Institute in Germany and the Institutional Review Boards at the Rockefeller University, the University of Cologne, Weill Cornell Medical College and the Brigham and Women's Hospital Boston. The control cohort of untreated viremic HIV-1-infected subjects were study participants in Seattle Vaccine Unit Observational Protocols 4712 and 4325 (P.I. MJ McElrath). Subjects were recruited from the Seattle, WA community through advertisements and referrals from care providers. All participants provided written informed consent before participation in the studies and the studies were conducted in accordance with Good Clinical Practice.

ELISA of 3BNC117 serum levels

3BNC117 serum concentrations were determined using a validated sandwich ELISA as previously described (*16*). Lower limit of accuracy of the ELISA assay was determined to be 2 μg/ml. Patients for whom background > 0.50 μg/ml at Day 0 baseline was detected (Subjects 2A4, 2B3, 2C2, 2C4, 2C1 and 2E2) were excluded from the illustration of 3BNC117 serum levels shown in Figure 1A.

Virus cultures

Virus from study participants was obtained by co-culture of patient peripheral blood mononuclear cells (PBMCs) with healthy donor PBMCs as previously described (*42*). Healthy donor PBMCs were obtained from patients by leukapheresis under study protocol MNU-0628 at Rockefeller University. All donors provided written informed consent before participation. Healthy donor PBMCs were pre-stimulated at a density of 5 x 10^6 cells ml^{-1} in IMDM containing 10% FBS, 1% Penicillin-Streptomycin and PHA at 1 μg ml^{-1} for 2-3 days at 37°C and 5% CO_2. 6 x 10^6 of the stimulated donor PBMCs were then transferred to IMDM containing 10% FBS, 1% Penicillin-Streptomycin, 10 IU IL-2 ml^{-1} and 5 μg ml^{-1} polybrene and co-cultured with 5-10 x 10^6 patient PBMCs at 37°C and 5% CO_2. Media was replaced weekly and the presence of p24 in culture supernatant was quantified by Lenti-X p24 Rapid Titer Kit (Clontech). Cultures exceeding 1 ng/ml of p24 per ml of supernatant were frozen and stored at -80°C. Determination of tissue culture infectious dose 50 ($TCID_{50}$) and subsequent testing for sensitivity of autologous viruses to different broadly neutralizing antibodies and autologous serum IgG was carried out using a TZM.bl neutralization assay according to established protocols (*43, 44*). All neutralization assays were run in duplicates.

Antibody neutralization testing by TZM.bl

To determine heterologous neutralizing activity, patient sera or purified IgG were tested against a defined panel of pseudoviruses in TZM.bl assay as previously described (*21, 43*). IgG was purified from patient plasma using Protein G Sepharose 4 Fast Flow according to manufacturer's instructions. All neutralization assays were conducted in a laboratory meeting Good Clinical Laboratory Practice (GCLP) Quality Assurance criteria

(Michael S. Seaman, Beth Israel Deaconess Medical Center). Samples from 3BNC117-treated patients and controls were tested in intermingled batches at a variety of different time points. All neutralization assays were run in duplicates.

Antibody-strain coverage curves

Total IgG coverage curves ((1/AUC) vs. percent of strains neutralized) were generated using Antibody database v2.0 (*45*).

Area under the curve calculation of antibody neutralization

The area under the curve (AUC) was calculated using the neutralization (y-axis) at a particular measured concentration level (x-axis) using the R package flux (version 0.2.1.). In order to have equal distance between concentration levels, concentration levels were logarithmized using the natural logarithm. Since mean neutralization activity of both treated and the control group towards murine leukemia virus (MuLV) was around 7%, a neutralization of 10% was considered background noise. In addition, we tested at which concentration level the standard deviation (sd) of the assay exceeded variation due to neutralization. For neutralization by total IgG, only concentrations greater than 6.173 ug/ml were taken into consideration since the measurement variation exceeded changes in antibody neutralization below this level (i.e. concentration levels used to calculate AUCs were 500 µg/ml, 166.667 µg/ml, 55.556 µg/ml, 18.519 µg/ml and 6.173 µg/ml).

Statistical tests of heterologous neutralization AUC data

In Figure 2A, mean AUC difference values of off/on-ART patients were compared to mean AUC differences of control patients with an unpaired Wilcoxon test (rank-sum test).

For Figure 2B pairwise testing was performed if AUC difference values are different for different groups (i.e., control group, treated while off-ART, and treated while on-ART). The model was built using generalized estimating equations for a model with intercept and group as the only covariate (*38*). Equicorrelated measurements within a cluster (patient) were assumed and normal distribution was used. p-values were estimated for the null-hypothesis that the weight parameter of the group covariate is zero. GEEQBOX version 1.0 was used for calculations (*38*).

For Figure 2D, AUC difference values of controls were compared to the AUC difference values of the off-ART (panel 2) or on-ART (panel 3) 3BNC117-treated patients using an unpaired Wilcoxon test (rank-sum test). This was performed for each virus separately. A star is shown if the p-value was significant after Bonferroni correction ((p-value <= (0.05 / # tests)) = (p-value <= 0.003846154)).

To exclude that leftover amounts of 3BNC117 could explain the neutralization increase observed in 3BNC117-treated individuals, the Spearman rank correlation of neutralization changes in each patient with the neutralization profile of 3BNC117 was calculated (Fig. S5, Table S8). Rank correlation was calculated using both IC_{50} and AUC values. p-values were corrected for multiple testing by Bonferroni-correction (p-value <= (0.05 / # tests)) = (p-value <= 0.0019)).

Analysis of confounding variables

To determine whether the observed difference in AUC improvement between 3BNC117-treated individuals and viremic controls was confounded by variables such as age, time of infection and starting AUC, a systematic confounder analysis was performed. For each variable and each virus (Table S7), we tested whether the variable is predictive for the observed AUC difference. To do so, a linear regression model (intercept only/null model) was compared to an extended model that considered the variable in question. A likelihood ratio test asking whether the extended model is superior to the linear model was performed (Table S7). To compute the bayes factors for the likelihood ratio test, the R package BayesFactor (version 0.9.2) was used. A total of 8 features was tested across each of 13 viruses (total = 104 tests).

Single genome sequencing (SGS) of HIV-1 *env* genes

HIV-1-RNA was extracted from patient plasma using the Qiagen MinElute Virus Spin Kit according to manufacturer's instructions. Extracted RNA was subjected to env-specific cDNA Synthesis using SuperScript III Reverse Transcriptase and primer envB3out 5'– TTGCTACTTGTGATTGCTCCATGT 3'. Remaining RNA was digested using RNAseH for 20 minutes at 37 °C before diluted cDNA was subjected to two-rounds of nested PCR with gp160-specific primers. First round PCR was performed in a 20 ul volume containing 1x High Fidelity buffer, 2 mM MgSO4, 0.2 mM dNTPs and 0.5 units of High Fidelity Platinum Taq using 0.2 uM each of primers envB5out 5'TAGAGCCCTGGAAGCATCCAGGAAG 3' and envB3out 5' TTGCTACTTGTGATTGCTCCATGT 3'. PCR conditions were 94 °C, 2min; (94 °C, 15s; 55 °C 30s; 68 °C, 4min) x 35; 68 °C, 15min. Second round PCR was performed using 1 ul of PCR 1 and 0.2 uM of primers of envB5in 5' TTAGGCATCTCCTATGGCAGGAAGAAG 3' and envB3in 5'– GTCTCGAGATACTGCTCCCACCC 3'. PCR conditions were the same as PCR-1 except for 45 cycles and an increased annealing temperature of 58 °C. PCR2 products were checked using 1% 96-well E-Gels (Invitrogen). Bands from PCRs with amplification efficiencies lower than 30% were subjected to library preparation using the Illumina Nextera DNA Sample Prepration Kit (Illumina) as described (*46*). Briefly, 10 ng of DNA per band were subjected to tagmentation, ligated to barcoded sequencing adapters using the Illumina Nextera Index Kit and then purified using AmPure Beads XP (Agencourt). 96 different purified samples were pooled into one library and then subjected to paired-end MiSeq Sequencing using the Illumina Miseq Nano 300 cycle kits at a concentration of 12 pM.

Bioinformatic processing of MiSeq env-sequences

Sequence adapters were removed using Cutadapt v1.8.3. Read assembly for each virus was performed in three steps. First, de novo assembly was performed using Spades v3.6.1 to yield long contig files. Contigs longer than 255bp were subsequently aligned to an HIV envelope reference sequence and a consensus sequence was generated using Geneious 8. Finally, reads were re-aligned to the consensus sequence to close gaps and a final consensus was generated. Sequences with double peaks (cutoff consensus identity for any residue <75%) were omitted from downstream analyses.

Analysis of viral evolution

Alignments of *env* nucleotide sequences were generated using ClustalW (version 2.11) (*47*) or via manual alignment using Geneious (version 8.1.6) sequence analysis software (*48*). Regions that could not be unambiguously aligned were removed for phylogenetic analysis and diversity calculations. Evolutionary model classes for maximum likelihood phylogenetic analyses were selected using jModelTest (*49*). Maximum likelihood phylogenetic trees were generated using PhyML (version 3) (*50*) with joint estimation of model parameter values and phylogenies. Within-patient trees were midpoint rooted and colored using the Rainbow Tree webtool through the Los Alamos National Laboratory (LANL) HIV Database (http://www.hiv.lanl.gov/content/sequence/RAINBOWTREE/rainbowtree.html). Pairwise genetic distances were calculated using PhyML (*50*) in the DIVEIN webtool (http://indra.mullins.microbiol.washington.edu/DIVEIN/diver.html) (*41*) using the Jukes-Cantor model of nucleotide substitution or the HIVw model of amino acid substitution (*51*). Pairwise genetic diversity was compared among samples using a two-sample U-statistic test (*39, 40*) in the DIVEIN webtool (*41*).

Highlighter plots were generated using the LANL nucleotide and amino acid Highlighter tools (http://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html). To generate Day 0 consensus sequences, each site was assigned the residue present in the highest frequency in Day 0 sequences. Gaps were treated as characters and sites without a single, most common residue were assigned ”X.”

Longitudinal modified weblogo plots of 3BNC117 contacts were generated using LASSIE (Longitudinal Antigenic Sequences and Sites from Intrahost Evolution) (*28*). Sites under significant selection (cutoff ≥80% new amino acid residue in particular position relative to Day 0 consensus) were determined and graphed using LASSIE.

Generation of CMV-env based pseudoviruses

CMV-env expression cassettes were generated according to an established protocol (*52*). Briefly, the CMV promoter was amplified from pcDNA 3.1D/V5-His-TOPO Expression vector using the primers CMVenv 5’AGTAATCAATTACGGGGTCATTAGTTCAT 3’and CMVenv1A 5’CATAGGAGATGCCTAAGCCGGTGGAGCTCTGCTTATATAGACCTC 3’. The PCR product was purified using the Macherey-Nagel PCR and Gel Purification Kit.

1 ul of first round PCR product was amplified using primers env1ATOPO 5’ CACC GGCTTAGGCATCTCCTATGGCAGGAAGAA 3’ and Rev19 5’ ACTTTTTGACCACTTGCCACCCAT 3’ in a 20 ul volume containing 1x High Fidelity Buffer, 2 mM MgSO4, 0.2 mM dNTPs, 0.5 units of High Fidelity Platinum and 0.2 uM of each primer. Cycling conditions were 94 °C, 2min; (94 °C, 15s; 55 °C 30s; 68 °C, 4min) x 35; 68 °C, 10min. The presence of env was validated by analysis on a 0.7% Agarose gel and the product was purified using the Macherey-Nagel Gel and PCR Purification Kit. 10 ng of envelope and 0.5 ng of CMV were then subjected to overlapping PCR with primers CMVenv and Rev19 in triplicates. Total reaction volume

was 50 ul containing 1x High Fidelity Buffer, 0.2 uM MgS04, 0.2 mM dNTPs, 1 U of Platinum Taq High Fidelity and 0.4 uM of each primer. PCR was carried out at 94 °C, 2min; (94 °C, 30s; 60 °C 30s; 68 °C, 4min) x 25; 68 °C, 10min. 500 ng of CMV-env were co-transfected with pSG3Δenv in 6-well plates into 293T cells and supernatant was harvested after 48h. Supernatants were subjected to neutralization testing by TZM.bl as described above.

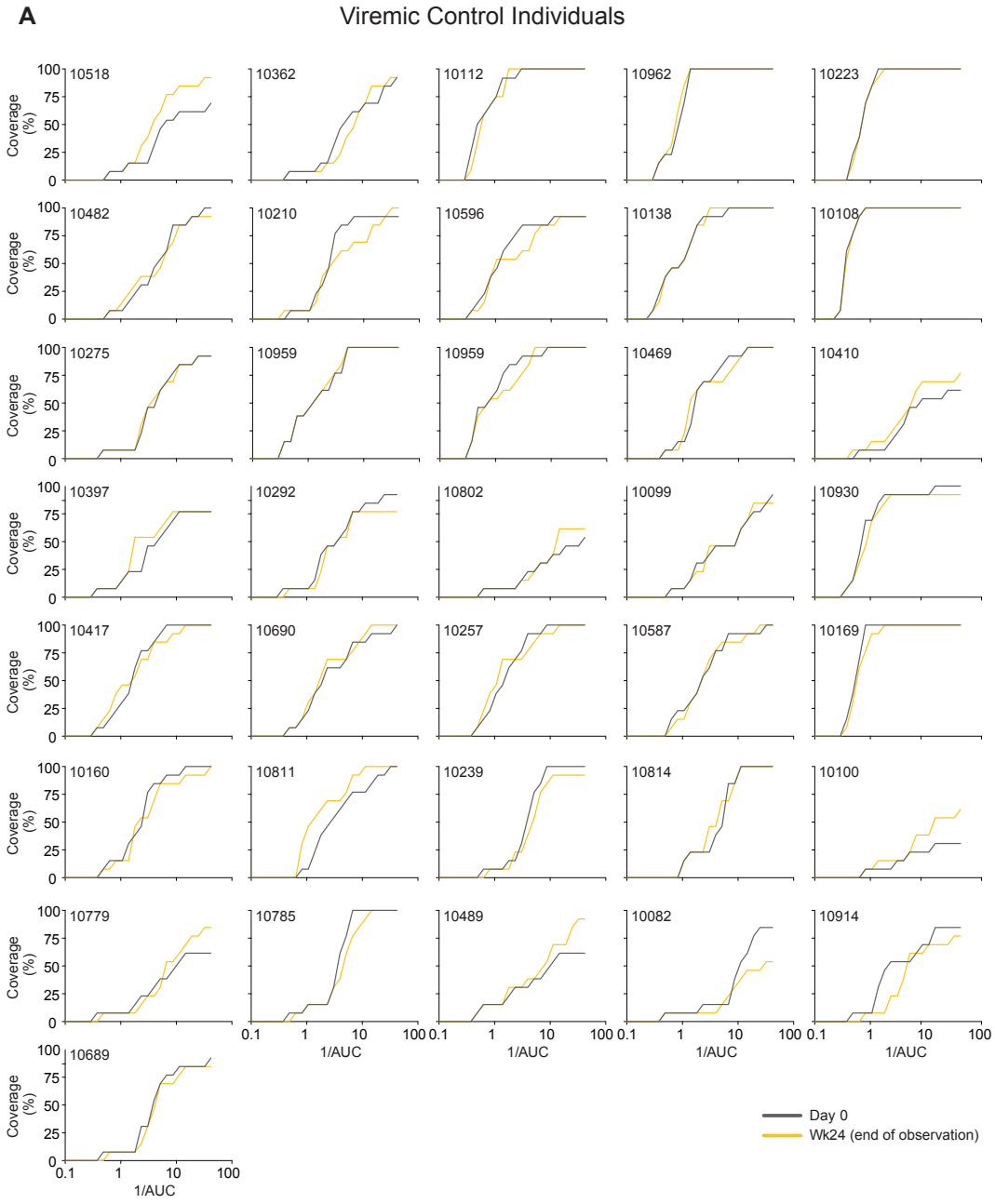


Fig. S1.

Dynamics of neutralization potency and breadth in each individual. Graphs show relationship between neutralization potency (1/AUC, x-axis) and breadth of neutralization (% of strains neutralized, y-axis) on Day 0 (grey) and at end of observation period (respective color for each group of study). **A**. untreated viremic control individuals (yellow). **B**. 3BNC117-treated individuals off-ART (dark blue) and 3BNC117-treated individuals on-ART (light blue)

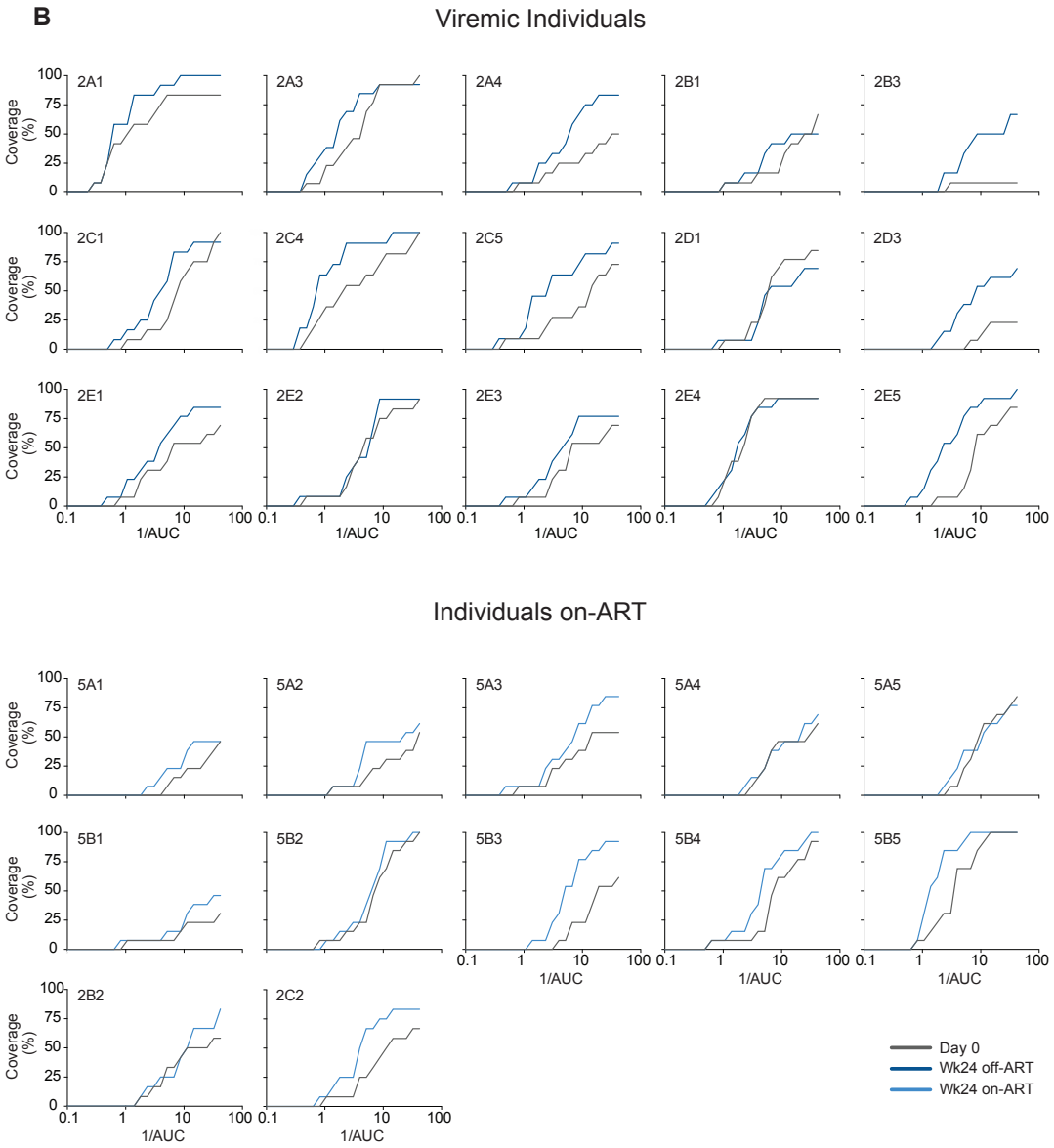


Fig. S1., continued

Dynamics of neutralization potency and breadth in each individual. Graphs show relationship between neutralization potency (1/AUC, x-axis) and breadth of neutralization (% of strains neutralized, y-axis) on Day 0 (grey) and at end of observation period (respective color for each group of study). **A**. untreated viremic control individuals (yellow). **B**. 3BNC117-treated individuals off-ART (dark blue) and 3BNC117-treated individuals on-ART (light blue)

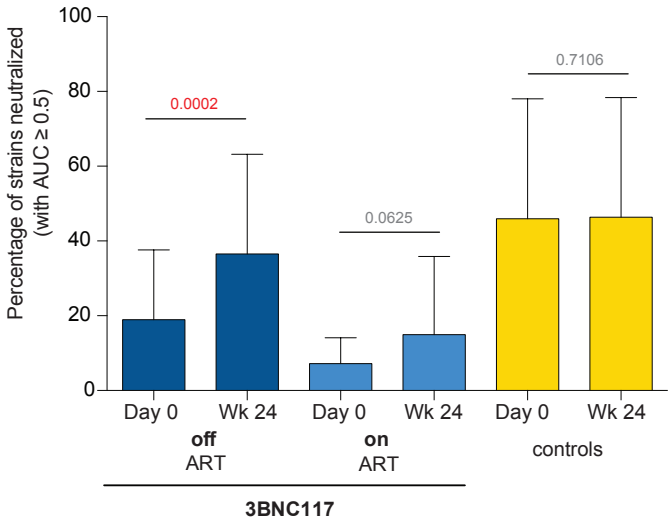


Fig. S2.

Dynamics of neutralization breadth by group of study. Bar plot illustrating the mean percentage of viruses neutralized (cutoff AUC ≥ 0.5) on Day0 and at the end of observation period by group of study. p-values determined using two-tailed Wilcoxon signed-rank test (red = significant, grey = n.s.). Raw values in Table S5.

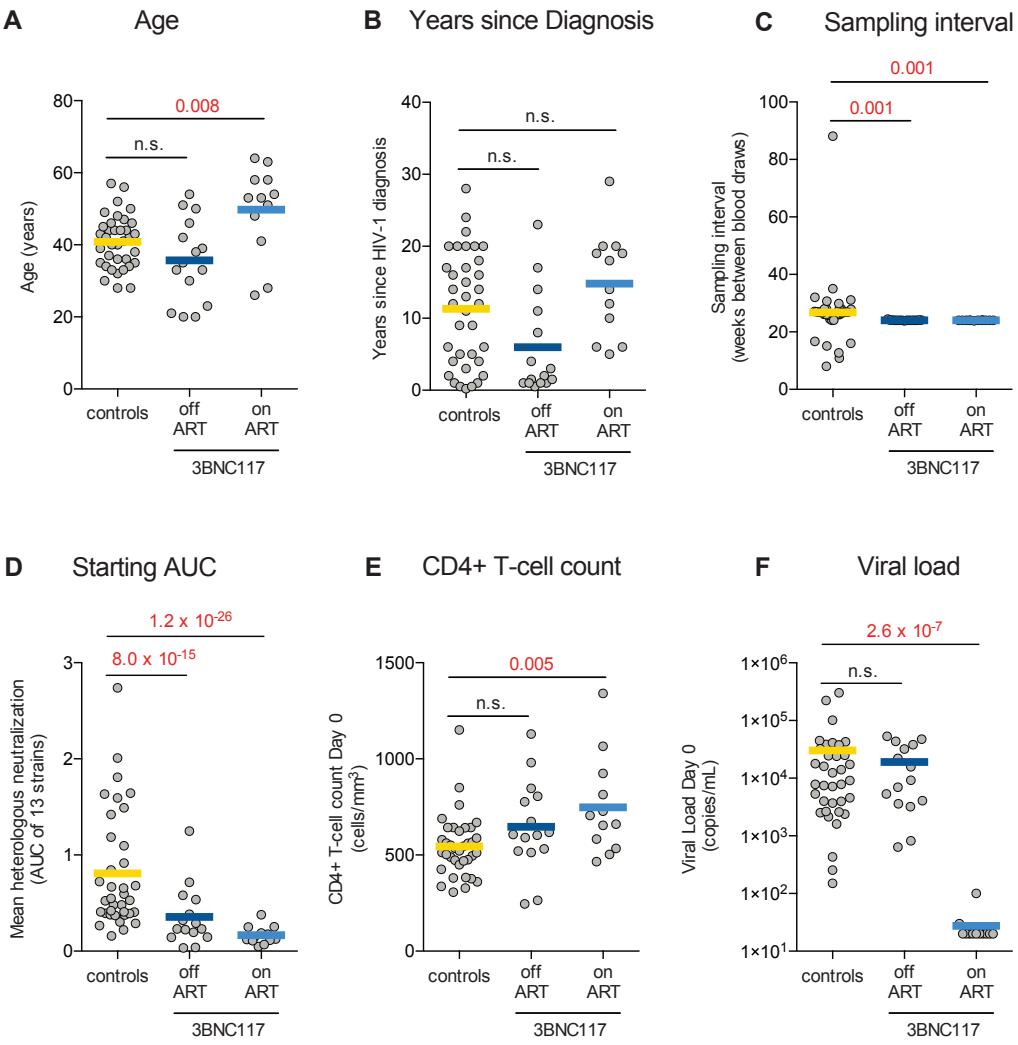


Fig. S3.

Comparison of clinical parameters between groups. Scatter dot plot of clinical parameters by patient group. Colored bars represent mean values of each group. p-values (indicated in red when significant) determined by unpaired Wilcoxon rank-sum test. Two hypotheses were tested for each parameter yielding a Bonferroni-corrected significance threshold of 0.025 with significance level $\alpha = 0.05$. n.s. = not significant. All parameters were tested in a confounder analysis and none of the parameter differences were found to be significant confounders (Table S7).

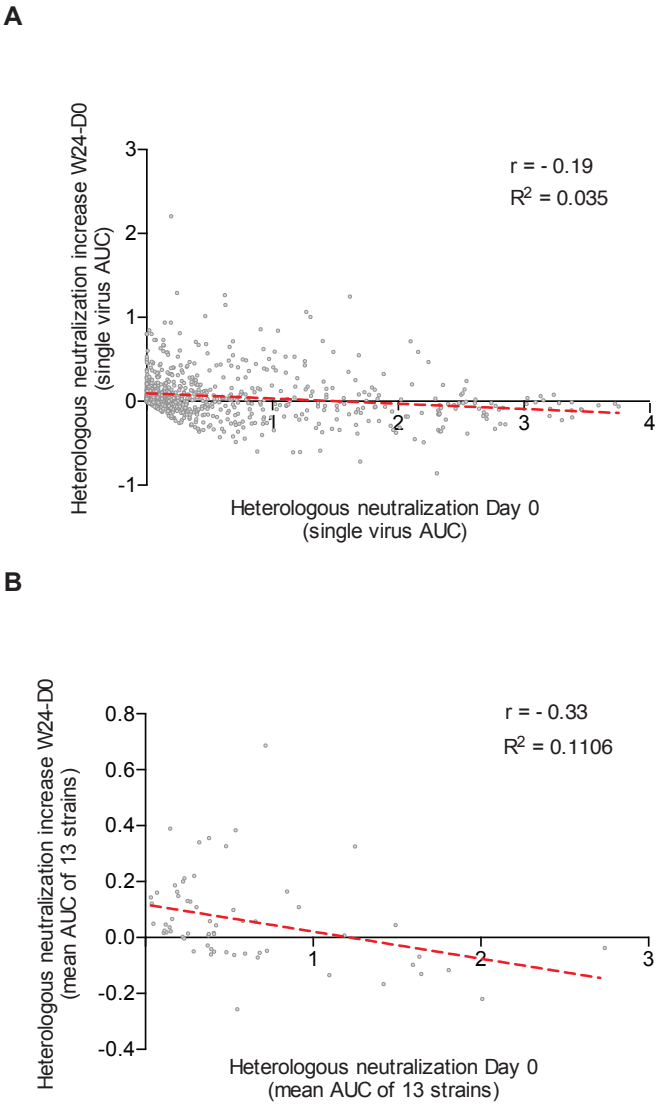


Fig. S4.

Correlation between starting AUC and AUC increase. **A.** Scatter plot of Day 0 AUC data (x-axis) and AUC difference W24-D0 (y-axis). Each dot shows the AUC data for a single virus from a single patient. **B.** Scatter plot of mean day 0 AUC data (x-axis) and mean AUC difference W24-D0 (y-axis). Each dot shows the mean data across all 13 viral strains for a particular patient. A linear fit model (dashed red line) is added with corresponding r and R2 in black at the top right.

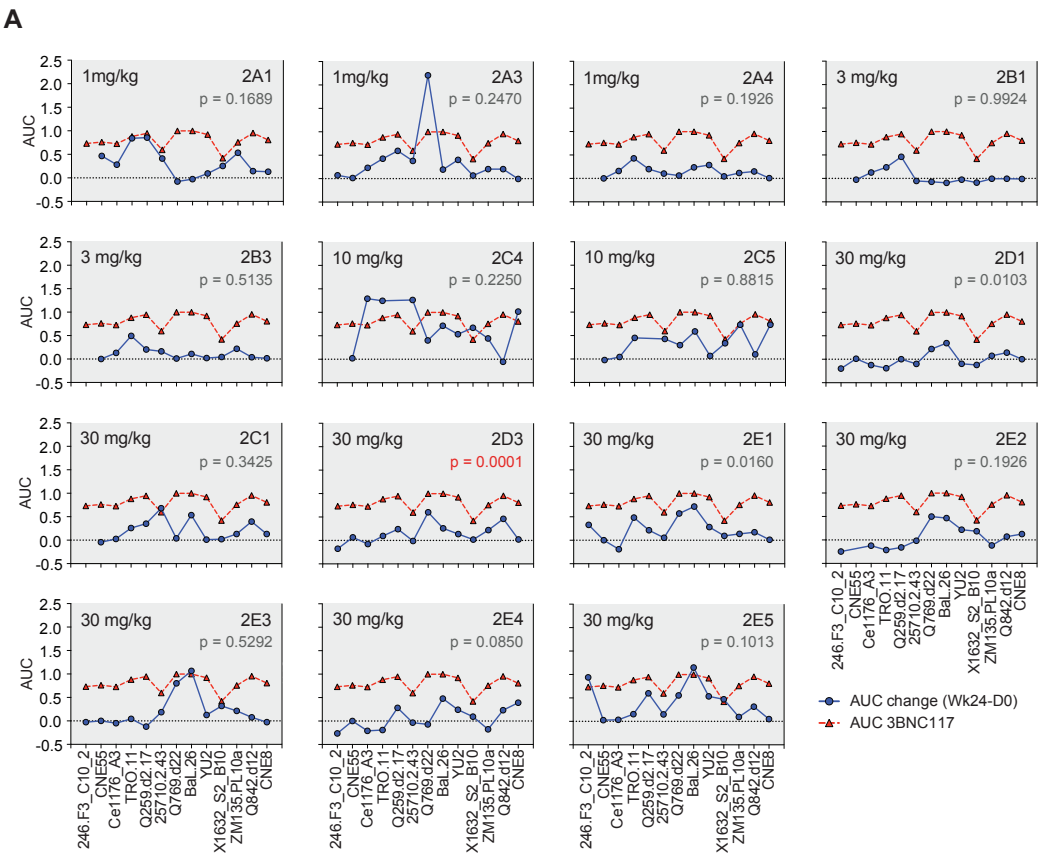


Fig. S5.

Neutralization profile comparison of AUC change in 3BNC117-treated individuals with neutralization profile of 3BNC117. For each individual the individual's AUC change (Wk24 - Day 0 AUC) against every strain is plotted in blue (**A.** viremic individuals, **B.** individuals on-ART). The AUC profile of 3BNC117 against those same strains is plotted in red as a comparator. The neutralization profiles were compared by assessing their Spearman rank correlation. p-values indicated in top right corner of graph (grey = n.s., red = significant). Bonferroni-corrected threshold: 0.0019. Full rank correlation and p-values in Table S8.

B

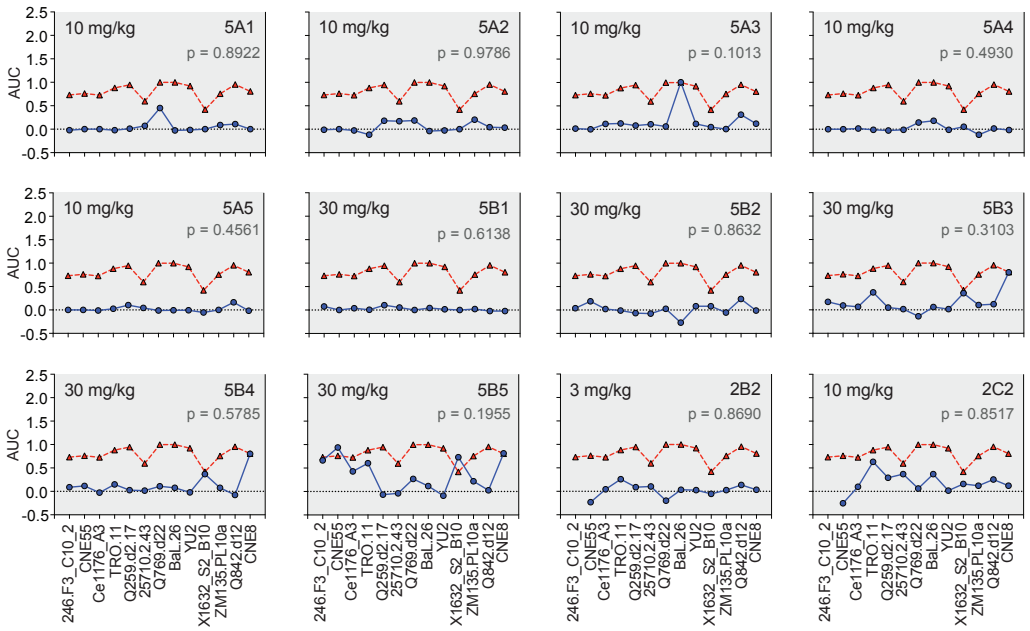


Fig. S5., continued

Neutralization profile comparison of AUC change in 3BNC117-treated individuals with neutralization profile of 3BNC117. For each individual the individual's AUC change (Wk24 - Day 0 AUC) against every strain is plotted in blue (**A.** viremic individuals, **B.** individuals on-ART). The AUC profile of 3BNC117 against those same strains is plotted in red as a comparator. The neutralization profiles were compared by assessing their Spearman rank correlation. p-values indicated in top right corner of graph (grey = n.s., red = significant). Bonferroni-corrected threshold: 0.0019. Full rank correlation and p-values in Table S8.

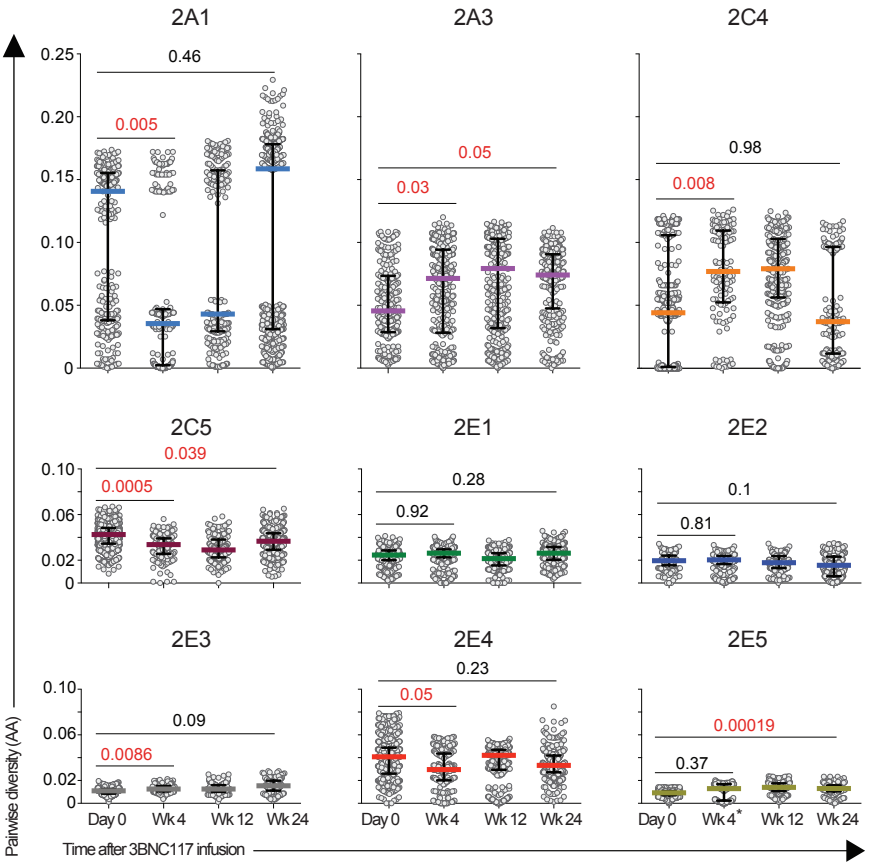


Fig. S6.

Pairwise diversity over time (amino acid level). Scatter plots depicting pairwise amino acid sequence diversity of plasma env sequences on d0, and wk 4 (*wk6 for 2E5), 12 and 24 after infusion. Each dot represents the pairwise genetic difference between two sequences at a given timepoint. Colored bars indicate median diversity, while black bars indicates the interquartile range. p-values were determined using a two-sample U-statistic based Z-test (*39-41*).

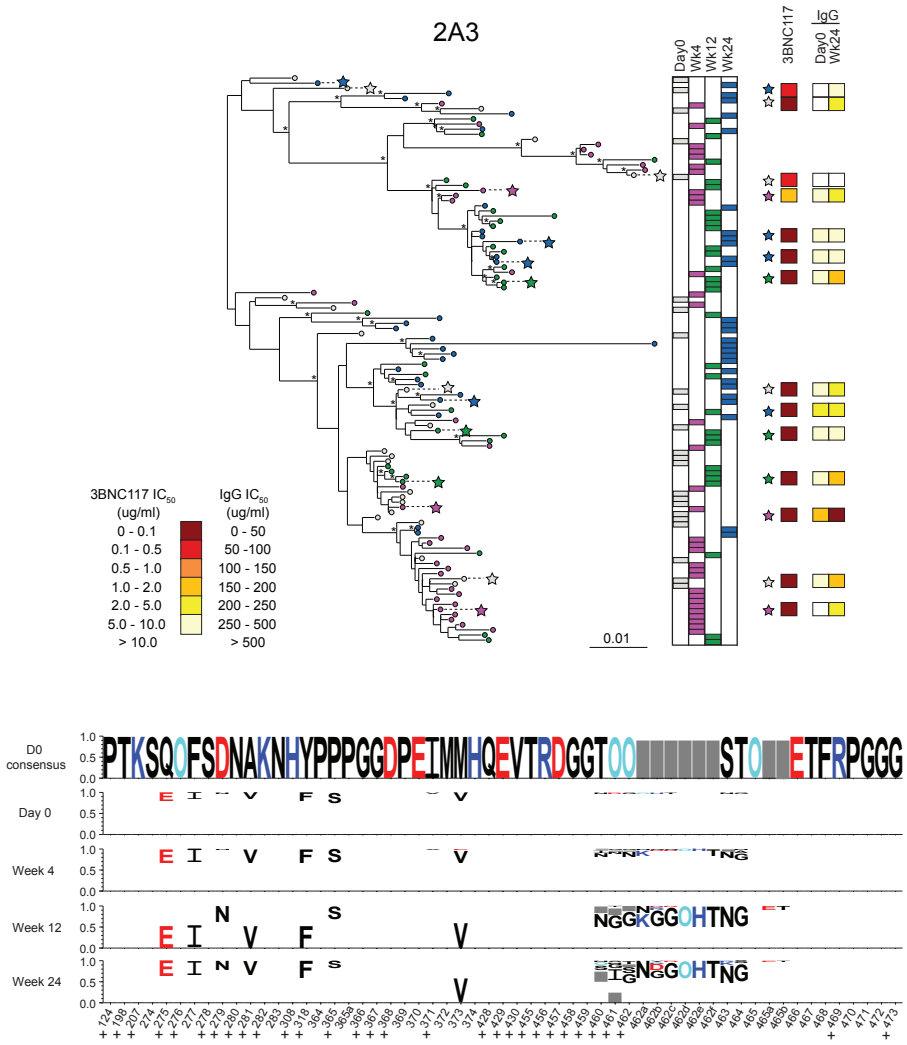


Fig. S7.

Viral evolution in subjects 2A3, 2C5, 2C4 and 2E2-2E5. **Top.** Maximum-likelihood phylogenetic trees of plasma-derived env sequences from respective subjects sampled on d0, wk4 (wk6 in case of 2E5), wk12 and wk24 after 3BNC117 infusion. Black asterisks indicate nodes with significant boostrap values (bootstrap support ≥ 70%). Bar graphs (middle) indicate the timepoints from which sequences were derived. Heat maps (right) show the 3BNC117 IC_{50}, d0 IgG IC_{50} and wk 24 IgG IC_{50} values against autologous envs as indicated by colored stars. **Bottom.** Weblogo plots that illustrate amino acid changes in and directly adjacent to 3BNC117 contact residues over time. White boxes indicate that sequence matches to the d0 consensus, grey boxes indicate gaps in alignment. Colors indicate basic (dark blue) and acidic (red) residues and a turquoise "O" is used instead of "N" to indicate a potential N-glycosylation site. Logo plots were generated using LASSIE (*28*). + symbols indicate contact residues confirmed by two crystal structures (*26, 27*).

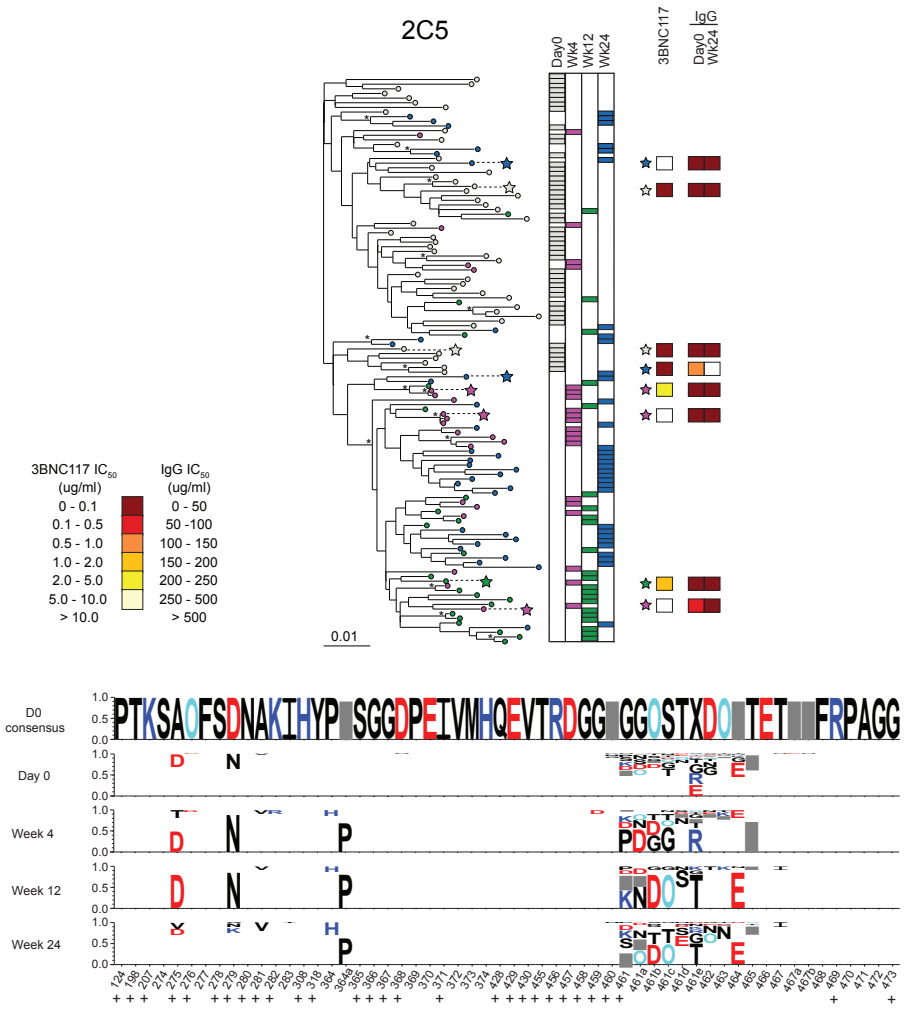


Fig. S7., continued

Viral evolution in subjects 2A3, 2C5, 2C4 and 2E2-2E5. **Top.** Maximum-likelihood phylogenetic trees of plasma-derived env sequences from respective subjects sampled on d0, wk4 (wk6 in case of 2E5), wk12 and wk24 after 3BNC117 infusion. Black asterisks indicate nodes with significant boostrap values (bootstrap support ≥ 70%). Bar graphs (middle) indicate the timepoints from which sequences were derived. Heat maps (right) show the 3BNC117 IC_{50}, d0 IgG IC_{50} and wk 24 IgG IC_{50} values against autologous envs as indicated by colored stars. **Bottom.** Weblogo plots that illustrate amino acid changes in and directly adjacent to 3BNC117 contact residues over time. White boxes indicate that sequence matches to the d0 consensus, grey boxes indicate gaps in alignment. Colors indicate basic (dark blue) and acidic (red) residues and a turquoise "O" is used instead of "N" to indicate a potential N-glycosylation site. Logo plots were generated using LASSIE (*28*). + symbols indicate contact residues confirmed by two crystal structures (*26, 27*).

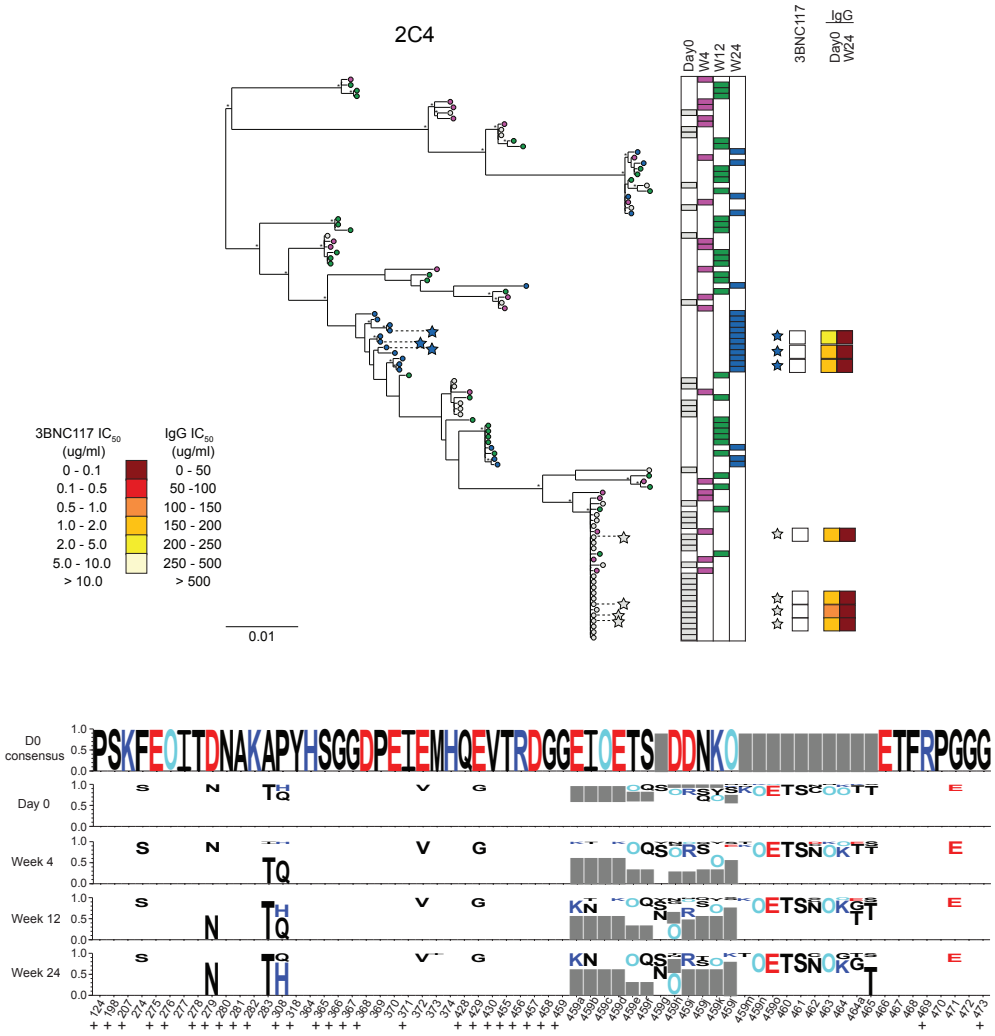


Fig. S7., continued

Viral evolution in subjects 2A3, 2C5, 2C4 and 2E2-2E5. **Top.** Maximum-likelihood phylogenetic trees of plasma-derived env sequences from respective subjects sampled on d0, wk4 (wk6 in case of 2E5), wk12 and wk24 after 3BNC117 infusion. Black asterisks indicate nodes with significant boostrap values (bootstrap support ≥ 70%). Bar graphs (middle) indicate the timepoints from which sequences were derived. Heat maps (right) show the 3BNC117 IC_{50}, d0 IgG IC_{50} and wk 24 IgG IC_{50} values against autologous envs as indicated by colored stars. **Bottom.** Weblogo plots that illustrate amino acid changes in and directly adjacent to 3BNC117 contact residues over time. White boxes indicate that sequence matches to the d0 consensus, grey boxes indicate gaps in alignment. Colors indicate basic (dark blue) and acidic (red) residues and a turquoise "O" is used instead of "N" to indicate a potential N-glycosylation site. Logo plots were generated using LASSIE (*28*). + symbols indicate contact residues confirmed by two crystal structures (*26, 27*).

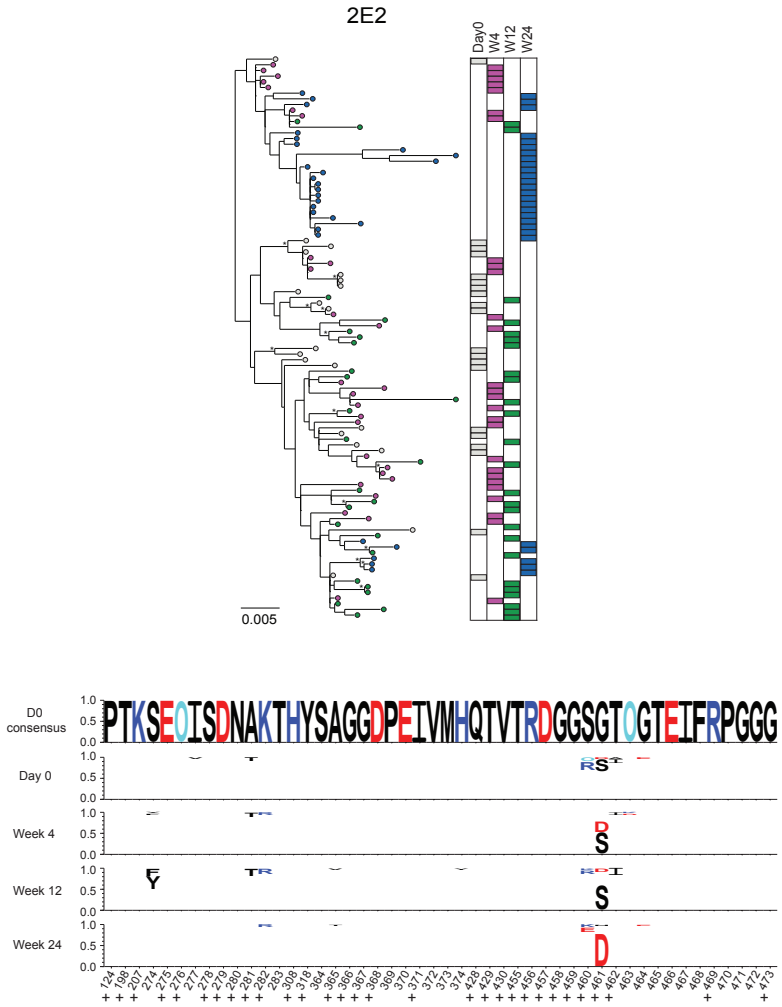


Fig. S7., continued

Viral evolution in subjects 2A3, 2C5, 2C4 and 2E2-2E5. **Top.** Maximum-likelihood phylogenetic trees of plasma-derived env sequences from respective subjects sampled on d0, wk4 (wk6 in case of 2E5), wk12 and wk24 after 3BNC117 infusion. Black asterisks indicate nodes with significant boostrap values (bootstrap support ≥ 70%). Bar graphs (middle) indicate the timepoints from which sequences were derived. **Bottom.** Weblogo plots that illustrate amino acid changes in and directly adjacent to 3BNC117 contact residues over time. White boxes indicate that sequence matches to the d0 consensus, grey boxes indicate gaps in alignment. Colors indicate basic (dark blue) and acidic (red) residues and a turquoise "O" is used instead of "N" to indicate a potential N-glycosylation site. Logo plots were generated using LASSIE (*28*). + symbols indicate contact residues confirmed by two crystal structures (*26, 27*).

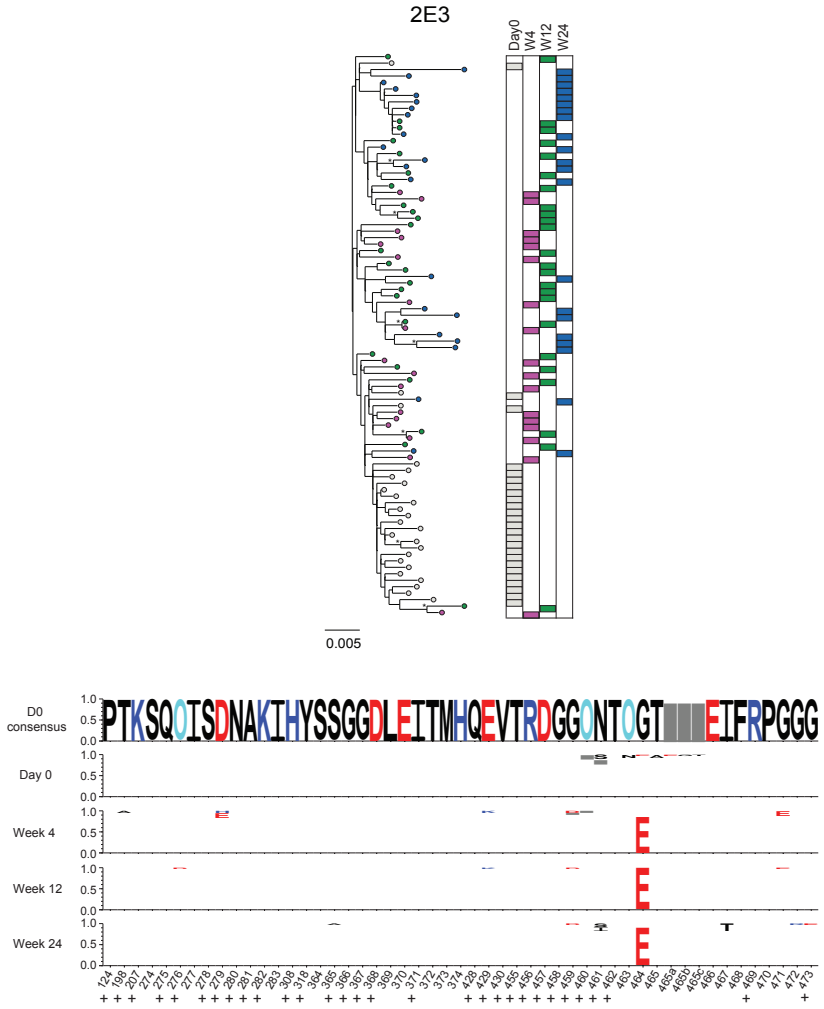


Fig. S7., continued

Viral evolution in subjects 2A3, 2C5, 2C4 and 2E2-2E5. **Top.** Maximum-likelihood phylogenetic trees of plasma-derived env sequences from respective subjects sampled on d0, wk4 (wk6 in case of 2E5), wk12 and wk24 after 3BNC117 infusion. Black asterisks indicate nodes with significant boostrap values (bootstrap support ≥ 70%). Bar graphs (middle) indicate the timepoints from which sequences were derived. **Bottom.** Weblogo plots that illustrate amino acid changes in and directly adjacent to 3BNC117 contact residues over time. White boxes indicate that sequence matches to the d0 consensus, grey boxes indicate gaps in alignment. Colors indicate basic (dark blue) and acidic (red) residues and a turquoise "O" is used instead of "N" to indicate a potential N-glycosylation site. Logo plots were generated using LASSIE (*28*). + symbols indicate contact residues confirmed by two crystal structures (*26, 27*).

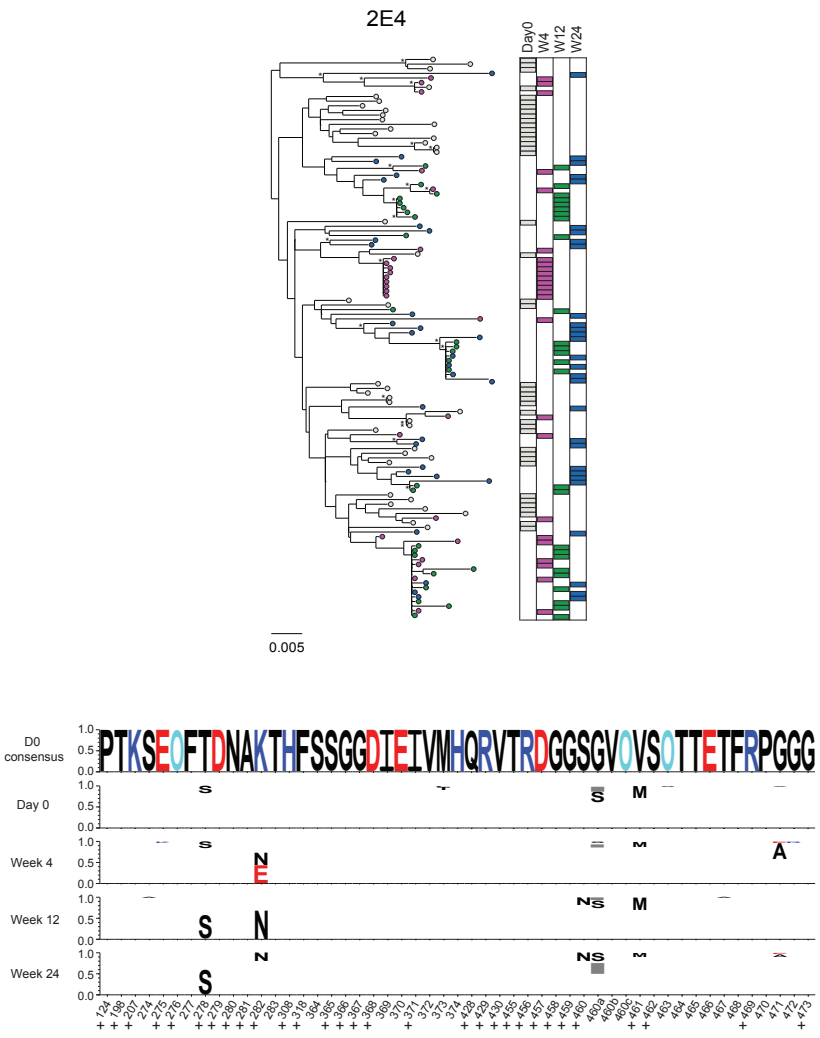


Fig. S7., continued

Viral evolution in subjects 2A3, 2C5, 2C4 and 2E2-2E5. **Top.** Maximum-likelihood phylogenetic trees of plasma-derived env sequences from respective subjects sampled on d0, wk4 (wk6 in case of 2E5), wk12 and wk24 after 3BNC117 infusion. Black asterisks indicate nodes with significant boostrap values (bootstrap support ≥ 70%). Bar graphs (middle) indicate the timepoints from which sequences were derived. **Bottom.** Weblogo plots that illustrate amino acid changes in and directly adjacent to 3BNC117 contact residues over time. White boxes indicate that sequence matches to the d0 consensus, grey boxes indicate gaps in alignment. Colors indicate basic (dark blue) and acidic (red) residues and a turquoise "O" is used instead of "N" to indicate a potential N-glycosylation site. Logo plots were generated using LASSIE (*28*). + symbols indicate contact residues confirmed by two crystal structures (*26, 27*).

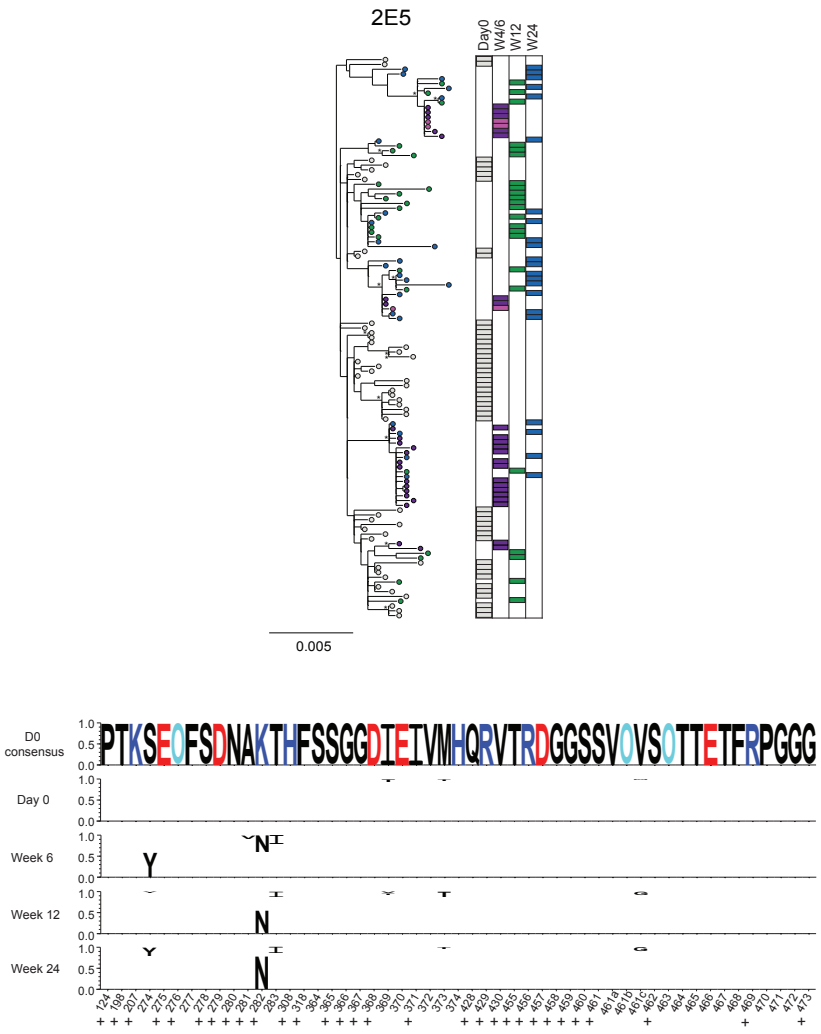


Fig. S7., continued

Viral evolution in subjects 2A3, 2C5, 2C4 and 2E2-2E5. **Top.** Maximum-likelihood phylogenetic trees of plasma-derived env sequences from respective subjects sampled on d0, wk4 (wk6 in case of 2E5), wk12 and wk24 after 3BNC117 infusion. Black asterisks indicate nodes with significant boostrap values (bootstrap support ≥ 70%). Bar graphs (middle) indicate the timepoints from which sequences were derived. For 2E5 pink indicates wk4, dark purple indicates wk6 sequences. **Bottom.** Weblogo plots that illustrate amino acid changes in and directly adjacent to 3BNC117 contact residues over time. White boxes indicate that sequence matches to the d0 consensus, grey boxes indicate gaps in alignment. Colors indicate basic (dark blue) and acidic (red) residues and a turquoise "O" is used instead of "N" to indicate a potential N-glycosylation site. Logo plots were generated using LASSIE (*28*). + symbols indicate contact residues confirmed by two crystal structures (*26, 27*).

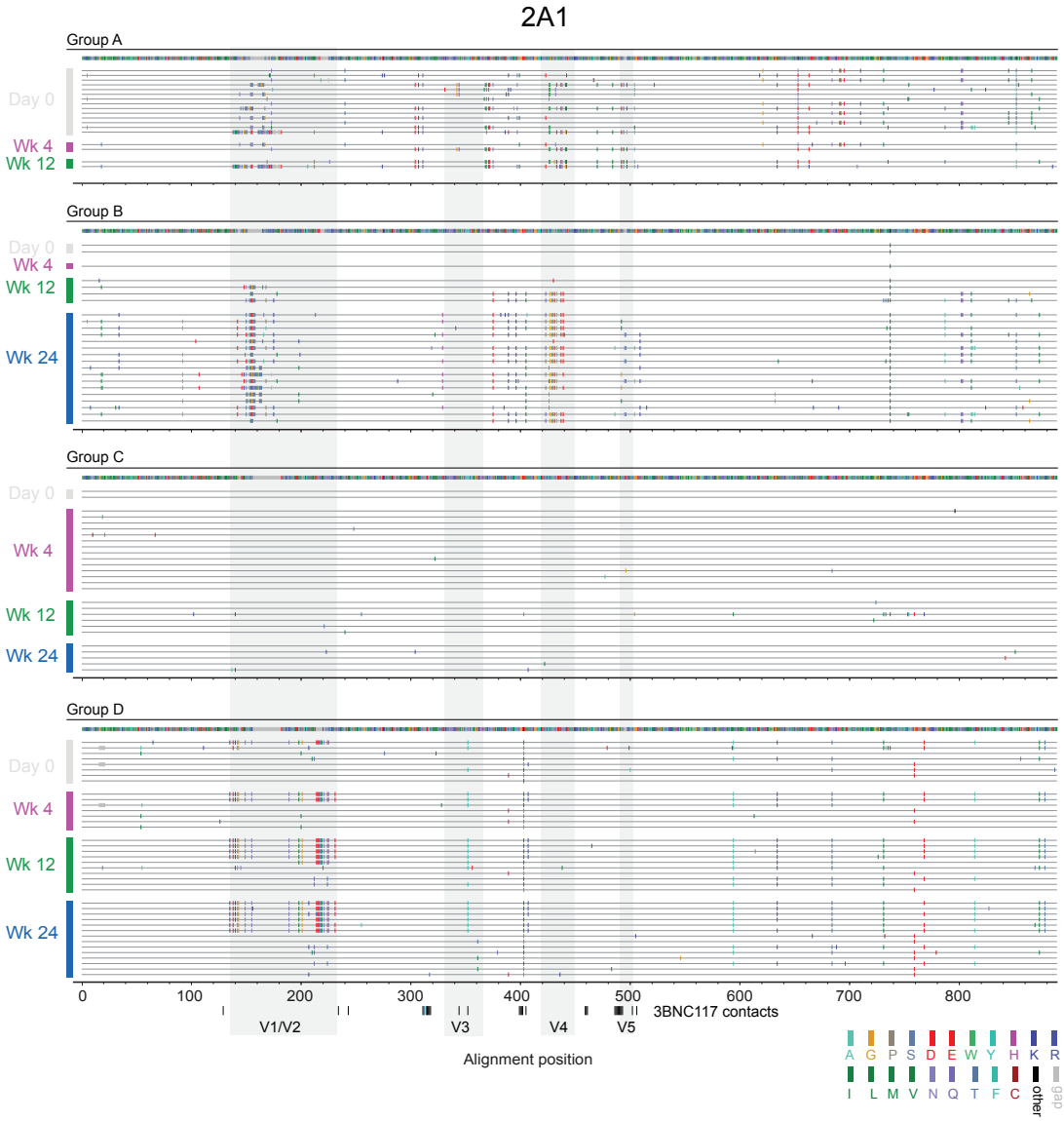


Fig. S8.

Viral sequence evolution in subject 2A1. Amino acid highlighter plots of 2A1 plasma env sequences separated by group (A-D, Fig. 4) illustrate that viral selection in 2A1 occurs mainly by 'clade shifting' with little sequence evolution within 'clades'. Similar patterns of shifting groups of viruses are seen in 2A3 and 2C4 where tree structure is even more complex making clear-cut group analyses challenging. Horizontal lines represent individual sequences and tic marks denote amino acid differences from the respective d0 consensus sequence (cutoff >50% identity) of each group. The amino acid color code key is at the bottom right. Bars to the left indicate the timepoint that each sequence belongs to and is ordered chronologically within each group from top to bottom. Grey-green boxes demarcate variable loops and black/light blue tic marks below the plot indicate 3BNC117 amino acid and glycan contacts, respectively, as determined by two crystal structures (*26, 27*).

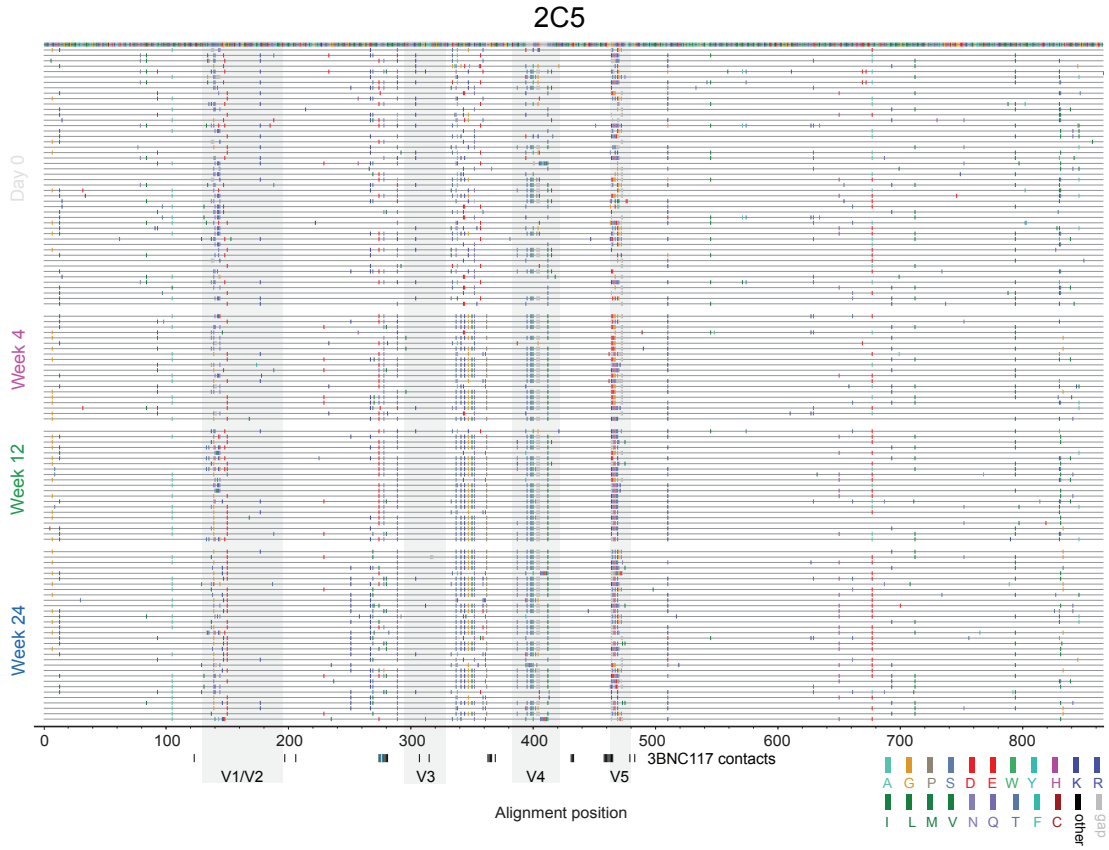


Fig. S9.

Viral sequence evolution in subjects 2C5 and 2E1-2E5. Amino acid highlighter plot for each patient (2C5, 2E1-2E5) depicts sequence changes over time relative to respective d0 consensus (cutoff >50% identity). Horizontal lines represent individual sequences and tic marks denote amino acid differences from consensus sequence. Amino acid color code key is at bottom right. Grey-green boxes demarcate variable loops and black/light blue tic marks underneath plots indicate 3BNC117 amino acid and glycan contacts, respectively, as determined by two crystal structures (*26, 27*). In contrast to subjects 2A1, 2A3 and 2C4, subjects 2C5 and 2E1-2E5 exhibit less complex phylogenies allowing for more stringent calling of a single day0 consensus sequence.

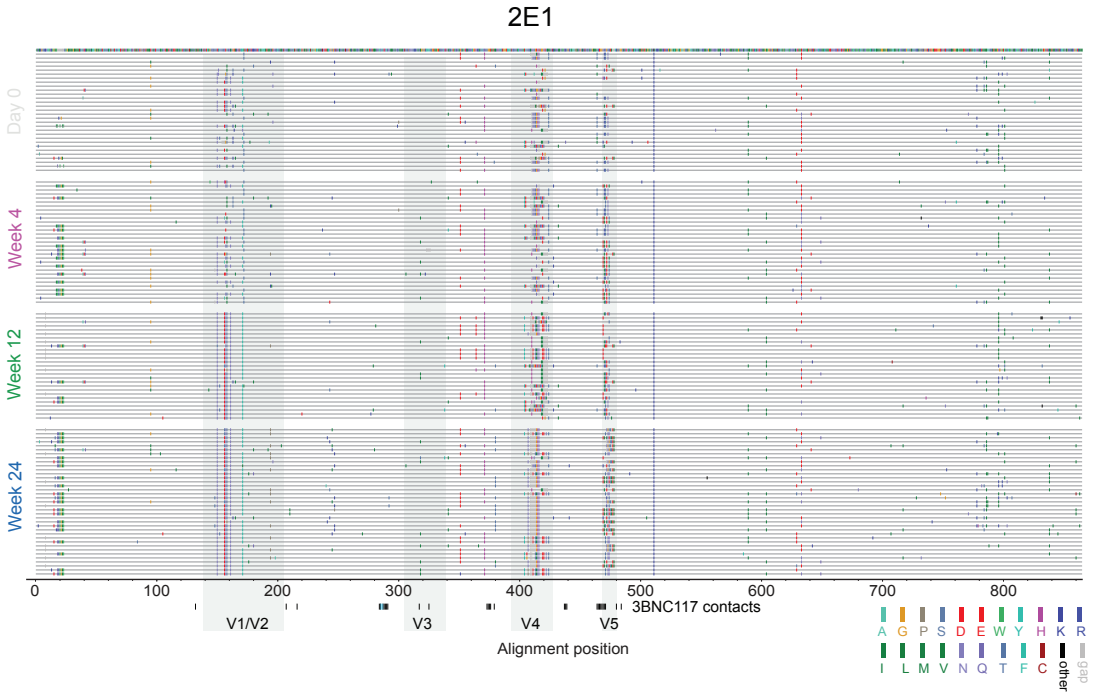


Fig. S9., continued

Viral sequence evolution in subjects 2C5 and 2E1-2E5. Amino acid highlighter plot for each patient (2C5, 2E1-2E5) depicts sequence changes over time relative to respective d0 consensus (cutoff >50% identity). Horizontal lines represent individual sequences and tic marks denote amino acid differences from consensus sequence. Amino acid color code key is at bottom right. Grey-green boxes demarcate variable loops and black/light blue tic marks underneath plots indicate 3BNC117 amino acid and glycan contacts, respectively, as determined by two crystal structures (*26, 27*). In contrast to subjects 2A1, 2A3 and 2C4, subjects 2C5 and 2E1-2E5 exhibit less complex phylogenies allowing for more stringent calling of a single day0 consensus sequence.

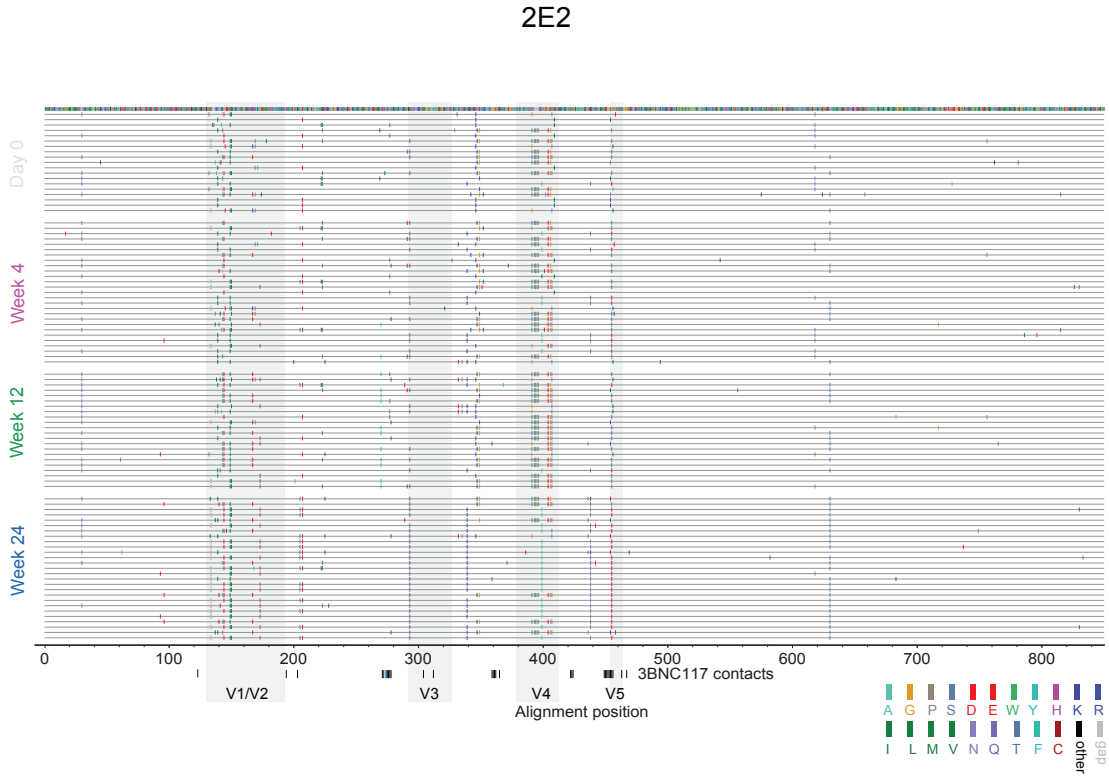


Fig. S9., continued

Viral sequence evolution in subjects 2C5 and 2E1-2E5. Amino acid highlighter plot for each patient (2C5, 2E1-2E5) depicts sequence changes over time relative to respective d0 consensus (cutoff >50% identity). Horizontal lines represent individual sequences and tic marks denote amino acid differences from consensus sequence. Amino acid color code key is at bottom right. Grey-green boxes demarcate variable loops and black/light blue tic marks underneath plots indicate 3BNC117 amino acid and glycan contacts, respectively, as determined by two crystal structures (*26, 27*). In contrast to subjects 2A1, 2A3 and 2C4, subjects 2C5 and 2E1-2E5 exhibit less complex phylogenies allowing for more stringent calling of a single day0 consensus sequence.

2E3

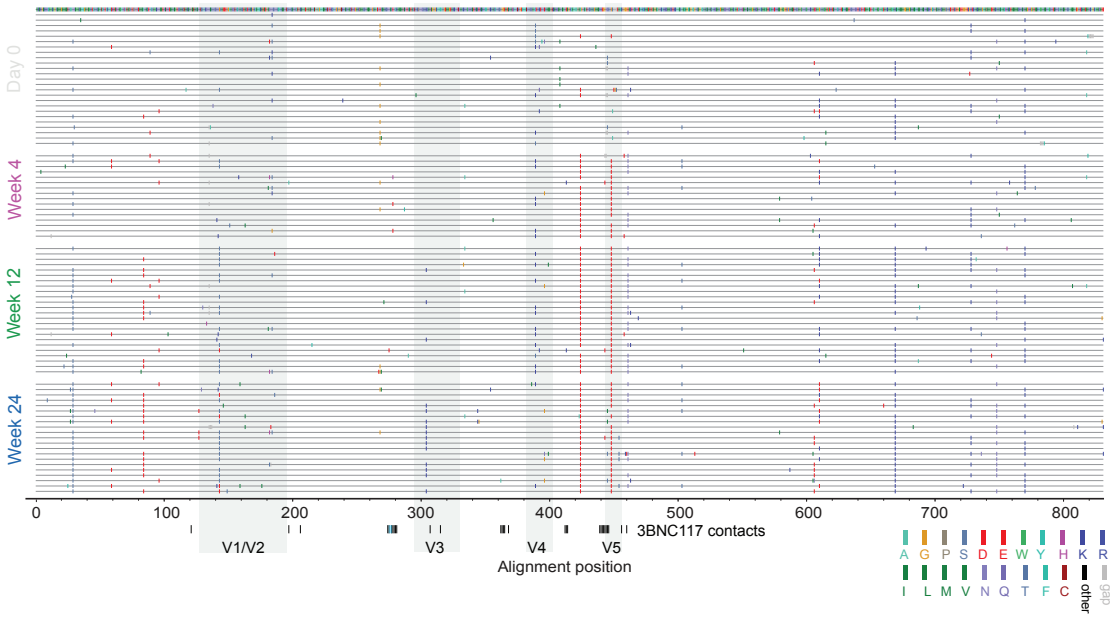


Fig. S9., continued

Viral sequence evolution in subjects 2C5 and 2E1-2E5. Amino acid highlighter plot for each patient (2C5, 2E1-2E5) depicts sequence changes over time relative to respective d0 consensus (cutoff >50% identity). Horizontal lines represent individual sequences and tic marks denote amino acid differences from consensus sequence. Amino acid color code key is at bottom right. Grey-green boxes demarcate variable loops and black/light blue tic marks underneath plots indicate 3BNC117 amino acid and glycan contacts, respectively, as determined by two crystal structures (*26, 27*). In contrast to subjects 2A1, 2A3 and 2C4, subjects 2C5 and 2E1-2E5 exhibit less complex phylogenies allowing for more stringent calling of a single day0 consensus sequence.

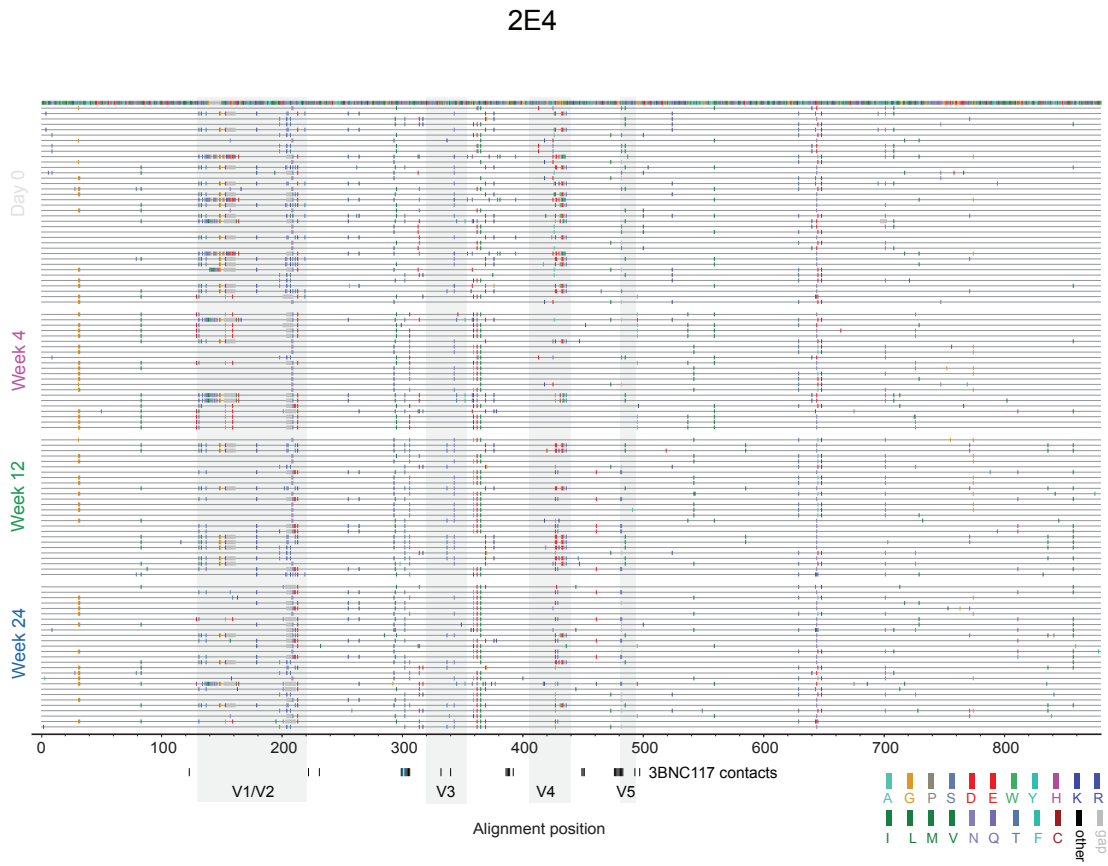


Fig. S9., continued

Viral sequence evolution in subjects 2C5 and 2E1-2E5. Amino acid highlighter plot for each patient (2C5, 2E1-2E5) depicts sequence changes over time relative to respective d0 consensus (cutoff >50% identity). Horizontal lines represent individual sequences and tic marks denote amino acid differences from consensus sequence. Amino acid color code key is at bottom right. Grey-green boxes demarcate variable loops and black/light blue tic marks underneath plots indicate 3BNC117 amino acid and glycan contacts, respectively, as determined by two crystal structures (*26, 27*). In contrast to subjects 2A1, 2A3 and 2C4, subjects 2C5 and 2E1-2E5 exhibit less complex phylogenies allowing for more stringent calling of a single day0 consensus sequence.

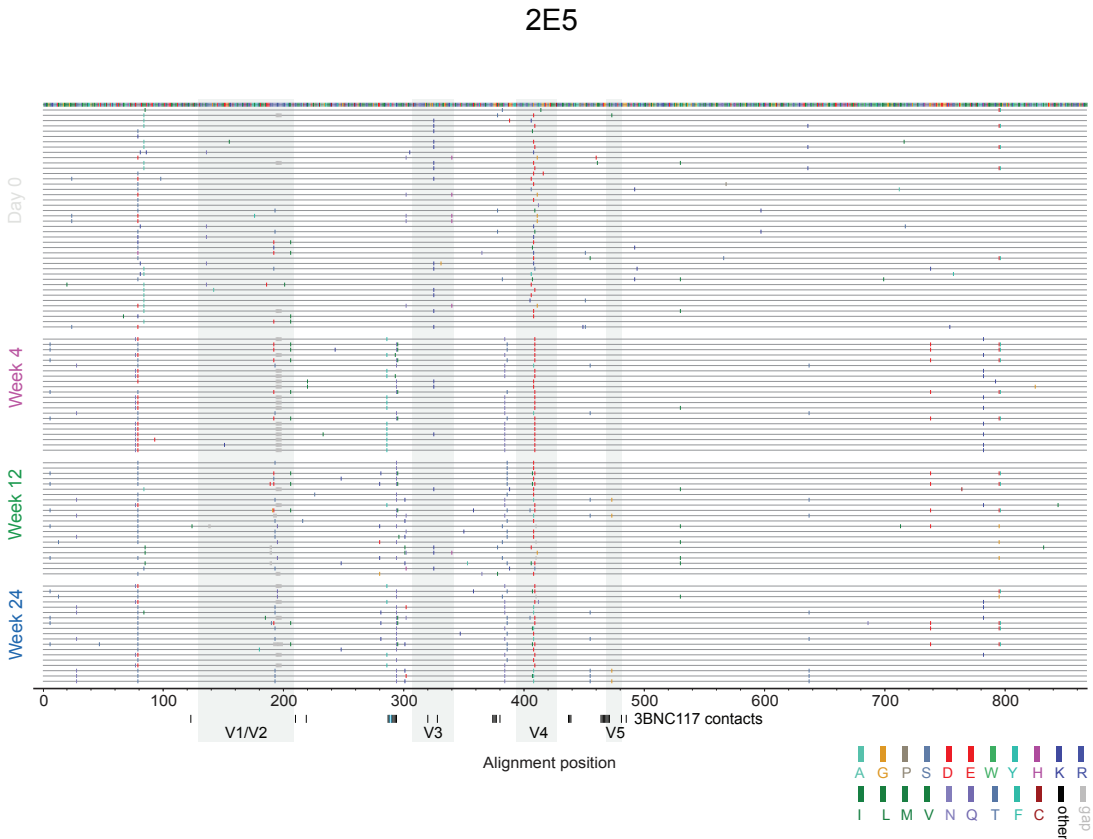


Fig. S9., continued

Viral sequence evolution in subjects 2C5 and 2E1-2E5. Amino acid highlighter plot for each patient (2C5, 2E1-2E5) depicts sequence changes over time relative to respective d0 consensus (cutoff >50% identity). Horizontal lines represent individual sequences and tic marks denote amino acid differences from consensus sequence. Amino acid color code key is at bottom right. Grey-green boxes demarcate variable loops and black/light blue tic marks underneath plots indicate 3BNC117 amino acid and glycan contacts, respectively, as determined by two crystal structures (*26, 27*). In contrast to subjects 2A1, 2A3 and 2C4, subjects 2C5 and 2E1-2E5 exhibit less complex phylogenies allowing for more stringent calling of a single day0 consensus sequence.

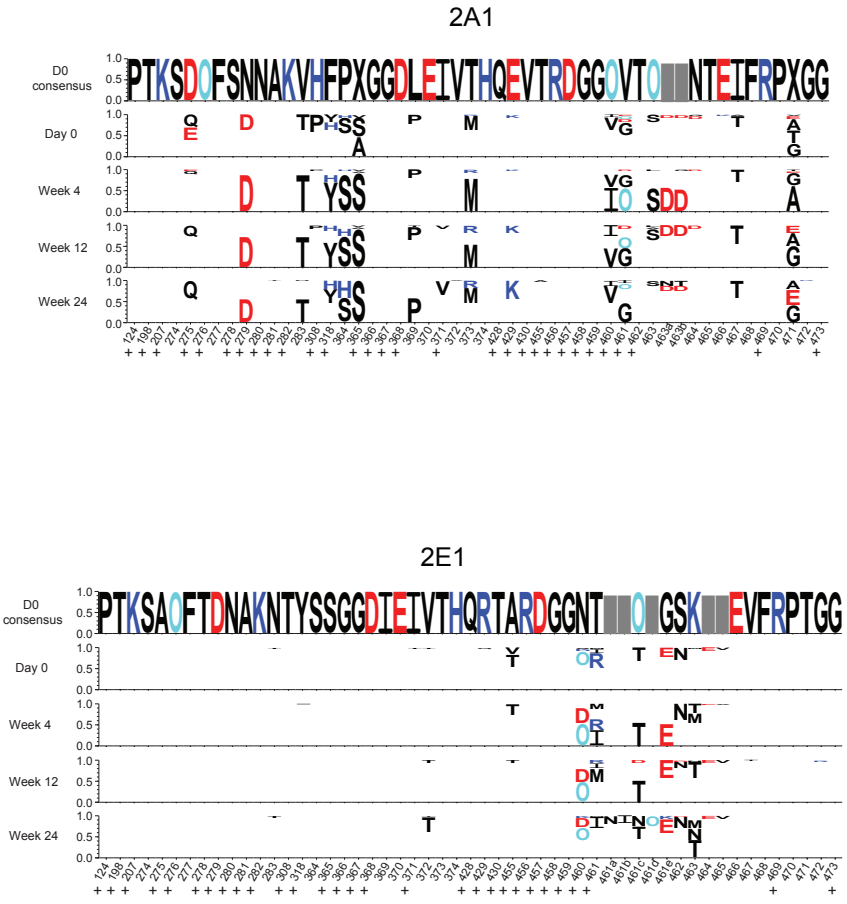


Fig. S10.

Full 3BNC117 contact site logo plots for subjects 2A1 and 2E1. Top. Weblogo plots that illustrate amino acid changes in and directly adjacent to 3BNC117 contact residues over time. White boxes indicate that sequence matches to the d0 consensus, grey boxes indicate gaps in alignment. Colors indicate basic (dark blue) and acidic (red) residues and a turquoise "O" is used instead of "N" to indicate a potential N-glycosylation site. Logo plots were generated using LASSIE (*28*). + symbols indicate 3BNC117 contact residues confirmed by two crystal structures (*26, 27*).

Table S1A.
Clinical characteristics of 3BNC117-treated subjects

Subject ID	3BNC117 dose	Age	Years since HIV Diagnosis	ART regimen pre-therapy	Clade	Sampling Interval (days)	HIV-RNA level (copies/ml)Day 0	HIV-RNA level (copies/ml) Week 24	abs. $CD4^+$ T cell count (day 0; cells/mm^3)	abs. $CD4^+$ T cell count (Week 24; cells/mm^3)
2A1	1 mg/kg	35	11	ART naïve	B	168	3210	2330	674	790
2A3	1 mg/kg	39	14	Off ART	B	171	43650	2870	520	800
2A4	1 mg/kg	42	8	ART naïve	B	167	5340	<20***	607	671
2B1	3 mg/kg	20	1	Off ART	ND	168	4090	<20***	264*	335
2B2	3 mg/kg	48	20	DRV/r/TDF/FTC	ND	167	100	<20	706	578
2B3	3 mg/kg	20	1	ART naïve	B	168	38190	<20***	777	627
2C2	10 mg/kg	51	12	ATV/r/3TC/ZDV	ND	169	30	140	728	804
2C4	10 mg/kg	54	23	Off ART	B	168	820	530	805	821
2C5	10 mg/kg	50	4	ART naïve	B	168	9260	24240	245*	203
2D1	30 mg/kg	33	3	ART naïve	B	168	53470	<20***	980	885
2C1	30 mg/kg	51	17	Off ART	B	168	47650	54950	1129	1026
2D3	30 mg/kg	33	0.5	ART naïve	B	168	640	<20***	618	482
2E1	30 mg/kg	21	2	ART naïve	B	169	15780	16500	847	660
2E2	30 mg/kg	46	1.5	ART naïve	B	169	6990	6590	513	480
2E3	30 mg/kg	23	1.5	ART naïve	BF	168	22030	35200	590	850
2E4	30 mg/kg	38	1	ART naïve	B	169	32220	69300	603	670
2E5	30 mg/kg	30	1	ART naïve	B	169	3610	6110	532	600
5A1	10 mg/kg	28	6	EFV/TDF/FTC	ND	168	<20	<20	814	661
5A2	10 mg/kg	58	19	EFV/TDF/FTC	ND	168	<20	<20	660	619
5A3	10 mg/kg	54	18	EVG/c/FTC/TDF	ND	168	<20	<20	1341	1203
5A4	10 mg/kg	53	5	EFV/TDF/FTC	ND	168	<20	<20	466	574
5A5	10 mg/kg	41	4	EFV/TDF/FTC	ND	168	<20	<20	924	624
5B1	30 mg/kg	58	28	EFV/TDF/FTC	ND	168	<20	<20	654	760
5B2	30 mg/kg	53	19	EFV/ZDV/3TC	ND	168	<20	<20	1065	1053
5B3	30 mg/kg	26	5	EFV/TDF/FTC	ND	168	<20	<20	583	603
5B4	30 mg/kg	64	9	ATV/r/TDF/FTC	ND	168	<20	<20	503	411
5B5	30 mg/kg	63	28	RAL/DRV/r/ETR	ND	168	<20	<20	534	821

* Absolute $CD4^+$ T cell count was 309 and 302 cells/mm^3 at screening. ND - Not Determined

*** Start of ART (Weeks post infusion): 2A3 (24), 2A4 (12), 2B1 (8), 2B3 (12), 2D1 (20), 2D3 (8)

ART-abbreviations: DRV darunavir, TDF tenofovir, FTC emtricitabine, r ritonavir, ATV atazanavir, 3TC lamivudine, ZDV zidovudine, EFV efavirenz, EVG elvitegravir, c cobicistat, RAL raltegravir, ETR etravirine

Table S1B.
Clinical characteristics of untreated viremic subjects

Subject ID	3BNC117 dose	Age	Years since HIV Diagnosis	ART regimen pre-therapy	Clade	Sampling Interval (days)	HIV-RNA level (copies/ml)Day 0	HIV-RNA level (copies/ml) Week 24	abs. $CD4^+$ T cell count (day 0; cells/mm^3)	abs. $CD4^+$ T cell count (Week 24; cells/mm^3)
10518	/	44	24	/	ND	182	38500	10000	643	536
10362	/	56	22	/	ND	190	2380	3670	643	818
10112	/	42	20	/	ND	171	7170	33500	513	551
10962	/	38	11	/	ND	168	7808	14112	489	468
10223	/	44	20	/	ND	196	4590	1300	473	444
10482	/	57	14	/	ND	183	3690	2014	642	640
10210	/	49	16	/	ND	189	3965	31495	546	412
10596	/	43	14	/	ND	188	1607	2269	503	488
10138	/	47	16	/	ND	196	17304	10149	536	532
10108	/	44	18	/	ND	196	7810	37800	850	501
10275	/	36	17	/	ND	181	24600	18220	643	648
10959	/	50	15	/	ND	189	303200	61000	760	621
10367	/	44	20	/	ND	195	432	1460	579	605
10469	/	35	13	/	ND	188	24000	573	1150	682
10410	/	46	4	/	ND	182	42920	62740	522	403
10397	/	39	20	/	ND	182	5327	13520	495	500
10292	/	42	28	/	ND	187	2152	3985	624	642
10802	/	40	2	/	ND	187	150	629	380	302
10099	/	35	12	/	ND	187	9095	12160	669	717
10930	/	52	17	/	ND	188	101500	263700*	562	448*
10417	/	36	5	/	ND	56	16070	41311	328	N/A
10690	/	30	2	/	ND	112	17820	97982	361	N/A
10257	/	32	6	/	ND	196	255	52091	449**	397
10587	/	28	1	/	ND	117	222465	772642	377	N/A
10169	/	34	12	/	ND	617	12423	34836	476	N/A
10160	/	34	0.5	/	ND	106	3984	5495	337	N/A
10811	/	43	20	/	ND	215	2520	5610	307	377
10239	/	46	3	/	ND	245	2660	8659	561	N/A
10814	/	45	5	/	ND	181	45000	33900	567	430
10100	/	37	0.2	/	ND	76	32360	13014	689	N/A
10779	/	33	4	/	ND	89	14029	25234	512	N/A
10785	/	33	1	/	ND	218	7438	148000	550	314
10489	/	41	6	/	ND	169	41100	1750	588	N/A
10082	/	40	9	/	ND	209	37600	104000	469	N/A
10914	/	48	9	/	ND	224	25000	21300	383	N/A
10689	/	28	0.5	/	ND	188	2600	137000	426***	336

* Absolute $CD4^+$ T cell count and HIV-RNA level were measured from a different sample from a timepoint 2 weeks later.

** Absolute $CD4^+$ T cell count was measured from a different sample from a timepoint 2 years earlier. ND - Not Determined. N/A - Not Available

*** Absolute $CD4^+$ T cell count was measured from a different sample from a timepoint 3 months earlier.

Table S2.
Autologous neutralization data of patient IgG against PBMC co-culture virus

		Virus Day 0		Virus Week 4*	
Subject ID	IgG tested	IC_{50} (µg/ml)	AUC	IC_{50} (µg/ml)	AUC
2A1	D0 IgG	>500	0.071	>500	0.11
	w24 IgG	>500	0.209	686.4	0.324
2A3	D0 IgG	>500	**0.571**	>500	0
	w24 IgG	**385.6**	**0.972**	>500	0.115
2C4	D0 IgG	>500	0.203	>500	0.28
	w24 IgG	>500	**0.505**	**286.5**	**0.659**
2C5	D0 IgG	>500	0.354	>500	0.143
	w24 IgG	**287.0**	**1.022**	**226.3**	**0.807**
2E1	D0 IgG	>500	0	>500	0
	w24 IgG	**417.8**	**0.857**	**311.7**	**0.593**
2E2	D0 IgG	>500	0.209	>500	0.319
	w24 IgG	**417.6**	0.401	**194.5**	**0.84**
2E3	D0 IgG	>500	**0.527**	>500	0.308
	w24 IgG	**96.7**	**1.335**	**269.3**	**0.615**
2E4	D0 IgG	>500	0.033	>500	0.005
	w24 IgG	>500	0.038	>500	0.055
2E5	D0 IgG	>500	0.181	>500	0.099
	w24 IgG	**462.6**	0.456	>500	0.253

IgG IC_{50} titers (µg/ml)

0 - 25 µg/ml
25 -100 µg/ml
100 - 400 µg/ml
400 - 500 µg/ml
>500 µg/ml

AUC

> 2
1-2
0.5 - 1
0 - 0.5

* for 2E5 week 6 virus was used instead of week 4

Table S3.
Panel of pseudoviruses used to determine neutralizing activity

HIV-1 Strains	Clade	Tier*
246-F3_C10_2	AC	Tier 2
25710-2.43	C	Tier 2
BaL.26	B	Tier 1B
Ce1176_A3	C	Tier 2
CNE55	AE	Tier 2
CNE8	AE	Tier 2
Q259.d2.17	A	Tier 2
Q769.d22	A	Tier 2
Q842.d12	A	Tier 2
TRO.11	B	Tier 2
X1632_S2_B10	G	Tier 2
YU2.DG	B	Tier 2
ZM135M.PL10a	C	Tier 2

*Tier classification for standard TZM-bl, based on (*21, 43*)

Table S4A.

Heterologous TZM.bl neutralization data of purified patient IgG (IC50 values, µg/ml)

Viremic control patients

Heterologous Strain	Clade	10518 Day 0	10518 Week 26 (182d)	10362 Day 0	10362 Week 27 (190d)	10112 Day 0	10112 Week 24 (171d)	10962 Day 0	10962 Week 24 (168d)	10223 Day 0	10223 Week 28 (196d)
YU2.DG	B	>500	>500	>500	>500	28.7	45.6	69.1	65.2	117.0	130.9
BaL.26	B	43.9	42.8	18.6	19.0	4.4	8.9	9.2	8.1	29.1	26.1
Q769.d22	A1	>500	>500	>500	>500	9.5	23.3	69.4	54.9	75.8	69.8
Q259.d2.17	A1	>500	>500	>500	>500	210.9	233.6	132.0	127.4	232.5	176.5
Q842.d12	A1	>500	>500	>500	>500	10.8	26.7	16.3	17.5	24.4	26.3
ZM135M.PL10a	C	>500	>500	>500	>500	183.1	94.6	233.4	124.8	192.1	490.7
25710-2.43	C	373.9	>500	>500	>500	22.8	42.0	98.7	92.3	79.5	95.6
CNE8	AE	>500	>500	>500	>500	115.4	233.3	163.9	136.4	92.4	84.2
TRO.11	B	>500	>500	>500	>500	19.9	30.0	28.9	27.9	41.1	57.4
X1632_S2_B10	G	>500	>500	>500	>500	>368	>500	258.4	394.5	44.3	70.8
Ce1176_A3	C	>500	>500	>500	>500	28.5	68.7	204.2	161.7	89.8	75.6
246-F3_C10_2	AC	>500	>500	>500	>500	77.4	109.7	278.6	189.5	120.3	123.2
CNE55	AE	>500	>500	>500	>500	*N/A*	196.1	93.1	90.7	34.7	57.1

Heterologous Strain	Clade	10482 Day 0	10482 Week 26 (183d)	10210 Day 0	10210 Week 27 (189d)	10596 Day 0	10596 Week 27 (188d)	10138 Day 0	10138 Week 28 (196d)	10108 Day 0	10108 Week 28 (196d)
YU2.DG	B	>500	>500	384.0	277.2	67.6	83.2	20.9	23.5	54.8	45.6
BaL.26	B	58.9	52.8	19.6	16.5	9.4	8.6	3.4	3.3	1.9	2.1
Q769.d22	A1	>500	>500	>500	>500	>500	>500	38.0	31.4	4.3	5.1
Q259.d2.17	A1	>500	>500	>500	>500	360.2	>500	291.0	371.5	7.2	7.7
Q842.d12	A1	>500	>500	348.5	393.9	89.4	234.6	7.7	11.2	5.2	6.7
ZM135M.PL10a	C	>500	>500	>500	>500	>500	>500	>500	>500	52.5	37.7
25710-2.43	C	>500	>500	480.4	472.1	32.7	61.4	248.2	214.1	14.3	15.1
CNE8	AE	>500	>500	>500	>500	>500	>500	327.4	384.8	27.1	29.9
TRO.11	B	276.6	>500	>500	>500	123.2	91.2	15.0	19.0	5.4	6.4
X1632_S2_B10	G	>500	>500	>500	>500	271.8	228.1	24.2	33.2	6.7	8.6
Ce1176_A3	C	>500	>500	>500	>500	60.3	72.7	>500	>500	94.9	94.1
246-F3_C10_2	AC	>500	>500	>500	>500	>500	>500	193.3	304.6	23.0	22.8
CNE55	AE	>500	>500	>500	>500	>500	>500	189.1	216.5	18.1	21.1

Heterologous Strain	Clade	10275 Day 0	10275 Week 26 (181d)	10959 Day 0	10959 Week 27 (189d)	10367 Day 0	10367 Week 28 (195d)	10469 Day 0	10469 Week 27 (188d)	10410 Day 0	10410 Week 26 (182d)
YU2.DG	B	>500	>500	41.8	41.6	24.8	22.0	439.4	431.5	>500	>500
BaL.26	B	27.1	34.2	10.7	11.8	5.0	6.0	26.4	35.6	56.0	32.0
Q769.d22	A1	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	>500	>500	166.5	222.0	89.9	139.0	>500	>500
Q842.d12	A1	>500	>500	16.3	17.8	170.9	>500	295.7	373.6	>500	>500
ZM135M.PL10a	C	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
25710-2.43	C	>500	>500	320.5	334.3	23.7	24.1	302.2	273.4	>500	276.4
CNE8	AE	>500	>500	493.4	428.4	>500	>500	>500	>500	>500	>500
TRO.11	B	>500	>500	51.1	45.2	11.3	10.5	>500	>500	>500	>500
X1632_S2_B10	G	>500	>500	136.9	105.6	102.5	90.4	>500	>500	>500	>500
Ce1176_A3	C	>500	>500	>500	>500	20.4	21.6	>500	>500	>500	>500
246-F3_C10_2	AC	>500	>500	227.2	209.5	36.8	46.0	>500	>500	>500	>500
CNE55	AE	>500	>500	51.0	60.9	>500	>500	>500	>500	>500	>500

Heterologous Strain	Clade	10397 Day 0	10397 Week 26 (182d)	10292 Day 0	10292 Week 27 (187d)	10802 Day 0	10802 Week 27 (187d)
YU2.DG	B	160.4	133.8	298.1	279.9	>500	>500
BaL.26	B	7.4	8.0	17.7	21.1	>500	>500
Q769.d22	A1	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	>500	>500	>500	>500
Q842.d12	A1	>500	438.3	313.0	273.4	>500	>500
ZM135M.PL10a	C	>500	>500	>500	>500	>500	>500
25710-2.43	C	182.4	320.5	429.6	445.7	>500	>500
CNE8	AE	>500	>500	378.7	485.3	>500	>500
TRO.11	B	422.6	422.4	482.1	>500	>500	>500
X1632_S2_B10	G	>500	>500	>500	>500	>500	>500
Ce1176_A3	C	>500	>500	>500	>500	56.4	61.9
246-F3_C10_2	AC	>500	>500	>500	>500	>500	>500
CNE55	AE	>500	>500	>500	>500	>500	>500

IgG IC_{50} titers (ug/ml)

0 - 25 µg/ml

25 -100 µg/ml

100 - 400 µg/ml

400 - 500 µg/ml

>500 µg/ml

Table S4A.
Heterologous TZM.bl neutralization data of purified patient IgG (IC50 values, µg/ml), continued

Heterologous Strain	Clade	10099 Day 0	10099 Week 27 (187d)	10930 Day 0	10930 Week 27 (188d)	10417 Day 0	10417 Week 8 (56d)	10690 Day 0	10690 Week 16 (112d)	10257 Day 0	10257 Week 28 (196d)
YU2.DG	B	>500	>500	102.2	122.5	405.2	>500	95.2	92.9	211.3	169.4
BaL.26	B	52.3	42.1	15.8	16.8	6.4	5.5	24.2	33.4	25.7	26.2
Q769.d22	A1	>500	>500	204.5	343.0	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	199.1	250.1	>500	>500	>500	>500	295.5	183.0
Q842.d12	A1	>500	>500	61.8	77.1	333.4	301.8	455.4	366.8	136.0	70.2
ZM135M.PL10a	C	>500	>500	>500	>500	219.6	133.7	>500	>500	254.0	155.6
25710-2.43	C	432.6	244.7	235.7	326.7	161.8	90.7	226.9	202.7	169.4	94.1
CNE8	AE	443.1	>500	29.6	39.2	48.5	32.9	365.6	295.1	150.2	117.5
TRO.11	B	187.5	198.6	81.9	122.4	81.7	68.8	174.4	166.6	67.7	62.2
X1632_S2_B10	G	>500	>500	54.4	109.2	>500	>500	228.0	122.1	369.1	175.5
Ce1176_A3	C	>500	>500	340.1	405.1	367.7	445.0	363.8	347.3	>500	>500
246-F3_C10_2	AC	>500	>500	56.4	61.9	>500	>500	>500	>500	>500	>500
CNE55	AE	>500	>500	87.7	83.6	476.1	453.9	>500	>500	>500	>500

Heterologous Strain	Clade	10587 Day 0	10587 Week 17 (117d)	10169 Day 0	10169 Week 88 (617d)	10160 Day 0	10160 Week 15 (106d)	10811 Day 0	10811 Week 31 (215d)	10239 Day 0	10239 Week 35 (245d)
YU2.DG	B	>500	>500	29.2	49.5	267.5	396.7	325.0	389.9	>500	>500
BaL.26	B	44.4	50.8	4.7	6.4	28.4	28.2	80.0	71.8	55.2	75.7
Q769.d22	A1	140.5	276.7	25.2	39.9	>500	>500	>500	498.6	>500	>500
Q259.d2.17	A1	>500	>500	72.2	102.9	>500	>500	>500	>500	>500	>500
Q842.d12	A1	252.5	194.6	21.2	33.5	>500	>500	476.7	146.2	>500	>500
ZM135M.PL10a	C	>500	>500	67.4	163.2	>500	>500	>500	>500	>500	>500
25710-2.43	C	377.1	416.7	74.2	37.1	199.7	300.0	254.5	84.6	245.7	326.5
CNE8	AE	>500	>500	45.7	67.3	>500	>500	>500	>500	>500	>500
TRO.11	B	66.0	102.2	13.9	16.1	52.4	72.5	294.6	71.8	>500	>500
X1632_S2_B10	G	>500	>500	68.2	41.3	>500	>500	>500	>500	>500	>500
Ce1176_A3	C	>500	448.5	96.3	150.0	301.2	438.5	426.7	117.2	>500	>500
246-F3_C10_2	AC	>500	>500	122.8	245.4	>500	>500	>500	278.4	>500	>500
CNE55	AE	>500	311.7	75.2	100.2	>500	>500	472.9	255.5	>500	>500

Heterologous Strain	Clade	10814 Day 0	10814 Week 26 (181d)	10100 Day 0	10100 Week 11 (76d)	10779 Day 0	10779 Week 13 (89d)	10785 Day 0	10785 Week 31 (218d)	10489 Day 0	10489 Week 24 (169d)
YU2.DG	B	>500	>500	>500	>500	>500	>500	>500	>500	221.2	239.7
BaL.26	B	126.4	180.7	77.5	96.4	35.0	26.8	23.4	24.6	13.8	20.9
Q769.d22	A1	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Q842.d12	A1	>500	>500	>500	>500	>500	>500	473.0	>500	>500	>500
ZM135M.PL10a	C	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
25710-2.43	C	>500	>500	>500	>500	>500	>500	149.9	151.3	399.0	360.7
CNE8	AE	228.4	233.6	>500	>500	205.3	310.0	>500	>500	>500	>500
TRO.11	B	134.6	175.1	>500	>500	>500	>500	>500	>500	59.2	73.2
X1632_S2_B10	G	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Ce1176_A3	C	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
246-F3_C10_2	AC	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
CNE55	AE	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500

Heterologous Strain	Clade	10082 Day 0	10082 Week 30 (209d)	10914 Day 0	10914 Week 32 (224d)	10689 Day 0	10689 Week 27 (188d)
YU2.DG	B	>500	>500	170.5	>500	>500	>500
BaL.26	B	34.2	38.8	29.7	77.8	29.0	43.2
Q769.d22	A1	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	>500	>500	>500	>500
Q842.d12	A1	>500	>500	>500	>500	>500	>500
ZM135M.PL10a	C	>500	>500	209.9	419.4	>500	>500
25710-2.43	C	>500	>500	213.9	456.7	488.1	419.1
CNE8	AE	>500	>500	276.4	>500	>500	>500
TRO.11	B	487.1	>500	307.7	>500	>500	>500
X1632_S2_B10	G	>500	>500	>500	>500	>500	>500
Ce1176_A3	C	>500	>500	329.3	>500	>500	>500
246-F3_C10_2	AC	>500	>500	>500	>500	>500	>500
CNE55	AE	>500	>500	>500	>500	>500	>500

IgG IC_{50} titers (ug/ml)
0 - 25 µg/ml
25 -100 µg/ml
100 - 400 µg/ml
400 - 500 µg/ml
>500 µg/ml

Table S4A.
Heterologous TZM.bl neutralization data of purified patient IgG (IC50 values, µg/ml), continued
3BNC117-treated viremic patients

Heterologous Strain	Clade	2A1		2A3		2A4		2B1		2B3	
		Day -7	Week 24	Day 0	Week 24	Day -7	Week 24	Day -7	Week 24	Day -7	Week 24
YU2.DG	B	**24.9**	**22.2**	>500	**363.4**	>500	>500	>500	>500	>500	>500
BaL.26	B	**3.3**	**3.2**	**28.4**	**18.9**	**89.5**	**71.3**	**135.0**	**164.8**	**365.9**	**310.8**
Q769.d22	A1	**102.6**	**97.0**	>500	**19.4**	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	**170.9**	**70.4**	**92.5**	**84.7**	>500	>500	>500	>500	>500	>500
Q842.d12	A1	**22.2**	**18.3**	**475.0**	**417.8**	>500	>500	>500	>500	>500	>500
ZM135M.PL10a	C	>500	**234.5**	>500	>500	>500	>500	>500	>500	>500	>500
25710-2.43	C	>500	**324.1**	**110.8**	**136.0**	>500	>500	>500	>500	>500	>500
CNE8	AE	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
TRO.11	B	**137.7**	**52.9**	**218.6**	**203.0**	>500	>500	>500	>500	>500	**495.5**
X1632_S2_B10	G	**65.1**	**48.3**	>500	>500	>500	>500	>500	>500	>500	>500
Ce1176_A3	C	>500	>500	**482.2**	**446.7**	>500	>500	>500	>500	>500	>500
246-F3_C10_2	AC	*N/A*	**297.2**	**445.4**	**283.7**	*N/A*	>500	*N/A*	>500	*N/A*	>500
CNE55	AE	>500	**204.8**	>500	>500	>500	>500	>500	>500	>500	>500

Heterologous Strain	Clade	2C4		2C5		2D1		2C1		2D3	
		Day -7	Week 24	Day -7	Week 24	Day 0	Week 24	Day -7	Week 24	Day -7	Week24
YU2.DG	B	**120.6**	**69.8**	>500	>500	>500	>500	>500	>500	>500	>500
BaL.26	B	**43.0**	**15.7**	**40.2**	**15.1**	**114.6**	**69.4**	**111.2**	**51.1**	>500	>500
Q769.d22	A1	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Q842.d12	A1	**82.0**	**86.7**	>500	>500	>500	>500	>500	**423.1**	>500	**369.1**
ZM135M.PL10a	C	>500	**419.5**	>500	**225.8**	>500	>500	>500	>500	>500	>500
25710-2.43	C	**241.8**	**49.6**	**305.9**	**159.6**	>500	>500	**426.3**	**125.4**	>500	>500
CNE8	AE	**325.3**	**83.3**	>500	**240.4**	>500	>500	>500	>500	>500	>500
TRO.11	B	**69.8**	**13.3**	>500	**316.7**	>500	>500	>500	**486.2**	>500	>500
X1632_S2_B10	G	>500	**378.2**	>500	>500	>500	>500	>500	>500	>500	>500
Ce1176_A3	C	>500	**79.0**	>500	>500	>500	>500	>500	>500	>500	>500
246-F3_C10_2	AC	*N/A*	**190.1**	*N/A*	>500	>500	>500	*N/A*	>500	>500	>500
CNE55	AE	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500

Heterologous Strain	Clade	2E1		2E2		2E3		2E4		2E5	
		Day 0	Week 24	Day 0	Week 24	Day 0	Week 24	Day 0	Week 24	Day 0	Week 24
YU2.DG	B	>500	>500	>500	>500	>500	>500	>500	**495.7**	>500	>500
BaL.26	B	**87.4**	**41.1**	**28.9**	**16.8**	**85.2**	**24.2**	**103.0**	**87.7**	**393.8**	**66.1**
Q769.d22	A1	**407.3**	**181.7**	>500	>500	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	>500	>500	>500	>500	>500	**366.6**	>500	**480.4**
Q842.d12	A1	>500	>500	>500	>500	>500	>500	**467.2**	**333.4**	>500	**463.1**
ZM135M.PL10a	C	>500	>500	>500	>500	>500	>500	**366.3**	**495.1**	>500	>500
25710-2.43	C	**439.1**	**397.9**	>500	>500	>500	**471.6**	**148.4**	**158.7**	>500	>500
CNE8	AE	>500	>500	>500	>500	>500	>500	**177.8**	**98.5**	>500	>500
TRO.11	B	**483.4**	**135.8**	>500	>500	>500	>500	**188.4**	**193.4**	>500	>500
X1632_S2_B10	G	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Ce1176_A3	C	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
246-F3_C10_2	AC	>500	>500	>500	>500	>500	>500	**277.1**	**483.5**	>500	>500
CNE55	AE	>500	>500	NT	>500	>500	>500	>500	>500	>500	>500

IgG IC_{50} titers (ug/ml)
0 - 25 µg/ml
25 -100 µg/mll
100 - 400 µg/ml
400 - 500 µg/ml
>500 µg/ml

Table S4A.
Heterologous TZM.bl neutralization data of purified patient IgG (IC50 values, µg/ml), continued

3BNC117-treated patients on-ART

Heterologous Strain	Clade	5A1 Day 0	5A1 Week 24	5A2 Day 0	5A2 Week 24	5A3 Day 0	5A3 Week 24	5A4 Day 0	5A4 Week 24	5A5 Day 0	5A5 Week 24
YU2.DG	B	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
BaL.26	B	>500	>500	**155.0**	**177.1**	**83.9**	**27.3**	>500	**404.5**	>500	>500
Q769.d22	A1	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	>500	>500	>500	**430.6**	>500	>500	>500	>500
Q842.d12	A1	>500	>500	>500	>500	>500	**388.4**	>500	**492.9**	>500	>500
ZM135M.PL10a	C	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
25710-2.43	C	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
CNE8	AE	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
TRO.11	B	>500	>500	>500	>500	>500	**416.3**	>500	>500	>500	>500
X1632_S2_B10	G	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Ce1176_A3	C	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
246-F3_C10_2	AC	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
CNE55	AE	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500

Heterologous Strain	Clade	5B1 Day 0	5B1 Week 24	5B2 Day 0	5B2 Week 24	5B3 Day 0	5B3 Week 24	5B4 Day 0	5B4 Week 24	5B5 Day 0	5B5 Week 24
YU2.DG	B	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
BaL.26	B	**129.2**	**134.1**	**127.9**	**152.5**	>500	>500	**66.8**	**46.4**	**124.8**	**89.5**
Q769.d22	A1	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Q842.d12	A1	>500	>500	>500	**435.0**	>500	>500	>500	>500	**359.3**	**298.1**
ZM135M.PL10a	C	>500	>500	>500	>500	>500	>500	>500	>500	>500	**460.0**
25710-2.43	C	>500	>500	>500	>500	>500	>500	>500	>500	**497.9**	**497.9**
CNE8	AE	>500	>500	>500	>500	>500	>500	>500	>500	>500	**219.8**
TRO.11	B	>500	>500	>500	>500	>500	>500	>500	>500	>500	**487.3**
X1632_S2_B10	G	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
Ce1176_A3	C	>500	>500	>500	>500	>500	>500	>500	>500	>500	>500
246-F3_C10_2	AC	>500	>500	>500	>500	>500	>500	>500	>500	>500	**279.2**
CNE55	AE	>500	>500	>500	>500	>500	>500	>500	>500	>500	**381.2**

Heterologous Strain	Clade	2B2 Day -7	2B2 Week 24	2C2 Day -7	2C2 Week 24
YU2.DG	B	>500	>500	>500	>500
BaL.26	B	**237.0**	**231.1**	**134.3**	**75.8**
Q769.d22	A1	>500	>500	>500	>500
Q259.d2.17	A1	>500	>500	>500	>500
Q842.d12	A1	>500	>500	>500	>500
ZM135M.PL10a	C	>500	>500	>500	>500
25710-2.43	C	>500	>500	>500	>500
CNE8	AE	>500	>500	>500	>500
TRO.11	B	>500	>500	>500	**353.6**
X1632_S2_B10	G	>500	>500	>500	>500
Ce1176_A3	C	>500	>500	>500	>500
246-F3_C10_2	AC	*N/A*	>500	*N/A*	>500
CNE55	AE	>500	>500	>500	>500

IgG IC_{50} titers (ug/ml)
0 - 25 µg/ml
25 -100 µg/ml
100 - 400 µg/ml
400 - 500 µg/ml
>500 µg/ml

Table S4B.
Heterologous TZM.bl neutralization data of purified patient IgG (AUC)
Viremic control patients

Heterologous Strain	Clade	10518 Day 0	10518 Week 26 (182d)	10362 Day 0	10362 Week 27 (190d)	10112 Day 0	10112 Week 24 (171d)	10962 Day 0	10962 Week 24 (168d)	10223 Day 0	10223 Week 28 (196d)
YU2.DG	B	0.24	0.18	0.09	0.13	**2.08**	**1.90**	**1.43**	**1.51**	**1.13**	**1.09**
BaL.26	B	**1.79**	**1.96**	**2.61**	**2.64**	**3.30**	**3.08**	**3.13**	**3.19**	**2.39**	**2.36**
Q769.d22	A1	0.17	0.40	0.37	0.24	**2.70**	**2.49**	**1.57**	**1.60**	**1.52**	**1.41**
Q259.d2.17	A1	0.03	0.19	0.05	0.07	**0.73**	**0.71**	**1.11**	**1.06**	**0.79**	**0.87**
Q842.d12	A1	0.31	0.27	0.16	**0.51**	**2.77**	**2.44**	**2.80**	**2.70**	**2.52**	**2.49**
ZM135M.PL10a	C	0.02	0.31	0.05	0.13	**0.87**	**1.31**	**0.79**	**1.28**	**0.93**	**0.71**
25710-2.43	C	**0.76**	**0.52**	0.32	0.29	**2.32**	**1.98**	**1.18**	**1.36**	**1.36**	**1.34**
CNE8	AE	0.00	0.00	0.00	0.04	**1.01**	**0.66**	**0.88**	**1.04**	**1.43**	**1.47**
TRO.11	B	0.09	0.43	0.23	0.22	**2.33**	**2.29**	**2.32**	**2.21**	**2.01**	**1.70**
X1632_S2_B10	G	0.32	**0.75**	**0.58**	0.15	0.39	**0.64**	**1.09**	**0.91**	**1.98**	**1.61**
Ce1176_A3	C	0.21	0.09	0.35	0.07	**1.94**	**1.65**	**0.84**	**0.99**	**1.43**	**1.58**
246-F3_C10_2	AC	0.00	0.04	0.03	0.09	**1.27**	**1.16**	**0.83**	**0.81**	**1.12**	**1.15**
CNE55	AE	0.00	0.23	0.26	0.00	*N/A*	*N/A*	**1.41**	**1.31**	**2.13**	**1.68**

Heterologous Strain	Clade	10482 Day 0	10482 Week 26 (183d)	10210 Day 0	10210 Week 27 (189d)	10596 Day 0	10596 Week 27 (188d)	10138 Day 0	10138 Week 28 (196d)	10108 Day 0	10108 Week 28 (196d)
YU2.DG	B	0.06	0.09	**0.51**	**0.64**	**1.51**	**1.46**	**2.64**	**2.55**	**1.73**	**1.84**
BaL.26	B	**1.75**	**1.75**	**2.54**	**2.78**	**3.09**	**3.08**	**3.59**	**3.54**	**3.75**	**3.69**
Q769.d22	A1	0.29	**0.63**	0.41	0.08	0.34	0.25	**1.90**	**1.83**	**3.46**	**3.36**
Q259.d2.17	A1	0.18	0.18	0.40	0.26	**0.71**	0.23	**0.66**	**0.68**	**3.28**	**3.26**
Q842.d12	A1	**0.59**	0.49	**0.90**	**0.89**	**1.40**	**1.03**	**3.20**	**3.03**	**3.38**	**3.19**
ZM135M.PL10a	C	0.13	0.09	0.38	0.46	**0.73**	0.36	**0.68**	**0.83**	**1.85**	**1.92**
25710-2.43	C	**0.54**	**0.72**	0.46	**0.58**	**2.20**	**1.71**	**0.81**	**0.79**	**2.74**	**2.72**
CNE8	AE	0.04	0.06	0.00	0.04	0.10	0.07	0.48	0.40	**2.34**	**2.26**
TRO.11	B	**0.93**	**0.95**	**0.73**	**0.58**	**1.20**	**1.47**	**2.70**	**2.54**	**3.38**	**3.27**
X1632_S2_B10	G	0.20	0.02	0.36	0.08	**0.73**	**1.13**	**2.32**	**2.16**	**3.21**	**3.12**
Ce1176_A3	C	0.31	0.24	**0.59**	0.35	**1.76**	**1.53**	0.15	0.34	**1.28**	**1.46**
246-F3_C10_2	AC	0.14	0.13	0.19	0.19	0.49	0.19	**0.92**	**0.65**	**2.54**	**2.53**
CNE55	AE	0.14	0.15	0.28	0.05	0.00	0.00	**1.19**	**1.01**	**2.68**	**2.50**

Heterologous Strain	Clade	10275 Day 0	10275 Week 26 (181d)	10959 Day 0	10959 Week 27 (189d)	10367 Day 0	10367 Week 28 (195d)	10469 Day 0	10469 Week 27 (188d)	10410 Day 0	10410 Week 26 (182d)
YU2.DG	B	0.05	0.16	**2.02**	**2.04**	**2.50**	**2.58**	**0.66**	**0.72**	0.21	0.13
BaL.26	B	**2.32**	**2.19**	**3.02**	**2.99**	**3.38**	**3.37**	**2.32**	**2.16**	**1.69**	**2.24**
Q769.d22	A1	0.10	0.40	0.40	0.45	0.35	0.23	0.31	0.47	0.00	0.00
Q259.d2.17	A1	0.14	0.09	0.23	0.21	**0.93**	**0.74**	**1.35**	**1.03**	0.00	0.01
Q842.d12	A1	0.41	**0.55**	**2.83**	**2.76**	**1.16**	**0.55**	**0.71**	**0.73**	0.49	0.33
ZM135M.PL10a	C	0.35	0.48	0.38	0.31	**0.78**	0.34	0.21	0.12	0.21	0.18
25710-2.43	C	0.46	0.40	**0.58**	**0.57**	**2.48**	**2.46**	**0.86**	**0.97**	0.41	**1.08**
CNE8	AE	0.02	0.00	0.23	0.34	0.14	0.22	0.08	0.08	0.00	0.00
TRO.11	B	0.24	0.27	**1.83**	**1.86**	**3.03**	**3.07**	**0.92**	**0.82**	0.00	0.03
X1632_S2_B10	G	0.20	0.22	**1.07**	**1.19**	**1.26**	**1.41**	**0.69**	**0.65**	0.29	0.23
Ce1176_A3	C	0.44	**0.55**	0.25	0.20	**2.61**	**2.59**	**0.67**	**0.75**	0.05	0.50
246-F3_C10_2	AC	0.17	0.11	**0.76**	**0.85**	**2.06**	**1.88**	0.19	0.19	0.14	0.27
CNE55	AE	0.42	0.05	**1.81**	**1.74**	**0.69**	0.27	0.45	0.11	0.00	0.17

Heterologous Strain	Clade	10397 Day 0	10397 Week 26 (182d)	10292 Day 0	10292 Week 27 (187d)	10802 Day 0	10802 Week 27 (187d)
YU2.DG	B	**0.90**	**1.04**	**0.60**	**0.65**	0.00	0.00
BaL.26	B	**3.15**	**3.16**	**2.68**	**2.55**	0.11	0.00
Q769.d22	A1	0.16	**0.74**	0.20	0.00	0.18	0.11
Q259.d2.17	A1	0.00	0.00	0.19	0.19	0.02	0.02
Q842.d12	A1	0.41	**0.60**	**0.63**	**0.66**	0.32	0.36
ZM135M.PL10a	C	0.02	0.00	0.00	0.00	0.01	0.08
25710-2.43	C	**0.96**	**0.70**	0.32	0.31	0.03	0.18
CNE8	AE	0.21	0.25	**0.52**	0.46	0.06	0.08
TRO.11	B	0.39	**0.60**	**0.71**	0.43	0.00	0.07
X1632_S2_B10	G	0.36	**0.65**	**0.74**	0.47	0.00	0.02
Ce1176_A3	C	0.14	0.19	0.15	0.16	**1.82**	**1.75**
246-F3_C10_2	AC	0.02	0.01	0.05	0.18	0.00	0.02
CNE55	AE	0.09	0.14	0.09	0.02	0.33	0.21

AUC
> 2
1-2
0.5 - 1
0 - 0.5

Table S4B.
Heterologous TZM.bl neutralization data of purified patient IgG (AUC), continued

Heterologous Strain	Clade	10099 Day 0	10099 Week 27 (187d)	10930 Day 0	10930 Week 27 (188d)	10417 Day 0	10417 Week 8 (56d)	10690 Day 0	10690 Week 16 (112d)	10257 Day 0	10257 Week 28 (196d)
YU2.DG	B	0.07	0.09	**1.24**	**1.08**	0.48	**0.51**	**1.31**	**1.37**	**0.75**	**0.83**
BaL.26	B	**1.81**	**1.95**	**2.77**	**2.72**	**3.24**	**3.31**	**2.38**	**2.16**	**2.44**	**2.40**
Q769.d22	A1	0.00	0.07	**1.27**	**1.06**	0.21	0.08	0.20	**0.52**	0.26	0.18
Q259.d2.17	A1	0.04	0.06	**0.86**	**0.60**	0.19	0.13	0.03	0.09	**0.50**	**0.87**
Q842.d12	A1	0.26	0.39	**1.85**	**1.61**	**0.59**	**1.03**	0.43	**0.63**	**1.26**	**1.69**
ZM135M.PL10a	C	0.06	0.06	0.08	0.01	**0.76**	**1.26**	0.19	0.15	**0.69**	**1.24**
25710-2.43	C	0.38	**0.64**	**0.65**	**0.54**	**1.01**	**1.41**	**0.73**	**0.77**	**0.97**	**1.48**
CNE8	AE	**0.56**	0.37	**2.33**	**2.18**	**1.88**	**2.21**	**0.54**	**0.54**	**0.99**	**1.18**
TRO.11	B	**0.81**	**0.85**	**1.53**	**1.19**	**1.48**	**1.64**	**1.06**	**1.02**	**1.69**	**1.77**
X1632_S2_B10	G	0.09	0.00	**1.68**	**1.29**	0.46	0.30	**0.81**	**1.18**	**0.61**	**0.90**
Ce1176_A3	C	**0.67**	0.41	**0.82**	**0.73**	**0.68**	**0.69**	**0.60**	**0.64**	0.28	0.31
246-F3_C10_2	AC	0.11	0.01	**1.82**	**1.75**	0.28	0.31	0.07	0.08	0.13	0.21
CNE55	AE	0.03	0.09	**1.57**	**1.54**	**0.64**	0.43	0.18	0.14	0.41	0.08

Heterologous Strain	Clade	10587 Day 0	10587 Week 17 (117d)	10169 Day 0	10169 Week 88 (617d)	10160 Day 0	10160 Week 15 (106d)	10811 Day 0	10811 Week 31 (215d)	10239 Day 0	10239 Week 35 (245d)
YU2.DG	B	0.40	0.29	**2.34**	**1.98**	**0.64**	**0.51**	0.43	0.48	0.21	0.20
BaL.26	B	**1.95**	**1.83**	**3.46**	**3.36**	**2.30**	**2.33**	**1.54**	**1.53**	**1.66**	**1.53**
Q769.d22	A1	**1.38**	**0.79**	**2.36**	**2.04**	0.44	0.28	0.21	**0.62**	0.12	0.49
Q259.d2.17	A1	0.18	0.21	**1.51**	**1.30**	0.16	0.03	0.04	0.10	0.12	0.01
Q842.d12	A1	**0.75**	**0.87**	**2.55**	**2.16**	0.26	0.20	0.39	**1.10**	0.23	0.22
ZM135M.PL10a	C	0.15	0.05	**1.70**	**0.98**	0.40	0.24	0.06	0.19	0.32	0.16
25710-2.43	C	0.49	0.42	**1.48**	**2.07**	**0.82**	**0.63**	**0.69**	**1.40**	**0.70**	**0.54**
CNE8	AE	0.28	0.35	**1.96**	**1.62**	0.41	0.32	0.07	0.25	0.18	0.12
TRO.11	B	**1.66**	**1.38**	**2.85**	**2.58**	**1.70**	**1.57**	**0.93**	**1.55**	0.29	0.26
X1632_S2_B10	G	**0.52**	0.49	**1.60**	**1.96**	0.40	**0.71**	0.29	0.16	0.35	0.19
Ce1176_A3	C	0.27	0.44	**1.47**	**1.14**	**0.83**	**0.69**	**0.66**	**1.31**	0.29	0.19
246-F3_C10_2	AC	0.04	0.07	**1.22**	**0.62**	0.08	0.07	0.16	**0.75**	0.21	0.10
CNE55	AE	**0.63**	**0.57**	**1.61**	**1.46**	0.40	**0.59**	**0.79**	**1.07**	0.41	0.29

Heterologous Strain	Clade	10814 Day 0	10814 Week 26 (181d)	10100 Day 0	10100 Week 11 (76d)	10779 Day 0	10779 Week 13 (89d)	10785 Day 0	10785 Week 31 (218d)	10489 Day 0	10489 Week 24 (169d)
YU2.DG	B	0.09	0.25	0.00	0.00	0.02	0.00	0.25	0.15	**0.70**	**0.63**
BaL.26	B	**1.21**	**0.99**	**1.36**	**1.30**	**2.81**	**2.54**	**2.19**	**2.05**	**2.11**	**2.31**
Q769.d22	A1	0.30	0.10	0.00	0.03	0.00	0.07	0.17	0.07	0.00	0.09
Q259.d2.17	A1	0.16	0.10	0.00	0.07	0.00	0.00	0.18	0.10	0.00	0.05
Q842.d12	A1	0.09	0.13	0.00	0.21	0.01	0.04	0.38	0.40	0.12	0.15
ZM135M.PL10a	C	0.20	0.34	0.00	0.00	0.25	0.25	0.29	0.25	0.11	0.10
25710-2.43	C	0.18	0.23	0.08	0.18	0.44	0.38	**1.14**	**1.14**	0.47	**0.60**
CNE8	AE	**0.79**	**0.81**	0.02	0.16	**0.70**	0.46	0.18	0.18	0.00	0.04
TRO.11	B	**1.11**	**1.05**	0.00	0.02	0.30	0.18	0.30	0.21	**1.68**	**1.59**
X1632_S2_B10	G	0.15	0.36	0.00	0.00	0.09	0.06	0.31	0.23	0.00	0.05
Ce1176_A3	C	0.29	0.35	0.20	0.01	0.08	0.19	0.36	0.38	0.21	0.29
246-F3_C10_2	AC	0.15	0.21	0.38	0.08	0.12	0.11	0.31	0.31	0.08	0.14
CNE55	AE	0.19	0.14	0.02	**0.87**	0.00	0.19	0.22	0.12	0.00	0.00

Heterologous Strain	Clade	10082 Day 0	10082 Week 30 (209d)	10914 Day 0	10914 Week 32 (224d)	10689 Day 0	10689 Week 27 (188d)
YU2.DG	B	0.08	0.00	**0.88**	0.28	0.10	0.08
BaL.26	B	**2.46**	**2.37**	**2.31**	**1.45**	**2.33**	**1.97**
Q769.d22	A1	0.13	0.01	0.08	0.20	**0.51**	0.22
Q259.d2.17	A1	0.06	0.04	0.02	0.01	0.26	0.25
Q842.d12	A1	0.06	0.00	0.13	0.04	0.44	0.35
ZM135M.PL10a	C	0.12	0.15	**0.77**	0.46	0.30	0.40
25710-2.43	C	0.11	0.07	**0.92**	**0.52**	0.44	0.45
CNE8	AE	0.00	0.00	**0.62**	0.25	0.01	0.00
TRO.11	B	0.45	0.20	**0.53**	0.20	0.16	0.24
X1632_S2_B10	G	0.13	0.13	0.08	0.00	0.03	0.00
Ce1176_A3	C	0.05	0.02	**0.60**	0.27	0.30	0.30
246-F3_C10_2	AC	0.10	0.11	0.18	0.10	0.24	0.31
CNE55	AE	0.00	0.00	0.00	0.00	0.20	0.11

AUC
> 2
1-2
0.5 - 1
0 - 0.5

Table S4B.
Heterologous TZM.bl neutralization data of purified patient IgG (AUC), continued
3BNC117-treated viremic patients

Heterologous Strain	Clade	2A1		2A3		2A4		2B1		2B3	
		Day -7	Week 24	Day 0	Week 24	Day -7	Week 24	Day -7	Week 24	Day -7	Week 24
YU2.DG	B	**2.51**	**2.60**	0.21	**0.61**	0.09	0.37	0.03	0.00	0.00	0.02
BaL.26	B	**3.72**	**3.70**	**2.35**	**2.55**	**1.41**	**1.64**	**1.03**	**0.94**	0.41	**0.52**
Q769.d22	A1	**1.72**	**1.64**	0.19	**2.40**	0.00	0.06	0.07	0.00	0.00	0.01
Q259.d2.17	A1	**0.91**	**1.76**	**1.02**	**1.61**	0.01	0.20	0.05	**0.51**	0.00	0.20
Q842.d12	A1	**2.53**	**2.67**	0.36	**0.56**	0.04	0.19	0.01	0.00	0.00	0.04
ZM135M.PL10a	C	0.25	**0.78**	0.12	0.32	0.01	0.12	0.09	0.08	0.00	0.22
25710-2.43	C	0.32	**0.73**	**0.96**	**1.33**	0.48	**0.58**	0.30	0.24	0.01	0.18
CNE8	AE	0.00	0.13	0.03	0.02	0.00	0.01	0.01	0.00	0.00	0.02
TRO.11	B	**1.04**	**1.88**	**0.67**	**1.10**	0.27	**0.70**	0.01	0.24	0.00	0.49
X1632_S2_B10	G	**1.66**	**1.91**	0.22	0.29	0.06	0.10	0.09	0.00	0.00	0.04
Ce1176_A3	C	0.00	0.28	0.21	0.44	0.00	0.16	0.02	0.15	0.01	0.14
246-F3_C10_2	AC	*N/A*	*N/A*	**0.53**	**0.60**	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*
CNE55	AE	0.34	**0.81**	0.12	0.13	0.00	0.00	0.03	0.00	0.02	0.02

Heterologous Strain	Clade	2C4		2C5		2D1		2C1		2D3	
		Day -7	Week 24	Day -7	Week 24	Day 0	Week 24	Day -7	Week 24	Day -7	Week24
YU2.DG	B	**1.10**	**1.64**	0.04	0.11	0.16	0.06	0.14	0.15	0.00	0.14
BaL.26	B	**2.10**	**2.81**	**2.16**	**2.75**	**1.11**	**1.45**	**1.18**	**1.71**	0.00	0.26
Q769.d22	A1	0.03	0.43	0.08	0.38	0.04	0.26	0.24	0.28	0.01	**0.61**
Q259.d2.17	A1	*N/A*	*N/A*	*N/A*	*N/A*	0.20	0.20	0.07	0.42	0.01	0.25
Q842.d12	A1	**1.42**	**1.36**	0.02	0.12	0.09	0.23	0.19	**0.59**	0.00	0.46
ZM135M.PL10a	C	0.11	**0.55**	0.13	**0.86**	0.16	0.24	0.09	0.23	0.09	0.31
25710-2.43	C	**0.62**	**1.88**	**0.55**	**0.98**	0.38	0.29	0.47	**1.15**	0.01	0.00
CNE8	AE	0.47	**1.48**	0.06	**0.79**	0.00	0.00	0.03	0.16	0.00	0.02
TRO.11	B	**1.61**	**2.86**	0.38	**0.83**	0.34	0.15	0.16	0.42	0.02	0.12
X1632_S2_B10	G	0.13	**0.80**	0.08	0.42	0.18	0.05	0.13	0.15	0.01	0.03
Ce1176_A3	C	0.24	**1.53**	0.00	0.04	0.12	0.00	0.04	0.07	0.08	0.00
246-F3_C10_2	AC	*N/A*	*N/A*	*N/A*	*N/A*	0.22	0.02	*N/A*	*N/A*	0.18	0.00
CNE55	AE	0.04	0.07	0.02	0.00	0.00	0.01	0.04	0.00	0.02	0.08

Heterologous Strain	Clade	2E1		2E2		2E3		2E4		2E5	
		Day 0	Week 24	Day 0	Week 24	Day 0	Week 24	Day 0	Week 24	Day 0	Week 24
YU2.DG	B	0.00	0.28	0.08	0.30	0.00	0.13	0.28	**0.52**	0.00	**0.53**
BaL.26	B	**1.40**	**2.12**	**2.24**	**2.71**	**1.27**	**2.33**	**1.26**	**1.74**	**0.63**	**1.77**
Q769.d22	A1	0.48	**1.05**	0.00	0.50	0.00	**0.80**	0.47	0.40	0.17	**0.73**
Q259.d2.17	A1	0.03	0.24	0.30	0.14	0.30	0.18	0.34	**0.62**	0.13	**0.73**
Q842.d12	A1	0.02	0.19	0.12	0.19	0.16	0.24	0.36	**0.59**	0.25	**0.56**
ZM135M.PL10a	C	0.00	0.14	0.24	0.13	0.15	0.37	**0.53**	0.35	0.12	0.21
25710-2.43	C	**0.60**	**0.65**	0.47	0.46	0.42	**0.61**	**1.02**	**0.99**	0.07	0.22
CNE8	AE	0.05	0.07	0.03	0.15	0.04	0.02	**0.93**	**1.32**	0.05	0.10
TRO.11	B	**0.65**	**1.14**	0.40	0.18	0.34	0.38	**0.99**	**0.80**	0.14	0.30
X1632_S2_B10	G	0.16	0.26	0.22	0.41	0.01	0.32	0.21	0.30	0.04	**0.51**
Ce1176_A3	C	0.21	0.02	0.12	0.00	0.05	0.00	0.34	0.13	0.12	0.15
246-F3_C10_2	AC	0.18	**0.51**	0.39	0.14	0.17	0.14	**0.83**	**0.57**	0.19	**1.13**
CNE55	AE	0.00	0.00	*N/A*	*N/A*	0.00	0.00	0.00	0.00	0.00	0.03

AUC
> 2
1-2
0.5 - 1
0 - 0.5

Table S4B.
Heterologous TZM.bl neutralization data of purified patient IgG (AUC), continued
3BNC117-treated patients on-ART

Heterologous Strain	Clade	5A1 Day 0	5A1 Week 24	5A2 Day 0	5A2 Week 24	5A3 Day 0	5A3 Week 24	5A4 Day 0	5A4 Week 24	5A5 Day 0	5A5 Week 24
YU2.DG	B	0.04	0.02	0.03	0.01	0.07	0.18	0.04	0.03	0.06	0.05
BaL.26	B	0.05	0.02	**0.91**	**0.87**	**1.30**	**2.31**	0.26	0.44	0.12	0.11
Q769.d22	A1	0.01	0.46	0.01	0.20	0.01	0.07	0.03	0.18	0.21	0.20
Q259.d2.17	A1	0.20	0.21	0.09	0.27	0.33	0.41	0.23	0.20	0.09	0.20
Q842.d12	A1	0.18	0.29	0.25	0.29	0.20	**0.52**	0.36	0.38	0.18	0.34
ZM135M.PL10a	C	0.00	0.09	0.01	0.21	0.07	0.07	0.17	0.05	0.04	0.04
25710-2.43	C	0.01	0.08	0.05	0.23	0.12	0.23	0.18	0.17	0.42	0.46
CNE8	AE	0.00	0.00	0.00	0.03	0.01	0.13	0.02	0.00	0.09	0.08
TRO.11	B	0.11	0.09	0.17	0.05	0.36	0.49	0.12	0.11	0.24	0.27
X1632_S2_B10	G	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.05	0.14	0.09
Ce1176_A3	C	0.00	0.00	0.03	0.00	0.00	0.12	0.00	0.02	0.03	0.02
246-F3_C10_2	AC	0.03	0.01	0.01	0.00	0.00	0.02	0.00	0.00	0.00	0.00
CNE55	AE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Heterologous Strain	Clade	5B1 Day 0	5B1 Week 24	5B2 Day 0	5B2 Week 24	5B3 Day 0	5B3 Week 24	5B4 Day 0	5B4 Week 24	5B5 Day 0	5B5 Week 24
YU2.DG	B	0.00	0.02	0.03	0.11	0.00	0.02	0.07	0.05	0.28	0.19
BaL.26	B	**1.18**	**1.22**	**1.37**	**1.10**	0.15	0.21	**1.77**	**1.85**	**1.32**	**1.43**
Q769.d22	A1	0.00	0.00	**0.54**	**0.57**	0.27	0.14	0.30	0.41	**0.73**	**0.99**
Q259.d2.17	A1	0.00	0.10	0.18	0.11	0.00	0.05	0.01	0.04	0.29	0.23
Q842.d12	A1	0.10	0.08	0.18	0.41	0.08	0.20	0.19	0.12	**0.56**	**0.59**
ZM135M.PL10a	C	0.00	0.02	0.18	0.13	0.03	0.14	0.18	0.25	0.30	**0.52**
25710-2.43	C	0.14	0.20	0.29	0.21	0.06	0.08	0.18	0.20	0.48	0.44
CNE8	AE	0.02	0.00	0.05	0.04	0.00	**0.80**	0.04	**0.84**	0.26	**1.08**
TRO.11	B	0.00	0.01	0.13	0.12	0.00	0.37	0.06	0.21	0.26	**0.87**
X1632_S2_B10	G	0.00	0.00	0.10	0.19	0.00	0.36	0.04	0.41	0.07	**0.80**
Ce1176_A3	C	0.01	0.04	0.08	0.10	0.06	0.13	0.12	0.09	0.10	**0.53**
246-F3_C10_2	AC	0.03	0.11	0.12	0.15	0.07	0.24	0.16	0.25	0.14	**0.80**
CNE55	AE	0.00	0.00	0.07	0.25	0.16	0.26	0.14	0.26	0.14	**1.08**

Heterologous Strain	Clade	2B2 Day -7	2B2 Week 24	2C2 Day -7	2C2 Week 24
YU2.DG	B	0.00	0.03	0.00	0.02
BaL.26	B	**0.68**	**0.71**	**1.03**	**1.40**
Q769.d22	A1	0.20	0.00	0.01	0.07
Q259.d2.17	A1	0.00	0.09	0.00	0.29
Q842.d12	A1	0.00	0.14	0.04	0.30
ZM135M.PL10a	C	0.11	0.14	0.16	0.29
25710-2.43	C	0.39	0.49	0.28	**0.65**
CNE8	AE	0.00	0.03	0.00	0.12
TRO.11	B	0.04	0.30	0.11	**0.74**
X1632_S2_B10	G	0.14	0.08	0.07	0.23
Ce1176_A3	C	0.02	0.07	0.12	0.22
246-F3_C10_2	AC	*N/A*	*N/A*	*N/A*	*N/A*
CNE55	AE	0.23	0.00	0.27	0.02

AUC
> 2
1-2
0.5 - 1
0 - 0.5

Table S5A.
Summary measures table breadth and potency Day 0 and Wk24
Untreated viremic control individuals

		Breadth				Overall neutralizing activity (Potency and breadth)	
Subject ID	**Status before therapy**	**Number of viruses neutralized with AUC > 0.5 pre-therapy**	**Percent of viruses neutralized with AUC > 0.5 pre-therapy**	**Number of viruses neutralized with AUC > 0.5 post-therapy**	**Percent of viruses neutralized with AUC > 0.5 post-therapy**	**Mean AUC Day 0**	**Mean AUC Wk 24**
Untreated viremic control individuals							
10518	Control	2	15%	3	23%	0.305	0.413
10362	Control	2	15%	2	15%	0.393	0.352
10112	Control	11	92%	12	100%	1.809	1.692
10962	Control	13	100%	13	100%	1.492	1.536
10223	Control	13	100%	13	100%	1.595	1.497
10482	Control	4	31%	4	31%	0.407	0.423
10210	Control	5	38%	5	38%	0.594	0.537
10596	Control	9	69%	7	54%	1.096	0.962
10138	Control	11	85%	11	85%	1.634	1.566
10108	Control	13	100%	13	100%	2.739	2.701
10275	Control	1	8%	3	23%	0.409	0.422
10959	Control	8	62%	8	62%	1.186	1.193
10367	Control	11	85%	9	69%	1.645	1.514
10469	Control	8	62%	8	62%	0.724	0.676
10410	Control	1	8%	2	15%	0.267	0.396
10397	Control	3	23%	7	54%	0.525	0.623
10292	Control	6	46%	3	23%	0.530	0.468
10802	Control	1	8%	1	8%	0.222	0.225
10099	Control	4	31%	3	23%	0.376	0.384
10930	Control	12	92%	12	92%	1.420	1.254
10417	Control	8	62%	8	62%	0.915	1.023
10690	Control	7	54%	9	69%	0.656	0.714
10257	Control	9	69%	9	69%	0.845	1.009
10587	Control	6	46%	5	38%	0.669	0.598
10169	Control	13	100%	13	100%	2.009	1.789
10160	Control	5	38%	7	54%	0.682	0.627
10811	Control	5	38%	8	62%	0.481	0.807
10239	Control	2	15%	2	15%	0.391	0.330
10814	Control	3	23%	3	23%	0.378	0.390
10100	Control	1	8%	2	15%	0.159	0.225
10779	Control	2	15%	1	8%	0.371	0.343
10785	Control	2	15%	2	15%	0.482	0.429
10489	Control	3	23%	4	31%	0.422	0.465
10082	Control	1	8%	1	8%	0.288	0.238
10914	Control	7	54%	2	15%	0.548	0.291
10689	Control	2	15%	1	8%	0.410	0.359
Average :		5.9	45.9%	6.0	46.4%	0.8	0.8

Table S5B.
Summary measures table breadth and potency Day 0 and Wk24
3BNC117-treated individuals

		Breadth				Overall neutralizing activity (Potency and breadth)	
Subject ID	**Status before therapy**	**Number of viruses neutralized with AUC > 0.5 pre-therapy**	**Percent of viruses neutralized with AUC > 0.5 pre-therapy**	**Number of viruses neutralized with AUC > 0.5 post-therapy**	**Percent of viruses neutralized with AUC > 0.5 post-therapy**	**Mean AUC Day 0**	**Mean AUC Wk 24**
Viremic individuals							
2A1	Viremic	7	58%	10	83%	1.250	1.575
2A3	Viremic	5	38%	8	62%	0.537	0.921
2A4	Viremic	1	8%	3	25%	0.196	0.345
2B1	Viremic	1	8%	2	17%	0.144	0.180
2B3	Viremic	0	0%	1	8%	0.038	0.159
2C4	Viremic	5	45%	9	82%	0.717	1.403
2C5	Viremic	2	18%	5	45%	0.321	0.662
2D1	Viremic	1	8%	1	8%	0.230	0.227
2C1	Viremic	1	8%	3	25%	0.232	0.444
2D3	Viremic	0	0%	1	8%	0.033	0.176
2E1	Viremic	3	23%	5	38%	0.292	0.513
2E2	Viremic	1	8%	1	8%	0.384	0.442
2E3	Viremic	1	8%	3	23%	0.224	0.425
2E4	Viremic	6	46%	8	62%	0.581	0.641
2E5	Viremic	1	8%	7	54%	0.147	0.537
Average :		2.3	19.1%	4.5	36.5%	0.4	0.6
Individuals on-ART							
5A1	on-ART	0	0%	0	0%	0.0482	0.0976
5A2	on-ART	1	8%	1	8%	0.1204	0.1665
5A3	on-ART	1	8%	2	15%	0.1897	0.3528
5A4	on-ART	0	0%	0	0%	0.1090	0.1259
5A5	on-ART	0	0%	0	0%	0.1255	0.1437
5B1	on-ART	1	8%	1	8%	0.1137	0.1382
5B2	on-ART	2	15%	2	15%	0.2535	0.2679
5B3	on-ART	0	0%	1	8%	0.0685	0.2294
5B4	on-ART	1	8%	2	15%	0.2510	0.3820
5B5	on-ART	3	23%	10	77%	0.3799	0.7352
2B2	on-ART	1	8%	1	8%	0.1501	0.1726
2C2	on-ART	1	8%	3	25%	0.1753	0.3616
Average :		0.9	7.2%	1.9	15.0%	0.2	0.3

Table S6.
TZM.bl neutralization data over time in subject 2A3 (IC50)

Heterologous strain	Clade	Day 0	Week 12	Week 16	Week 20	Week 24
YU2.DG	B	>500	**483.2**	**403.7**	**390.1**	**363.4**
BaL.26	B	**28.4**	**24.2**	**25.5**	**24.3**	**18.9**
Q769.d22	A1	>500	**452.0**	**266.2**	**82.8**	**19.4**
Q259.d2.17	A1	**92.5**	**61.4**	**71.9**	**93.0**	**84.7**
Q842.d12	A1	**475.0**	**499.7**	**436.0**	**426.8**	**417.8**
ZM135M.PL10a	C	>500	>500	>500	>500	>500
25710-2.43	C	**110.8**	**109.7**	**159.8**	**123.1**	**136.0**
CNE8	AE	>500	>500	>500	>500	>500
TRO.11	B	**218.6**	**435.4**	**401.5**	**257.5**	**203.0**
X1632_S2_B10	G	>500	>500	>500	>500	>500
Ce1176_A3	C	**482.2**	>500	**493.4**	**432.9**	**446.7**
246-F3_C10_2	AC	**445.4**	**430.4**	**420.3**	**299.9**	**283.7**
CNE55	AE	>500	>500	>500	>500	>500

IgG IC_{50} titers (µg/ml)

0 - 25 µg/ml
25 -100 µg/ml
100 - 400 µg/ml
400 - 500 µg/ml
>500 µg/ml

Table S7.
Analysis of potential confounding variables

	Variable (p-value of likelihood ratio test if alternative model superior to null-model)							
AUC change observed in viral strain	**Age**	**Time since diagnosis**	**Sampling interval**	**abs. $CD4^+$ T cell count ($cells/mm^3$) Day 0**	**abs. $CD4^+$ T cell count ($cells/mm^3$) Week 24**	**HIV-RNA level (copies/ml) Day 0**	**HIV-RNA level (copies/ml) Week 24**	**Starting AUC (AUC Day 0)**
246.F3_C10_2	1	1	1	1	1	1	1	1
CNE55	1	1	0.18	1	1	1	1	0.18
Ce1176_A3	1	1	1	1	1	1	1	1
TRO.11	1	1	0.41	1	1	1	0.23	1
Q259.d2.17	1	1	0.10	1	1	1	1	1
25710-2.43	1	1	1	1	1	1	1	0.66
Q769.d22	1	1	0.64	1	1	1	1	0.50
BaL.26	1	0.50	0.02	1	0.09	1	1	0.47
YU.2	1	1	0.52	1	1	1	0.69	1
X1632_S2_B10	1	1	0.05	1	1	1	0.92	0.17
ZM135M.PL10a	1	1	0.37	1	1	1	1	1
Q842.d12	1	1	1	1	0.56	1	1	0.0035
CNE8	0.46	1	0.58	1	1	1	1	1

Bonferroni-corrected significance threshold: 0.0005 with significance level $\alpha = 0.05$

variable is not predictive of AUC change

Table S8.
Rank correlation and p-values for Spearman rank correlation of patient and 3BNC117 neutralization

Subject ID	rank correlation (AUC-based)	p-value (AUC-based)	rank correlation (IC_{50}-based)	p-value (IC_{50}-based)
2A1	0.46	0.1689	-0.43	0.1369
2A3	-0.33	0.2470	0.35	0.2664
2A4	-0.52	0.1926	0.41	0.0818
2B1	0.02	0.9224	-0.03	0.9611
2B3	0.14	0.5135	-0.21	0.6641
2C4	0.39	0.2250	-0.40	0.2334
2C5	-0.12	0.8815	0.05	0.7287
2D1	-0.65	0.0103	0.70	0.0163
2C1	-0.30	0.3425	0.30	0.3355
2D3*	-0.84	0.0001	0.87	0.0003
2E1	-0.62	0.0160	0.66	0.0225
2E2	-0.43	0.1926	0.41	0.1659
2E3	-0.24	0.5292	0.19	0.4364
2E4	-0.48	0.0850	0.50	0.1002
2E5	-0.38	0.1013	0.48	0.1972
2B2	-0.14	0.8690	0.06	0.6641
2C2	-0.09	0.8517	0.06	0.7700
5A1	0.00	0.8922	0.04	1.0000
5A2	0.12	0.9786	-0.01	0.7029
5A3	-0.60	0.1013	0.48	0.0293
5A4	-0.19	0.4930	0.21	0.5339
5A5	-0.21	0.4561	0.23	0.4988
5B1	0.23	0.6138	-0.15	0.4505
5B2	0.01	0.8632	-0.05	0.9644
5B3	0.24	0.3103	-0.31	0.4330
5B4	0.26	0.5785	-0.17	0.3936
5B5	0.43	0.1955	-0.38	0.1377

p-value threshold (Bonferroni-corrected): 0.0019

*Significant rank correlation only detected in 2D3 (indicated in green).

Table S9.
Number of gp160 nucleotide sequences included in phylogenetic analysis

Subject	Day 0	Week 4 (6)	Week 12	Week 24	Total per patient
2A1	27	26	24	40	117
2A3	23	32	29	27	111
2C4	33	19	30	19	101
2C5	48	20	22	32	122
2E1	30	31	27	37	125
2E2	20	27	25	27	99
2E3	25	17	24	21	87
2E4	41	25	26	29	121
2E5	47	3 (22*)	23	22	117
Total per TP	294	222	230	254	1000

(*) due to low viral load week 6 plasma was sequenced

Table S10.
Envelope amino acid residues under significant selection in each individual

Subject ID	Sites Under Selection (≥80% different from Day 0 Consensus)
2A1	12,134-6, 138-46, 148-9, 149lmopstuv, 150-3, 155, 161, 164, 166, 168-71, 177, 178, 182, 185-188, 189bdef, 195, 209, 240, 270, 272, 279*, 283, 287, 291, 293, 306, 316-7, 318, 321a, 326, 333, 335, 336-7, 343, 344, 346-7, 360, 362, 364, 365, 373, 386, 395, 398, 401-5, 408, 411-3, 417, 442, 460, 461, 471, 698
2A3	137-8, 140, 141af, 142, 187e, 336, 410
2C4	59, 188ckl, 283, 308, 339, 346-7, 354, 356, 362, 386, 389, 392-6, 400-5, 410, 415, 459abhijkl, 640, 674, 689, 742
2C5	140, 275, 279, 290, 340, 344, 347, 350, 353, 355, 364a, 396, 397, 400-3, 405-7, 409, 461, 461abce, 464, 465, 500
2E1	141, 147-9, 150b, 160, 398, 401-4, 410, 460, 463, 500, 624
2E2	153, 295, 297, 344, 397-8, 411, 413, 444, 461, 636
2E3	32, 145, 440, 464, 683
2E4	187dehi, 337, 340, 343
2E5	80, 282, 400

*3BNC117 contact sites confirmed by crystal structures (*26, 27*) are in red

Table S11.
TZM.bl neutralization data of CMV-env pseudotyped viruses (IC_{50})

Subject ID	Virus ID (in order of appearance in tree top to bottom)	3BNC117 IC_{50} (µg/ml)	Day 0 IgG IC_{50} (µg/ml)	Week 24 IgG IC_{50} (µg/ml)
2A1	2A1-D0-0408TIT-E5_S57	0.119	101.1	33.4
	2A1_D0_0409_G3_S89	0.209	333.6	66.6
	2A1-D0-0409-B1_S73	0.116	161.5	70.1
	2A1-W4-0408TIT-G11_S19	0.157	233.4	109.2
	2A1-D0-0409-E7_S85	0.141	96.9	94.5
	2A1-W24-0409-H4_S33	0.306	185.1	199.5
	2A1-W24-0409-E1_S22	0.641	171.3	157.6
	2A1-W12-TIT0402-D1_S36	0.691	154.3	160.3
	2A1-W24-0409-G11_S30	0.585	175.6	187.1
	2A1-D0-0409-C2_S77	0.692	149.4	135.1
	2A1-W4-0408TIT-H2_S16	0.595	164.2	153.9
	2A1-D0-0409-G9_S90	0.651	165.8	154.5
	2A1-W4-0408TIT-E12-S4	0.184	198.4	96.3
	2A1-W4-0408TIT-G9_S12	0.181	266.1	143.2
	2A1-W4-0408TIT-G10_S13	0.128	164.9	85.4
	2A1-D0-0409-B4-S74	0.317	423.9	186.6
	2A1-W24-0409-H11_S34	0.240	176.1	113.6
	2A1-W4-0408TIT-F8_S6	0.159	239.1	127.1
	2A1-D0-0409-H9_S93	0.008	123.4	139.4
	2A1-D0-0409-E5_S83	0.017	187.2	180.6
2E1	2E1-D0-0409-E4_S58	0.038	107.7	77.3
	2E1-W4_G11_S39	0.185	257.7	23.8
	2E1-W4-TIT0408-D3_S63	0.048	338.1	29.1
	2E1-D0-0409-H8_S69	0.124	211.0	73.6
	2E1-W4-TIT0408-D7_S64	0.106	384.6	38.0
	2E1-W12-0409-H3_S20	0.025	>500.0	233.5
	2E1-W12-0409-G3_S16	0.063	>500.0	>500.0
	2E1-W24-0526PlE-A9_S72	0.052	333.9	154.8
	2E1-W24-0526PlE-B3_S74	0.080	>500.0	452.1
	2E1-D0-0421-G4_S4	0.091	421.4	97.6
	2E1-W4-TIT0408-E7_S69	0.078	360.5	33.5
	2E1-W12-0524Pl2-F6_S49	0.087	>500.0	94.6
	2E1-W24-0526PlE-G6_S85	0.064	266.7	189.6
	2E1-W24-0526PlE-F1e_S80	0.040	370.5	197.9
2C5	2C5-W24-0423-H6_S78	>20	22.5	30.2
	2C5-D0-0422-H4-S56	0.006	36.4	35.9
	2C5-D0-0422-F4-S44	0.011	20.3	34.5
	2C5-W24-0423-F2-S73	0.022	145.7	500.0
	2C5-W4-0423-A6-S33	3.131	18.1	20.6
	2C5-W4-0423-D11_S38	14.079	49.9	12.2
	2C5-W12-TIT0418-D2_S6	1.142	2.5	2.8
	2C5-W4-0422-G1_S91	>20	85.1	38.1
2A3	2A3-W24-TIT0406-B6_S68	0.195	>500.0	407.2
	2A3_D0_0404_C9_S75	0.055	543.9	210.4
	2A3-D0-0421-A3_S44	0.171	>500.0	>500.0
	2A3-W4-0421-E11_S55	1.024	255.2	218.5
	2A3-W24-TIT0406-B7_S69	0.022	411.3	419.5
	2A3-W24-TIT0406-A6e_S65	0.071	292.9	327.8
	2A3-W12-0421-A7_S84	0.051	255.2	182.0
	2A3-D0-0421-D3_S50	0.037	305.6	212.7
	2A3-W24-TIT0406-C11_S74	0.049	224.0	223.0
	2A3-W12-0421-C12_S94	0.055	293.6	263.7
	2A3-W12-0406-F2_S50	0.049	276.5	168.6
	2A3-W4-0406-E12_S31	0.030	169.3	11.1
	2A3_D0_0404_C1_S74	0.074	447.4	155.9
	2A3-W4-0406-A3_S14	0.099	>500.0	219.3
2C4	2C4-W24-0406-A5_S65	>20	240.8	36.1
	2C4-W24-0406-H10_S88	>20	156.7	39.2
	2C4-W24-0406-D6_S74	>20	153.9	39.0
	2C4-D0-0404-C1_S30	>20	195.7	21.9
	2C4-D0-0406-H7_S24	>20	155.3	20.7
	2C4-D0-0406-F3_S19	>20	147.6	20.4
	2C4-D0-0406-E10_S18	>20	170.9	23.8

3BNC117 IC_{50} (µg/ml)	IgG IC_{50} (µg/ml)
0 - 0.1	0 - 50
0.1 - 0.5	50 -100
0.5 - 1.0	100 - 150
1.0 - 2.0	150 - 200
2.0 - 5.0	200 - 250
5.0 - 10.0	250 - 500
> 10.0	> 500

A.3. Paper 3 - bNAb efficacy study II

Copyright clearance

The herein included manuscript is the version of record of the paper "HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption" as published in *Nature 535* and available at link 10.1038/nature18929.

LETTER

doi:10.1038/nature18929

HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption

Johannes F. Scheid[1,2]*, Joshua A. Horwitz[1]*, Yotam Bar-On[1], Edward F. Kreider[3], Ching-Lan Lu[1], Julio C. C. Lorenzi[1], Anna Feldmann[4], Malte Braunschweig[1], Lilian Nogueira[1], Thiago Oliveira[1], Irina Shimeliovich[1], Roshni Patel[1], Leah Burke[5], Yehuda Z. Cohen[1], Sonya Hadrigan[1], Allison Settler[1], Maggi Witmer-Pack[1], Anthony P. West Jr[6], Boris Juelg[7], Tibor Keler[8], Thomas Hawthorne[8], Barry Zingman[9], Roy M. Gulick[5], Nico Pfeifer[4], Gerald H. Learn[3], Michael S. Seaman[10], Pamela J. Bjorkman[6], Florian Klein[1,11,12], Sarah J. Schlesinger[1], Bruce D. Walker[7,13], Beatrice H. Hahn[3], Michel C. Nussenzweig[1,14] & Marina Caskey[1]

Interruption of combination antiretroviral therapy in HIV-1-infected individuals leads to rapid viral rebound. Here we report the results of a phase IIa open label clinical trial evaluating 3BNC117, a broad and potent neutralizing antibody against the CD4 binding site of the HIV-1 Env protein[1], during analytical treatment interruption in 13 HIV-1-infected individuals. Participants with 3BNC117-sensitive virus outgrowth cultures were enrolled. Results show that two or four 30 mg kg^{-1} 3BNC117 infusions, separated by 3 or 2 weeks, respectively, are generally well tolerated. Infusions are associated with a delay in viral rebound of 5–9 weeks after two infusions, and up to 19 weeks after four infusions, or an average of 6.7 and 9.9 weeks, respectively, compared with 2.6 weeks for historical controls ($P < 0.00001$). Rebound viruses arise predominantly from a single provirus. In most individuals, emerging viruses show increased resistance, indicating escape. However, 30% of participants remained suppressed until antibody concentrations waned below 20 μg ml^{-1}, and the viruses emerging in all but one of these individuals showed no apparent resistance to 3BCN117, suggesting failure to escape over a period of 9–19 weeks. We conclude that the administration of 3BNC117 exerts strong selective pressure on HIV-1 emerging from latent reservoirs during analytical treatment interruption in humans.

A fraction of HIV-1-infected individuals develops broad and potent serologic activity against the virus. Single-cell antibody cloning methods[2] have uncovered the source of this activity as broadly neutralizing antibodies (bNAbs), which target different sites on the HIV-1 envelope spike protein, gp160 (refs 1–3).

In animal models, bNAbs show potent prophylactic activity, suppress established viraemia, and delay viral rebound during analytical treatment interruption (ATI)[4–8]. In humans, a phase I clinical trial showed that 3BNC117 is generally safe and effective in transiently reducing viraemia in chronically HIV-1-infected individuals[9]. A single infusion of 3BNC117 was well tolerated, rapidly decreased viral loads in viraemic individuals by an average of 1.48 $\log_{10}$ copies per ml, with durable activity for 4 weeks[9]. In addition, 3BNC117 increased autologous antibody responses in HIV-1-infected individuals, and enhanced clearance of infected cells in humans and in humanized mice[10,11]. VRC01, a less potent bNAb that also targets the CD4-binding site, suppressed viraemia by 1.14 $\log_{10}$ (refs 12, 13 and Fig. 1a, b).

To investigate whether 3BNC117 can suppress viral rebound from the latent reservoir during ATI in chronically suppressed HIV-1 infected humans, we conducted a phase IIa open label clinical trial. To select participants with 3BNC117-sensitive viruses in their latent reservoirs, we performed bulk viral outgrowth cultures of peripheral blood mononuclear cells (PBMCs) from individuals whose viraemia was suppressed by combination antiretroviral therapy (ART). The resulting isolates were screened for sensitivity to 3BNC117 using the TZM-bl assay (Supplementary Table 1). Of 63 individuals screened, only 11% yielded viruses that were fully resistant to 3BNC117 ($IC_{50} > 20$ μg/ml), and 65% were sensitive to 3BNC117 IC_{50} at concentrations below 2.0 μg/ml. In contrast only 29% were similarly sensitive to VRC01 (Fig. 1a and b, Extended Data Fig. 1 and Supplementary Table 1).

We enrolled HIV-1 infected individuals who were on suppressive antiretroviral therapy (ART) with plasma viral loads <50 HIV-1 RNA copies per ml for at least 12 months, had CD4 counts >500 cells per mm^3, yielded 3BNC117-sensitive outgrowth viruses ($IC_{50} \leq 2.0$ μg ml^{-1}), and whose viral load at screen was <20 copies per ml (Extended Data Fig. 1, Supplementary Tables 2–4, and Methods). Participants were enrolled in two groups: eight in group A to receive two 30 mg kg^{-1} infusions three weeks apart, while seven in group B received up to four 30 mg kg^{-1} infusions at two-week intervals (Fig. 1c, d, Supplementary Table 2). Two group A participants had viral loads >20 copies per ml at the time of infusion and were excluded from further analysis (Supplementary Tables 2 and 4). Participants are numbered 701–715 (Supplementary Table 2).

ATI was started 2 days after the first 3BNC117 infusion. ART was reinitiated and infusions were stopped after two consecutive plasma viral load measurements exceeded 200 copies per ml. All individuals on non-nucleoside reverse transcriptase inhibitors (NNRTIs) were switched to an integrase-inhibitor-based regimen (dolutegravir plus tenoforvir disoproxil fumarate/emtricitabine) four weeks before ATI owing to the long half-life of NNRTIs (Supplementary Table 2).

Both dosing regimens were generally well tolerated. The majority of reported adverse events were transient and grade 1 in severity (Supplementary Table 5). The mean CD4 T-cell count at baseline (day 0) was 747 cells per mm^3, and the average change in CD4 T-cell counts between start of ATI and rebound was −127 cells per mm^3. Although CD4 T cells declined modestly during viral rebound in some participants, CD4 T-cells returned to baseline by week 12 in most participants (mean 828 cells per mm^3) (Extended Data Fig. 2 and Supplementary Table 4). Of 12 individuals tested, 5 showed measurable increases in the magnitude and/or breadth of T cell responses to HIV-1 12 weeks

[1]Laboratory of Molecular Immunology, The Rockefeller University, New York, New York 10065, USA. [2]Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA. [3]Departments of Medicine and Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [4]Department of Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Campus E1 4, 66123 Saarbrücken, Germany. [5]Division of Infectious Diseases, Weill Medical College of Cornell University, New York, New York 10065, USA. [6]Division of Biology, California Institute of Technology, Pasadena, California 91125, USA. [7]Ragon Institute of MGH, MIT and Harvard, Cambridge, Massachusetts 02139, USA. [8]Celldex Therapeutics, Inc., Hampton, New Jersey 08827, USA. [9]Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York 10467, USA. [10]Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Boston, Massachusetts 02215, USA. [11]Laboratory of Experimental Immunology, Center for Molecular Medicine Cologne (CMMC), University of Cologne, 50931 Cologne, Germany. [12]Department I of Internal Medicine, Center of Integrated Oncology Cologne-Bonn, University Hospital Cologne, 50937 Cologne, Germany. [13]Howard Hughes Medical Institute, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [14]Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10065, USA.
*These authors contributed equally to this work.

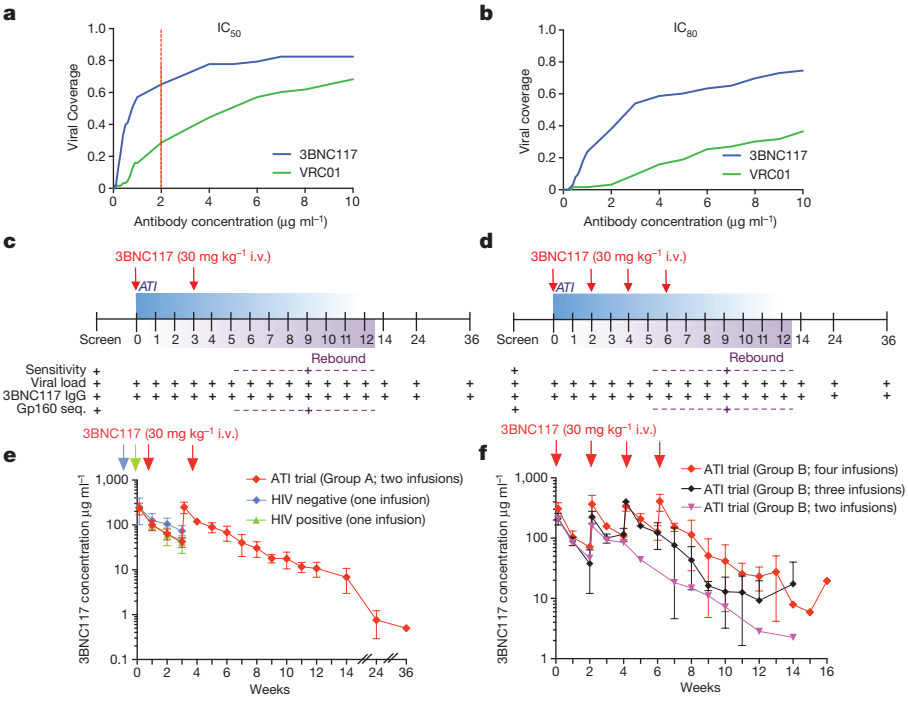


Figure 1 | 3BNC117 neutralization coverage, trial design and pharmacokinetics of 3BNC117 in HIV-1-infected individuals during ATI. **a**, **b**, Sensitivity of virus outgrowth cultures from 63 ART suppressed individuals to 3BNC117 and VRC01 (Supplementary Table 1). The *y*-axis shows the fraction of viral outgrowth culture supernatants neutralized by a given antibody concentration (*x*-axis) in Tzm-bl assays. Red line indicates cut-off IC_{50} (2 μg ml^{-1}) for participation in the trial. **c**, **d**, Diagrammatic representation of study groups A and B respectively. 3BNC117 infusions indicated by the red arrows, and sampling for PK and virologic studies indicated below. Numbers indicate study weeks. **e**, **f**, 3BNC117 levels as determined by ELISA for group A ($n = 6$, left panel, red), group B ($n = 7$, right panel, red ($n = 4$), black ($n = 2$) and purple ($n = 1$)), HIV-1 negative ($n = 3$, blue) and viraemic individuals ($n = 6$, green)[9]. Curves indicate mean 3BNC117 levels, error bars indicate standard deviation. Arrows indicate 3BNC117 infusions.

after ATI, relative to baseline (Extended Data Fig. 3). None of the participants experienced acute retroviral syndrome during rebound, and viraemia was re-suppressed below 20 copies per ml in all participants within 2–7 weeks after restarting ART (Supplementary Table 4). We conclude that up to four 30 mg kg^{-1} infusions of 3BNC117 during ATI are generally safe and well tolerated.

By anti-idiotype ELISA[9] the half-life of 3BNC117 during ATI was 19.6 days among group A participants, and 14.1 days among those in group B (Fig. 1e, f and Supplementary Table 4). These measurements are similar to previously reported values for 3BNC117 in HIV-1-infected individuals on ART[9] (Fig. 1e).

All six group A participants maintained viral loads below 200 copies per ml during the first 4 weeks, with rebound 5–9 weeks after ART interruption (Fig. 2a and Supplementary Table 4a). In group B, rebound occurred 3–19 weeks after ATI, with four out of seven (57%) participants remaining suppressed for at least 10 weeks (Fig. 2b and Supplementary Table 4b). The average time to rebound was 6.7 weeks in group A, 9.9 weeks in group B, and 8.4 weeks for all participants together, compared with 2.6 weeks for matched historical non-infused control individuals (Fig. 2c, Extended Data Fig. 4a, Supplementary Tables 4, 6 and 7). Altogether, 6 of the 13 infused individuals (46%) remained suppressed until at least 9 weeks after ATI. Relative to matched historical control individuals, the delay to rebound among all 3BNC117-infused participants was highly significant ($P < 0.00001$ weighted log-rank test, Fig. 2c, Extended Data Fig. 4, Supplementary Tables 4, 6 and 7 and Methods). We conclude that repeated infusions of 3BNC117 are generally safe and significantly delay HIV-1 rebound from the latent reservoir during ATI.

Time to viral rebound did not correlate with pre-ATI viral culture sensitivity to 3BNC117, nor to baseline levels of cell-associated HIV-1 DNA (Fig. 2d and Extended Data Fig. 4e). Therefore the significance of viral outgrowth sensitivity as an inclusion criterion is not clear. 3BNC117 levels at rebound were also variable, ranging from 6–168 μg ml^{-1}, but directly correlated with the IC_{80} of the emerging virus (Fig. 2a, b, e).

To determine whether rebound was associated with resistance to 3BNC117, we compared pre-infusion and rebound viral outgrowth cultures. A majority (8/13) of participants had rebound viruses that were more resistant to 3BNC117 (IC_{80} > threefold higher, Fig. 3a, c, Extended Data Fig. 5a, Supplementary Table 3). Among group A participants, all but one (707) had more resistant rebound viruses; however, among group B participants, four of seven (710, 711, 712 and 715) showed similar pre-infusion and rebound sensitivity to 3BNC117 (Fig. 3a, c, Extended Data Fig. 5a, Supplementary Table 3). Among these five individuals, 711 was the earliest to rebound at 3 weeks, despite having viruses that were surprisingly sensitive to 3BNC117 as measured by IC_{50} (Fig. 2b, Extended Data Fig. 5a, Supplementary Tables 3 and 4). However, 100% neutralization was not achieved against 711 rebound or pre-infusion viruses, even at high (50 μg ml^{-1}) antibody concentrations, suggesting that 3BNC117 was not fully therapeutic (Extended Data Fig. 5a, Supplementary Table 3). Thus, the only participant in the study to rebound within 3 weeks of ATI may have done so because of pre-existing resistance to 3BNC117 by the dominant virus in the reservoir.

The other four participants that showed no change between pre- and post-infusion culture sensitivity to 3BNC117, 707, 710, 712, and 715 rebounded relatively late at 9, 19, 16 and 11 weeks after ATI, respectively (Figs 2a, b, 3a, c, Extended Data Fig. 5, Supplementary Tables 3 and 4). In all of these individuals rebound was associated with relatively low antibody concentrations ranging from 6–41 μg ml^{-1} (mean 19.7 μg ml^{-1}). This antibody concentration represents 9.6-fold the mean IC_{80} for the rebounding viruses, which is consistent with previous reports on the relationship between suppressive 3BNC117 concentration and neutralization titre in macaques[14] (Fig. 2, Extended Data Fig. 5, Supplementary Tables 3 and 4).

To determine whether viral rebound during ATI was associated with resistance to other bNAbs undergoing clinical testing, we examined sensitivity to 10-1074 (ref. 15), which targets a different and non-overlapping epitope on the HIV-1 trimer (Fig. 3b, d, Extended Data Fig. 5, and Supplementary Table 3). With the exception of 703 and 711, the participants' rebound cultures did not show increased resistance to 10-1074. We conclude that rebound during ATI in the presence of 3BNC117 is infrequently associated with increased resistance to 10-1074.

To characterize viruses emerging from the latent reservoir further, we performed single genome sequencing (SGS) of viral RNA from the plasma and viral outgrowth cultures from eight individuals. Phylogenetic analysis of these sequences indicated that all of these eight trial participants were infected with epidemiologically unrelated viruses (Extended Data Fig. 6). Given the limited sampling of the pre-infusion reservoir, rebound viruses did not always fall within the radiation

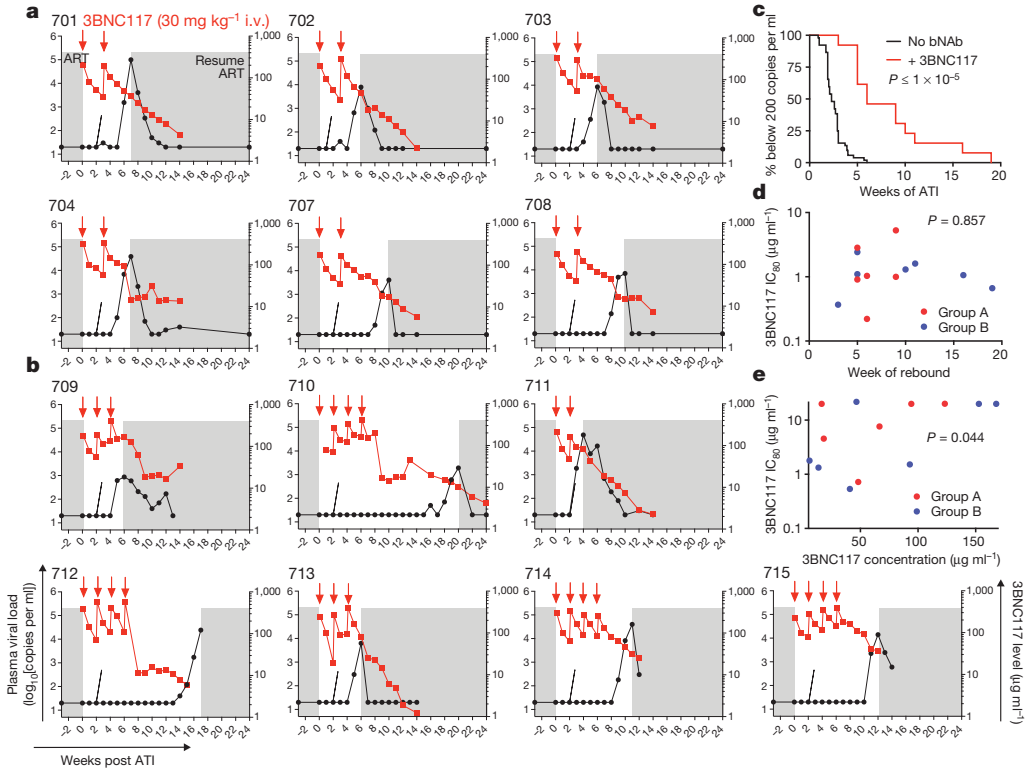


Figure 2 | Delay in viral rebound in the presence of 3BNC117. **a**, **b**, Plasma viral loads and 3BNC117 levels in group A and group B participants respectively. 3BNC117 infusions are indicated with red arrows. The left *y*-axis shows plasma viral loads in RNA copies per ml (black curves), and right *y*-axis shows antibody levels measured by ELISA (red curves). Average rebound time point (2.6 weeks, Supplementary Table 6) in 52 ACTG trial participants who underwent ATI without antibody treatment[29] is shown with dotted lines. Grey areas indicate ART therapy. **c**, Kaplan–Meier plot summarizing viral rebound in 52 ACTG trial participants who underwent ATI without antibody treatment (black, Supplementary Table 6), and the combination of all 13 participants (red) who underwent ATI with 3BNC117 infusions. The *y*-axis indicates the percentage of participants with viral levels below 200 RNA copies per ml, *x*-axis indicates weeks after ATI initiation. The *P* value is based on a bootstrap version of the weighted log-rank test adjusting for the potential confounders 'years on ART', and 'age' (Supplementary Table 7, Methods Statistical Analyses). **d**, Dot plot indicating the relationship between 3BNC117 sensitivity of pre-infusion outgrowth cultures at screening (*y*-axis, IC_{80} in μg ml^{-1}) and the week of rebound (*x*-axis). Group A ($n = 6$) and group B ($n = 7$) participants are coloured red and blue respectively. The *P* value was derived from calculating the Pearson correlation coefficients. **e**, Dot plot indicating the relationship between 3BNC117 sensitivity of rebound outgrowth cultures (*y*-axis, IC_{80} in μg ml^{-1}) and the 3BNC117 serum concentration at rebound (*x*-axis, in μg ml^{-1}). 704, 708, 709 and 713 did not reach IC_{80} at the concentrations tested and were assigned a value of 22 μg ml^{-1}. Group A ($n = 6$) and group B ($n = 7$) participants are coloured red and blue, respectively. The *P* value was derived from calculating the Pearson correlation coefficients.

of pre-infusion viral isolates (Fig. 3e, f, Extended Data Figs 7 and 8, Supplementary Figs 1 and 2).

Remarkably, in five of eight participants, all rebounding virus sequences clustered within a low diversity lineage, consistent with the clonal expansion of a single recrudescent virus (Fig. 3e, f, Extended Data Figs 7 and 8, Supplementary Table 8). These data contrast with individuals undergoing ATI in the absence of antibody infusion, where virus rebound is consistently polyclonal, indicating the activation of multiple latently infected cells[16–19]. Thus, in addition to delaying rebound, 3BNC117 appears to restrict the outgrowth of viral genotypes from the latent reservoir.

Six of eight participants sequenced had rebound viral outgrowth culture and/or plasma sequences that indicated 3BNC117 resistance. For example, in 704, all rebound viruses carried a serine at position 456 (Fig. 4a and Supplementary Figs 1 and 2), which may disrupt a highly conserved salt bridge that maintains the V5 loop's position and conformation[20–22]. Similarly, in 708 and 709, nearly all rebound viruses carried atypical residues at position 282, where a lysine residue typically forms a salt bridge with 3BNC117 (Fig. 4a and Supplementary Figs 1 and 2)[23]. However, documented 3BNC117 resistance mutations[24] were not universally identified among rebound viral strains (Fig. 4a and Supplementary Figs 1 and 2). Only a minor fraction (3 of 23) of sequences in the rebound population of participant 701 had potential resistance-conferring residues in Loop D (274F, 282R), while the remaining sequences did not (Fig. 4a and Supplementary Figs 1 and 2). Similarly, in 702 and 703, only a subset of rebound viruses carried a putative resistance-conferring A281D change[1,23]. Nevertheless, the frequency of this change increased markedly with time in both participants, indicating continued selection for 3BNC117 resistance (Fig. 4a and Supplementary Figs 1 and 2). For participants 707 and 711, no sequence features were identified that would indicate 3BNC117 resistance.

To determine the sensitivity of rebound viruses to 3BNC117, we performed TZM-bl neutralization assays using pseudoviruses typed with SGS Env genotypes (Fig. 4b, Extended Data Figs 7 and 8, Supplementary Figs 1 and 2 and Supplementary Table 9). With the exception of participant 707, who rebounded 9 weeks after ATI at very low 3BNC117 titres, Env genotypes at rebound were more resistant to 3BNC117 than pre-ATI (Fig. 4b, Extended Data Figs 7 and 8 and Supplementary Table 9). We conclude that viral rebound during ATI in the presence of 3BNC117 selects for the emergence of resistant variants, indicating strong selection pressure by this antibody on viral populations arising from the reservoir.

Antibody potency and half-life are directly correlated with HIV-1 prophylaxis in pre-clinical models. For example, VRC01, a CD4bs antibody that is less potent than 3BNC117[1], is less effective than 3BNC117 in preventing $SHIV_{AD8}$ infection in macaques[8,25]. Consistent with these observations, clinical trials with combinations of three less-potent

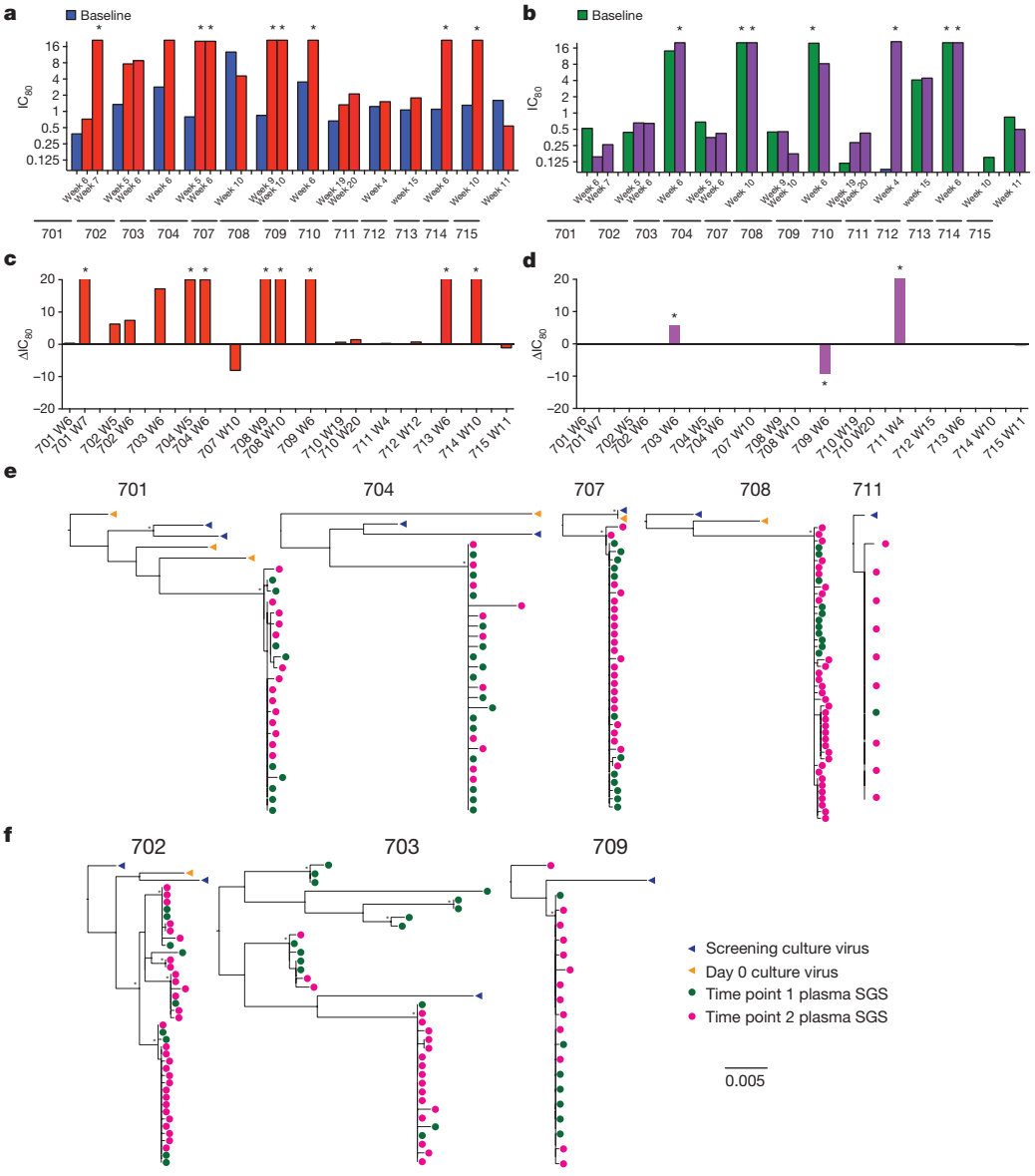


Figure 3 | Viral rebound during ATI and 3BNC117 treatment. **a**, **b**, Graph of 3BNC117 (**a**) or 10-1074 (**b**) IC_{80} titres against baseline and rebound outgrowth cultures. Blue and green bars represent average IC_{80} titres against screen and day 0 outgrowth cultures; red and purple bars represent IC_{80} titres against rebound outgrowth cultures from the indicated time points. Asterisks indicate cultures failing to reach an IC_{80} up to 20 μg ml^{-1}. **c**, **d**, Difference between rebound and pre-infusion culture IC_{80} titres (from **a**, **b**) for 3BNC117 (**c**) or 10-1074 (**d**). For cultures failing to reach an IC_{80} up to 20 μg ml^{-1}, a value of 20 μg ml^{-1} was assigned, and such cultures are marked with an asterisk. **e**, **f**, Clonality of the rebound virus. Maximum likelihood phylogenetic trees comparing pre-ATI single genome derived *env* sequences (blue and orange) to rebound plasma *env* sequences (green and pink) are shown for participants whose rebound comprised single (**e**) versus multiple (**f**) viruses (Supplementary Table 8). Pre-ATI culture sequences were inferred as described in the Methods section 'Statistical analyses'.

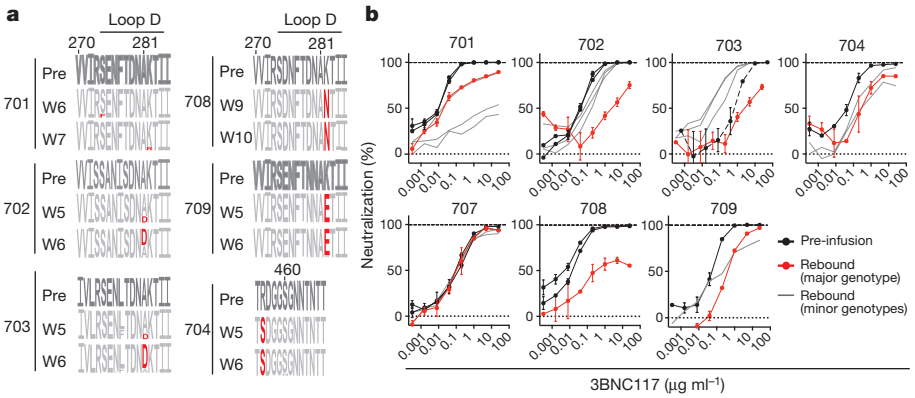


Figure 4 | 3BNC117 resistance in rebound viruses. **a**, Logogram shows *env* gp120 regions (amino acid positions; 270–285 and 455–467, according to HXBc2 numbering) indicating sequence changes from pre-infusion culture(s) (first row) to rebound sequences derived from plasma SGS at the indicated time points. The frequency of each amino acid is indicated by its height. Red residues represent mutations predicted to affect neutralization[30]. **b**, 3BNC117 neutralization sensitivity of pseudoviruses derived from pre-infusion or rebound SGS. Black lines represent pre-infusion virus *envs*; red lines represent the major *env* at rebound for each participant (Extended Data Figs 7 and 8, Supplementary Tables 8 and 9 and Methods); grey lines represent minor rebound *envs* in participants with multiple rebound viruses or variants that evolved after rebound (Extended Data Figs 7 and 8, Supplementary Table 9). Symbols reflect the means of two technical replicates; error bars denote standard deviation.

first-generation bNAbs showed limited effects on viral rebound in the setting of ATI in chronically infected individuals[26,27]. In addition, selective pressure as evidenced by escape mutations was only observed for one of the three antibodies used in the combination, 2G12[26,27]. In contrast, 3BNC117 alone significantly delayed viral rebound with nearly half of all individuals remaining below 200 copies per ml until at least 9 weeks, including four individuals who failed to develop resistance and only rebounded at low antibody concentrations. We speculate that the difference in efficacy between 3BNC117 and less potent bNAbs in the setting of ATI is due to increased potency and/or a longer half-life[1,9,13].

Nevertheless, the majority of the individuals we studied rebounded at high 3BNC117 serum concentrations. A single viral genotype displaying increased resistance to 3BNC117 established rebound in most cases. These viruses represent pre-existing dormant variants that emerged from the latent reservoir. The time to rebound did not correlate with the amount of viral DNA in circulating PBMCs; however, this is a poor measure of the HIV-1 reservoir, since most integrated proviruses in patients on ART are defective[28]. Instead, the delay in viral rebound may represent a measure of the frequency of 3BNC117-resistant variants in the latent reservoir.

Combinations of drugs are needed to maintain HIV-1 suppression in effective ART regimens. Similarly, combinations of antibodies were required to suppress viraemia in humanized mice[6,7]. We speculate that combinations of bNAbs will also be needed to increase the frequency of individuals that remain suppressed by antibody during ATI.

Whether 3BNC117 can also impact the size and composition of the latent reservoir during ATI will require additional studies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 May; accepted 15 June 2016.
Published online 22 June 2016.

1. Scheid, J. F. *et al.* Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333,** 1633–1637 (2011).
2. Scheid, J. F. *et al.* Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458,** 636–640 (2009).
3. Klein, F. *et al.* Antibodies in HIV-1 vaccine development and therapy. *Science* **341,** 1199–1204 (2013).
4. Barouch, D. H. *et al.* Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys. *Nature* **503,** 224–228 (2013).
5. Halper-Stromberg, A. *et al.* Broadly neutralizing antibodies and viral inducers decrease rebound from HIV-1 latent reservoirs in humanized mice. *Cell* **158,** 989–999 (2014).
6. Klein, F. *et al.* HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature* **492,** 118–122 (2012).
7. Horwitz, J. A. *et al.* HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. *Proc. Natl Acad. Sci. USA* **110,** 16538–16543 (2013).
8. Shingai, M. *et al.* 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 (2014).
9. Caskey, M. *et al.* Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature* **522,** 487–491 (2015).
10. Schoofs, T. *et al.* HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1. *Science* **352,** 997–1001 (2016).
11. Lu, C. L. *et al.* Enhanced clearance of HIV-1-infected cells by broadly neutralizing antibodies against HIV-1 *in vivo*. *Science* **352,** 1001–1004 (2016).
12. Kong, R. *et al.* Improving neutralization potency and breadth by combining broadly reactive HIV-1 antibodies targeting major neutralization epitopes. *J. Virol.* **89,** 2659–2671 (2015).
13. Lynch, R. M. *et al.* Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci. Transl. Med.* **7,** 319ra206 (2015).
14. Shingai, M. *et al.* Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. *Nature* **503,** 277–280 (2013).
15. Mouquet, H. *et al.* Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. *Proc. Natl Acad. Sci. USA* **109,** E3268–E3277 (2012).
16. Rothenberger, M. K. *et al.* Large number of rebounding/founder HIV variants emerge from multifocal infection in lymphatic tissues after treatment interruption. *Proc. Natl Acad. Sci. USA* **112,** E1126–E1134 (2015).
17. Kearney, M. F. *et al.* Lack of detectable HIV-1 molecular evolution during suppressive antiretroviral therapy. *PLoS Pathog.* **10,** e1004010 (2014).
18. Kearney, M. F. *et al.* Origin of rebound plasma HIV includes cells with identical proviruses that are transcriptionally active before stopping of antiretroviral therapy. *J. Virol.* **90,** 1369–1376 (2015).
19. Salantes, B. S. B; Bar, Katharine. Potency and Kinetics of Autologous HIV-1 Neutralizing Antibody Responses During ATI. *CROI Conference Abstracts* **Abstract #92** (2016).
20. West, A. P., Jr, Diskin, R., Nussenzweig, M. C. & Bjorkman, P. J. Structural basis for germ-line gene usage of a potent class of antibodies targeting the CD4-binding site of HIV-1 gp120. *Proc. Natl Acad. Sci. USA* **109,** E2083–E2090 (2012).
21. Diskin, R. *et al.* Restricting HIV-1 pathways for escape using rationally designed anti-HIV-1 antibodies. *J. Exp. Med.* **210,** 1235–1249 (2013).
22. Lynch, R. M. *et al.* HIV-1 fitness cost associated with escape from the VRC01 class of CD4 binding site neutralizing antibodies. *J. Virol.* **89,** 4201–4213 (2015).
23. Zhou, T. *et al.* Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity* **39,** 245–258 (2013).
24. Lyumkis, D. *et al.* Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* **342,** 1484–1490 (2013).
25. Gautam, R. *et al.* A single injection of anti-HIV-1 antibodies protects against repeated SHIV challenges. *Nature* **533,** 105–109 (2016).
26. Trkola, A. *et al.* Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nat. Med.* **11,** 615–622 (2005).
27. Mehandru, S. *et al.* Adjunctive passive immunotherapy in human immunodeficiency virus type 1-infected individuals treated with antiviral therapy during acute and early infection. *J. Virol.* **81,** 11016–11031 (2007).
28. Ho, Y. C. *et al.* Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* **155,** 540–551 (2013).
29. Li, J. Z. *et al.* The size of the expressed HIV reservoir predicts timing of viral rebound after treatment interruption. *AIDS* **30,** 343–353 (2016).
30. West, A. P., Jr *et al.* Computational analysis of anti-HIV-1 antibody neutralization panel data to identify potential functional epitope residues. *Proc. Natl Acad. Sci. USA* **110,** 10598–10603 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank the trial participants for their invaluable support. We thank the Rockefeller University Hospital Clinical Research Support Office and nursing staff for help with recruitment and study implementation, especially N. Buckley, A. Hurley, S. B. A. Shulman and L. Corregano; all members of the Nussenzweig laboratory, especially T. Schoofs, A. Halper-Stromberg, M. and Z. Jankovic; C. Unson-O'Brien, J. Dizon, R. Baptiste and R. Levin for sample processing and study coordination; A. Louie for regulatory support; P. Fast and H. Park for clinical monitoring; E. Giorgi and W. Fischer from Los Alamos National Laboratory; R. T. Gandhi, J. Li and The AIDS Clinical Trials Group (grant UM1 AI068636) and its Statistical and Data Management Center (grant UM1 AI068634). This study was supported by the following grants: Collaboration for AIDS Vaccine Discovery grant OPP1033115 (M.C.N.) and OPP1032144 (M.S.S.). Grant 8 UL1 TR000043 from the National Center for Advancing Translational Sciences (NCATS); NIH Clinical and Translational Science Award (CTSA) program; NIH Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID) 1UM1 AI100663-01 (M.C.N.) and 5UM1 AI100645-03 (B.H.H.); Bill and Melinda Gates Foundation grants OPP1092074 and OPP1124068 (M.C.N.); NIH HIVRAD P01 AI100148 (P.J.B. and M.C.N.); the Robertson Foundation to M.C.N. M.C.N. is a Howard Hughes Medical Institute Investigator. Ruth L. Kirschstein National Research Service Award F30 AI112426 (E.F.K.); F31 AI118555 (J.A.H.); The NIH Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID) 1UM1 AI00645 (B.H.H.); The University of Pennsylvania Center for AIDS Research (CFAR) Single Genome Amplification Service Center P30 AI045008 (B.H.H.); The NIH Scripps Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery (CHAVI-ID and 1UM1-AI100663) (B.D.W.).

Author Contributions M.C.N., J.F.S., J.A.H. and M.C. wrote the manuscript; J.F.S., M.C. and M.C.N. designed the trial; J.F.S., J.A.H., Y.B., J.C.C.L., L.N., Y.Z.C., C.-L.L. and M.B. performed tissue culture experiments and SGS amplifications; M.S.S. performed TZM-bl assays; J.F.S., J.A.H., Y.B., E.F.K., T.O., A.P.W., G.H.L., P.J.B., F.K., S.J.S., B.H.H., M.C.N. and M.C. analysed the data; E.F.K., G.H.L. and B.H.H. performed SGA analysis; I.S., R.P. and J.F.S. processed patient samples; L.B., S.H., A.S., M.W.-P., B.Z., R.M.G., S.J.S. and M.C. performed patient recruitment; A.F. and N.P. performed statistical analyses; B.J. and B.D.W. performed antigen-specific T cell experiments; T.K. and T.H. produced 3BNC117 and provided PK data.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.C.N. (nussen@rockefeller.edu) or M.C. (mcaskey@rockefeller.edu).

Reviewer Information *Nature* thanks S. Deeks, D. Richman and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Study design. An open-label, dose-escalation phase 2a study was conducted in HIV-1-infected participants (http://www.clinicaltrials.gov; NCT02446847). Study participants were enrolled sequentially according to eligibility criteria. Group A received 3BNC117 on days 0 and 21 at a dose of 30 mg/kg body weight at a rate of 250 ml/hour. Group B received 3BNC117 on days 0, 14, 28 and 42 at a dose of 30 mg/kg, as long as viral rebound did not occur. Antiretroviral therapy (ART) was discontinued 2 days after the first 3BNC117 infusion (day 2). Plasma HIV-1 RNA levels were monitored weekly, and ART was resumed when viral load increased to ≥200 c.p.m. in two consecutive weekly measurements.

Study participants were followed for 36 weeks after the first infusion. Safety data are reported until week 36 for participants enrolled in group A and until week 14 for participants enrolled in group B. All participants provided written informed consent before participation in the study and the study was conducted in accordance with Good Clinical Practice. The protocol was approved by the Federal Drug Administration in the USA and the Institutional Review Board at the Rockefeller University.

Study participants. All study participants were recruited at the Rockefeller University Hospital, New York, USA. Eligible participants were adults aged 18–65 years, HIV-1-infected and before enrolment had plasma HIV-1 RNA levels <50 c.p.m. for at least 12 months while on combination ART and <20 c.p.m. at the screening visit, and current CD4 count >500/μl. In addition, participant-derived HIV-1 isolates produced by co-culture of participant PBMCs with HIV-uninfected donor PBMCs were required to be neutralized by 3BNC117 with an IC_{50} <2 μg/ml in TZM-bl neutralization assays, as previously described[31]. An IC_{50} of <2 ug/ml was chosen as a cut-off based on previous PK data of 3BNC117 in humans[9] and data in macaques showing that antibody levels 10–100 times the IC_{50} value against infecting viral strains are necessary to control viral rebound[14]. However, given the limited diversity and representation of the latent reservoir in outgrowth cultures (Supplementary Fig. 2) and the fact that no correlation between pre-infusion IC_{50} and delay of viral rebound was found in this study, the significance of this criterion is unclear. Participants on an NNRTI-based ART regimen were switched to a study-provided integrase-inhibitor-based regimen (dolutegravir (Tivicay, ViiV Pharmaceuticals) + tenofovir disoproxil fumarate/emtricitabine (Truvada, Gilead Sciences) 4 weeks before treatment interruption due to the prolonged half-life of NNRTIs. Exclusion criteria included history of CD4 nadir <200 cells/μl, concomitant hepatitis B or C infections, previous receipt of a monoclonal antibody of any kind, or clinically relevant physical findings, medical conditions or laboratory abnormalities. Pregnant and breastfeeding women were not eligible.

Historical controls (ACTG trial participants). Viral rebound data from 52 participants who participated in four ACTG ATI studies without additional interventions (ACTG 371[32], A5024[33], A5068[34], and A5197[32]) were compared with viral rebound data in this study. Historical controls were selected based on similar inclusion criteria: age 18–65, Plasma HIV-1 RNA <50 c.p.m. for at least 12 months before ATI while on combination ART, CD4 count at time of ATI >500 cells/μl, CD4 nadir >200 cells/μl, weekly viral load measurements at least until viral rebound occurred.

Study procedures. The appropriate volume of 3BNC117 was calculated according to body weight, diluted in sterile normal saline to a total volume of 250 ml, and administered intravenously over 60 min. Study participants received 3BNC117 on days 0 and 21, or 0, 14, 28, and 42 and remained under monitoring at the Rockefeller University Hospital for 4 h after each infusion. Participants returned for frequent follow up visits for safety assessments, which included physical examination, measurement of clinical laboratory parameters such as haematology, chemistries, urinalysis, and pregnancy tests (for women). Plasma HIV-1 RNA levels were monitored weekly during the ATI period and CD4 counts were measured every other week (Supplementary Table 4). Study investigators evaluated and graded adverse events according to the DAIDS AE Grading Table and determined causality. Blood samples were collected before and at multiple times after 3BNC117 infusions. Samples were processed within 4 h of collection, and serum and plasma samples were stored at −80 °C. PBMCs were isolated by density gradient centrifugation. The absolute number of peripheral blood mononuclear cells was determined by an automated cell counter (Vi-Cell XR; Beckman Coulter), and cells were cryopreserved in fetal bovine serum plus 10% DMSO.

ART re-initiation criteria. Antiretroviral therapy was discontinued 2 days after the first 3BNC117 infusion (day 2). ART was re-initiated when HIV-1 RNA levels were found to be ≥200 c.p.m. and/or CD4 T cell counts decreased to <350 cells/μl and the result was confirmed with a repeat measurement.

Plasma HIV-1 RNA Levels. HIV-1 RNA levels in plasma were measured at the time of screening (within 49 days before the first infusion), day 0 (before infusion), and weekly until week 12, then at weeks 14, 24 and 36. Participants that remained virologically suppressed to <20 c.p.m. off ART beyond week 12, returned for weekly measurements of plasma HIV-1 RNA levels. HIV-1 RNA levels were determined using the Roche COBAS AmpliPrep/COBAS TaqMan HIV-1 Assay, Version 2.0, which detects between 20 and 1×10^7 c.p.m. This assay was performed at LabCorp.

$CD4^+$ and $CD8^+$ T cells. $CD4^+$ and $CD8^+$ T-cell counts were determined at screening, on day 0 (before infusion), and weeks 2, 3, 4, 6, 8, 10, 12, 14, and 36 by a clinical flow cytometry assay, performed at LabCorp. Cells were analysed by flow cytometry. Leukocytes were determined as $CD45^+$ cells. Percentage of cells positively stained for CD3, CD4, CD8 as well as the CD4/CD8 ratio were analysed with the BD Multiset software (BD Biosciences).

3BNC117 study drug. 3BNC117 is a recombinant, fully human IgG1κ mAb recognizing the CD4 binding site on the HIV-1 envelope[1]. The antibody was cloned from an HIV-1-infected viraemic controller in the International HIV Controller Study[1,35], expressed in Chinese hamster ovary cells (clone 5D5-5C10), and purified using standard methods. The 3BNC117 drug substance was produced at Celldex Therapeutics Fall River (MA) GMP facility, and the drug product was fill-finished at Gallus BioPharmaceuticals (NJ). The resulting purified 3BNC117 was supplied as a single use sterile 20 mg/ml solution for intravenous injection in 8.06 mM sodium phosphate, 1.47 mM potassium phosphate, 136.9 mM sodium chloride, 2.68 mM potassium chloride, and 0.01% polysorbate 80. 3BNC117 vials were shipped and stored at 4 °C.

Measurement of 3BNC117 serum levels. Serum levels of 3BNC117 were determined by a validated sandwich ELISA at Celldex Theapeutics as described previously[9]. Plates (Sigma-Aldrich PN: CLS3590 96-well, High Bind, polystyrene) were coated with 4 μg/ml of an anti-idiotypic antibody specifically recognizing 3BNC117 (anti-ID 1F1 mAb), and incubated overnight at 2–8 °C. After washing, plates were blocked for 1 h with 5% BSA. Serum samples, QCs and standards were added (1:50 minimum dilution in 5% BSA) and incubated for 1 h at room temperature. 3BNC117 was detected using an HRP-conjugated mouse anti-human IgG kappa chain specific antibody (Abcam PN: ab79115) and the HRP substrate tetra-methylbenzidine. 3BNC117 concentrations were then interpolated from a standard curve of 3BNC117 using a four-parameter logistic curve-fitting algorithm. The reference standard and positive controls were created from the drug product lot of 3BNC117 used in the clinical study.

Pharmacokinetic analysis. Blood samples were collected immediately before and at the end of infusions as well as on the day after infusion, weekly during the ATI period and at weeks 14, 24 and 36. 3BNC117 serum levels were measured by ELISA (Celldex Therapeutics).

Neutralization assay. Serum samples, viral supernatants, and control antibodies were tested against HIV-1 envelope pseudoviruses as previously described[36,37].

Cell-associated HIV-1 DNA. Participant's $CD4^+$ T-cells were isolated from 10 million cryopreserved PBMCs by negative magnetic selection (Miltenyi). Total DNA was extracted and quantitative PCR performed using *pol*- and CCR5-directed primers as previously described[7].

Virus cultures. Autologous virus was retrieved from HIV-1 infected individuals as previously described[31]. Briefly, healthy donor PBMCs were obtained by leukapheresis from a single donor. Cells were cultured at a concentration of 5×10^6/ml in Iscove's Modified Dulbecco's Medium (IMDM; Gibco) supplemented with 10% fetal bovine serum (FBS; HyClone, Thermo Scientific), 1% penicillin/streptomycin (Gibco), and 1 μg/ml phytohemagglutinin (PHA; Life Technologies) at 37 °C and 5% CO_2. After 2–3 days, 5×10^6 $CD8^+$ depleted cells were transferred into IMDM supplemented with 10% FBS, 1% penicillin/streptomycin, 5 μg/ml polybrene (Sigma), and 100 U/ml of IL-2. Cells were then co-incubated with $4–8 \times 10^6$ $CD4^+$ T cells from the study participants and 10 million irradiated healthy donor PBMCs that had been cultured together for 24 h prior in IMDM supplemented with 10% FBS, 1% penicillin/streptomycin, 100 U/ml IL-2 and 1 μg/ml PHA at 37 °C and 5% CO_2. Lymphoblasts were replenished weekly by adding 3 million healthy donor PHA stimulated $CD8^+$ depleted lymphoblasts. Culture supernatants were quantified using the Alliance HIV-1 p24 Antigen ELISA kit (PerkinElmer) according to the manufacturer's instructions. $TCID_{50}$s were determined for all HIV-1 containing supernatants[36,37] and then tested for sensitivity against 3BNC117 and other bNAbs in a TZM.bl neutralization assay. Blood samples and leukapheresis were collected under separate IRB-approved protocols and after volunteers provided informed consent.

Sequence analysis. HIV-1 RNA extraction and single genome amplification was performed as described previously[38]. In detail, HIV-1 RNA was extracted from plasma samples using the Qiagen MinElute Virus Spin kit (Qiagen) followed by first strand cDNA synthesis using SuperScript III reverse

transcriptase (Invitrogen Life Technologies) and the antisense primer env3out 5′-TTGCTACTTGTGATTGCTCCATGT-3′. gp160 *env* was amplified using envB5out 5′-TAGAGCCCTGGAAGCATCCAGGAAG-3′ and envB3out 5′-TTGCTACTTGTGATTGCTCCATGT-3′ in the first round and second round nested primers envB5in 5′-CACCTTAGGCATCTCCTATGGCAGGAAGAAG-3′ and envB3in 5′- GTCTCGAGATACTGCTCCCACCC-3′. PCRs were performed using a High Fidelity Platinum Taq (Invitrogen) at 94 °C, 2 min; (94 °C, 15 s; 55 °C 30 s; 68 °C, 4 min) × 35; 68 °C, 15 min. Second round PCR was performed with 1 μl of first PCR product as template and High Fidelity Platinum Taq at 94 °C, 2 min; (94 °C, 15 s; 55 °C 30 s; 68 °C, 4 min) × 35; 68 °C, 15 min. Sequence alignments, phylogenetic trees and mutation analysis of gp160 was performed by using Geneious Pro software, version 8.1.6 (Biomatters Ltd.)[39]. Sequence analysis was performed using Antibody database by A. West[30]. Logograms were generated using the Weblogo 3.0 tool[40].

Pseudovirus generation. Selected SGS from virus culture supernatants and plasma were used to generate pseudoviruses and tested for sensitivity to bNAbs in a TZM.bl assay[41]. To produce the pseudoviruses, plasmid DNA containing the cytomegalovirus (CMV) promoter was amplified by PCR using forward primer 5′-GTTGACATTGATTATTGACTAG and reverse primer 5′-CTTCCTGCCAT AGGAGATGCCTAAAGCTCTGCTTATATAGAC-CTC. The CMV promoter amplicon was fused to individual *env* SGS amplicons by PCR using forward primer 5′-AGTAATCAATTACGGGGTCATTAGTTCAT and reverse primer 5′-ACTTTTTGACCACTTGCCACCCAT. Fusion PCR was carried out using the Expand Long Template PCR System (Roche) in a 60 μl reaction consisting of 1 ng purified CMV promoter amplicon, 0.125 μl unpurified *env* SGA amplicon, 200 nM forward and reverse primers, 200 μM dNTP mix, 1× Buffer 1, and 1 μl DNA polymerase mix. PCR was run at 94 °C for 2 min; 25 cycles (94 °C for 12 s, 55 °C for 30 s, 68 °C for four minutes); and 72 °C for 10 min. Resulting amplicons were analysed by gel electrophoresis, purified without gel extraction, and co-transfected with pSG3Δ*env* into HEK293T cells to produce pseudoviruses as described previously[41].

Statistical analyses. Adverse events were summarized by the number of participants who experienced the event, by severity grade according to the DAIDS AE Grading Table and by relationship to 3BNC117 as determined by the investigator. PK-parameters were estimated by performing a non-compartmental analysis (NCA) using WinNonlin 6.3. Kaplan–Meier survival curves were used to compare time to rebound in trial participants to participants in previous ATI studies conducted by ACTG[29]. To exclude the possibility that the observed delay in rebound is confounded by clinical factors, we compared the clinical variables between the control (ACTG trial participants) and treated group using a two-sided Fisher's Exact test for categorical variables (gender and CD4 Nadir) and an unpaired Wilcoxon test (two-sided) for continuous variables (age, years on ART and CD4 count before ATI initiation) (Supplementary Table 7). Additionally, we tested for each potential confounder whether the variable is predictive for the rebound time. Therefore, we built a univariate survival regression model for each potential confounder and compared those models to a null model using a likelihood ratio test (LRT), which determines how much better the more complex model explains the data than the less complex model. Confounders were considered significant if the model with the potential confounder had an LRT *P* value of 0.05 or less, which was the case for 'years on ART' as well as 'age' for the comparison between the controls and the combined treatment group (Supplementary Table 7). We did not perform a standard Cox regression, since the proportional hazards assumption was not fulfilled for some of the variables. Rebound time was modelled using a log-normal distribution, which resulted in the best model fit as measured by Akaike information criterion (AIC) among several different distributions (Extended Data Fig. 4b–d). To determine the effect of the treatment after adjusting for the discovered confounders, we performed a weighted log-rank test[42]. Therefore, for each sample inverse probability weights based on the discovered confounders were estimated, which were used to re-weigh the variables of the log-rank statistic. We performed a bootstrapped version of the weighted log-rank test, as recommended in ref. 42 owing to the small sample size. We estimated the class probabilities using a lasso logistic regression model trained with the Matlab function lassoglm with five lambda values in a threefold cross-validation. To improve stability, the optimal lambda for the lasso logistic regression was determined only once using the original labels and used in all bootstrap runs to train the models that estimate the class probabilities.

Additionally, we performed an LRT at significance level $\alpha = 0.05$ based on a parametric survival regression model adjusted for the discovered confounders. In this analysis the treatment group still significantly predicted the delay in rebound (Supplementary Table 7). For the analyses, the R (version 3.2.1) packages survival (version 2.38-3) and fitdistrplus (version 1.0-6) were used and Matlab (version R2015b) for implementation of the weighted log-rank test.

Sequence and phylogenetic analysis. Nucleotide alignments were generated using ClustalW (v.2.11)[43] and manually adjusted using Geneious R8 (v.8.1.6)[39] and MacClade (v.4.08a)[44]. Sites that could not be unambiguously aligned were removed for all phylogenetic analyses. Optimal evolutionary model classes were determined using jModelTest (v.2.1.4)[45]. Maximum likelihood phylogenetic trees were generated using PhyML (v.3)[46] with joint estimation of evolutionary model parameter values and phylogenies. The tree comparing all participants was midpoint rooted and each within-subject tree was rooted on the basal branch as determined by the between-subject tree. Sequences with premature stop codons and frameshift mutations that fell in the gp120 surface glycoprotein region were excluded from all deduced protein analyses.

Sequences generated from the supernatants of viral outgrowth assays represented viruses that were present in the latent reservoir. Per assay, 4–8 million $CD4^+$ T cells were activated. In an HIV-infected person who is completely suppressed on antiretroviral therapy, it has been determined that 1×10^{-6} resting CD4 cells are latently infected with replication-competent virus[47]. Thus, one would expect to identify up to eight distinct viral isolates per individual culture. Single genome sequencing of the culture supernatants revealed sets of clonally related sequences, which appear as 'rakes' in a phylogenetic tree (Extended Data Figs 7 and 8). The most recent common ancestor of these rakes represents the reactivated virus that was present in the host (similar to the inference of infectious molecular clones as described in ref. 48). As shown in Extended Data Figs 7 and 8, sequences from culture reactions fall in 1–3 rakes within a given individual. We inferred each rake's most recent common ancestor (MRCA) by building a majority-rule consensus and treated it as a single virus from the participant's latent pool. These MRCAs were used to build the phylogenetic trees shown in Fig. 3.

Because mixed culture isolates replicated for 14 or more days, *in vitro* recombinants were observed. *In vitro* recombinants from culture reactions were identified and removed from the data set if they: (i) had two identifiable parental sequences within the same culture reaction; and (ii) exhibited three consecutive informative sites relative to one parent followed by three consecutive informative sites relative to another. We independently verified that a subset of these sequences showed evidence of recombination using the Recco tool (v.0.93)[49].

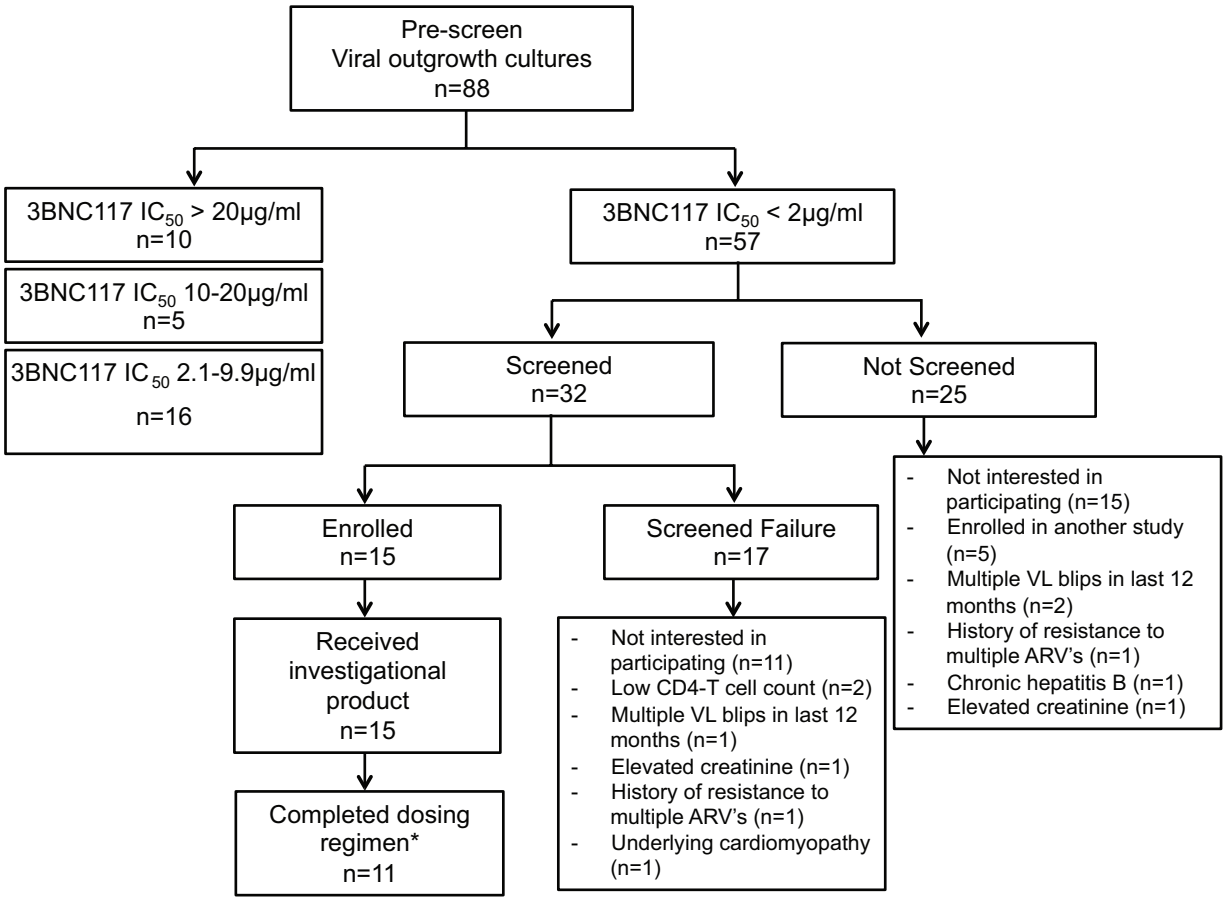
Assessment of rebound virus clonality. The Poisson Fitter v2 tool[50] is designed to test if a set of homogeneous sequences exhibits random diversification. If such a set of sequences exhibits a star-like phylogeny and a Poisson distribution of pairwise differences (Hamming distances), it can be inferred that a single virus gave rise to those sequences. Poisson Fitter v2 tests these and other parameters using maximum likelihood methods and performs a χ^2 goodness of fit test to obtain a *P* value. A non-significant *P* value signifies that the observed Hamming distances adhere to a Poisson distribution and it can be inferred that a single virus gave rise to rebound. Single genome derived *env* sequences from the plasma at the earliest time point post-rebound from each participant were tested using Poisson Fitter (Supplementary Table 8).

ELISPOT T-cell response analysis. Interferon gamma Elispot assays were performed as described[51]. Briefly, 96-well polyvinylidene plates (Millipore, Bedford, Mass.) were precoated with 0.5 g/ml of anti-IFNγ monoclonal antibody, 1-DIK (Mabtech, Stockholm, Sweden) and previously frozen PBMCs were plated at a concentration of 50,000 to 100,000 cells per well in a volume of 100 μl of R10 medium (RPMI 1640 (Sigma), 10% fetal calf serum (Sigma), 10 mM HEPES buffer (Sigma)) with antibiotics (2 mM L-glutamine, 50 U of penicillin-streptomycin/ml). Plates were incubated overnight at 37 °C, 5% CO_2, and developed as described[51]. Cells were tested against a panel of 410 B-clade overlapping 18-mer peptides (OLPs) spanning the entire HIV-1 genome (consensus sequence from 2001). These peptides were used in a matrix system of 11–12 peptides per pool to screen study participants for HIV-specific T cell responses. Confirmation of recognized individual peptides within a peptide pool was undertaken in an additional Elispot assay, as described[52]. Wells containing PBMCs and R10 medium alone were used as negative controls and were run in duplicate on each plate. Wells containing PBMCs and phytohemagglutinin (PHA) served as positive controls. The numbers of spots per well were counted using an automated Elispot plate reader (ImmunoSpot Reader System; Cellular Technology Limited, Shaker Heights, OH). Responses were regarded as positive if they had at least three times the mean number of spot forming cells (SFC) in the two negative control wells and had to be >50 SFC/10^6 PBMCs (ref. 51,52).

31. Laird, G. M. *et al.* Rapid quantification of the latent reservoir for HIV-1 using a viral outgrowth assay. *PLoS Pathog.* **9,** e1003398 (2013).
32. Volberding, P. *et al.* Antiretroviral therapy in acute and recent HIV infection: a prospective multicenter stratified trial of intentionally interrupted treatment. *AIDS* **23,** 1987–1995 (2009).
33. Kilby, J. M. *et al.* A randomized, partially blinded phase 2 trial of antiretroviral therapy, HIV-specific immunizations, and interleukin-2 cycles to promote efficient control of viral replication (ACTG A5024). *J. Infect. Dis.* **194,** 1672–1676 (2006).

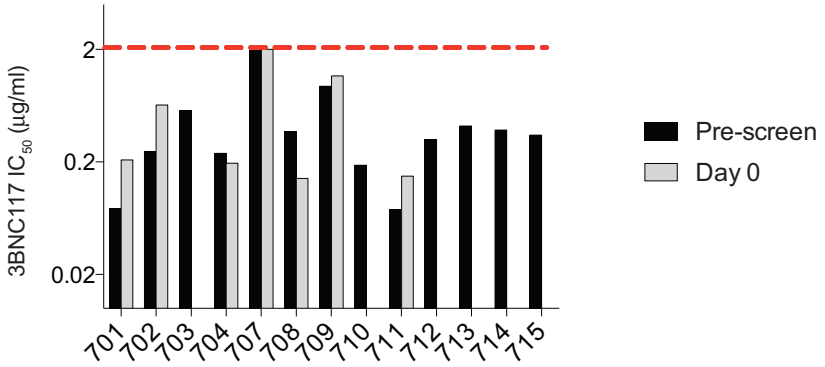
34. Jacobson, J. M. *et al.* Evidence that intermittent structured treatment interruption, but not immunization with ALVAC-HIV vCP1452, promotes host control of HIV replication: the results of AIDS Clinical Trials Group 5068. *J. Infect. Dis.* **194,** 623–632 (2006).
35. Pereyra, F. *et al.* The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science* **330,** 1551–1557 (2010).
36. Montefiori, D. C. Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. *Curr. Protoc. Immunol.* **Chapter 12**, Unit 12 11 (2005).
37. Li, M. *et al.* Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J. Virol.* **79,** 10108–10125 (2005).
38. Salazar-Gonzalez, J. F. *et al.* Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J. Virol.* **82,** 3952–3970 (2008).
39. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28,** 1647–1649 (2012).
40. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14,** 1188–1190 (2004).
41. Kirchherr, J. L. *et al.* High throughput functional analysis of HIV-1 env genes without cloning. *J. Virol. Methods* **143,** 104–111 (2007).
42. Xie, J. & Liu, C. Adjusted Kaplan-Meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Stat. Med.* **24,** 3089–3110 (2005).
43. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948 (2007).
44. Maddison, W. P. & Maddison, D. R. *MacClade – Analysis of Phylogeny and Character Evolution – Version 4.* (Sinauer Associates, Inc., 2001).
45. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9,** 772 (2012).
46. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59,** 307–321 (2010).
47. Chun, T. W. *et al.* Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* **387,** 183–188 (1997).
48. Parrish, N. F. *et al.* Phenotypic properties of transmitted founder HIV-1. *Proc. Natl Acad. Sci. USA* **110,** 6626–6633 (2013).
49. Maydt, J. & Lengauer, T. Recco: recombination analysis using cost optimization. *Bioinformatics* **22,** 1064–1071 (2006).
50. Giorgi, E. E. & Bhattacharya, T. A note on two-sample tests for comparing intra-individual genetic sequence diversity between populations. *Biometrics* **68,** 1323–1326, author reply 1326 (2012).
51. Altfeld, M. A. *et al.* Identification of dominant optimal HLA-B60- and HLA-B61-restricted cytotoxic T-lymphocyte (CTL) epitopes: rapid characterization of CTL responses by enzyme-linked immunospot assay. *J. Virol.* **74,** 8541–8549 (2000).
52. Addo, M. M. *et al.* Comprehensive epitope analysis of human immunodeficiency virus type 1 (HIV-1)-specific T-cell responses directed against the entire expressed HIV-1 genome demonstrate broadly directed responses, but no correlation to viral load. *J. Virol.* **77,** 2081–2092 (2003).

a

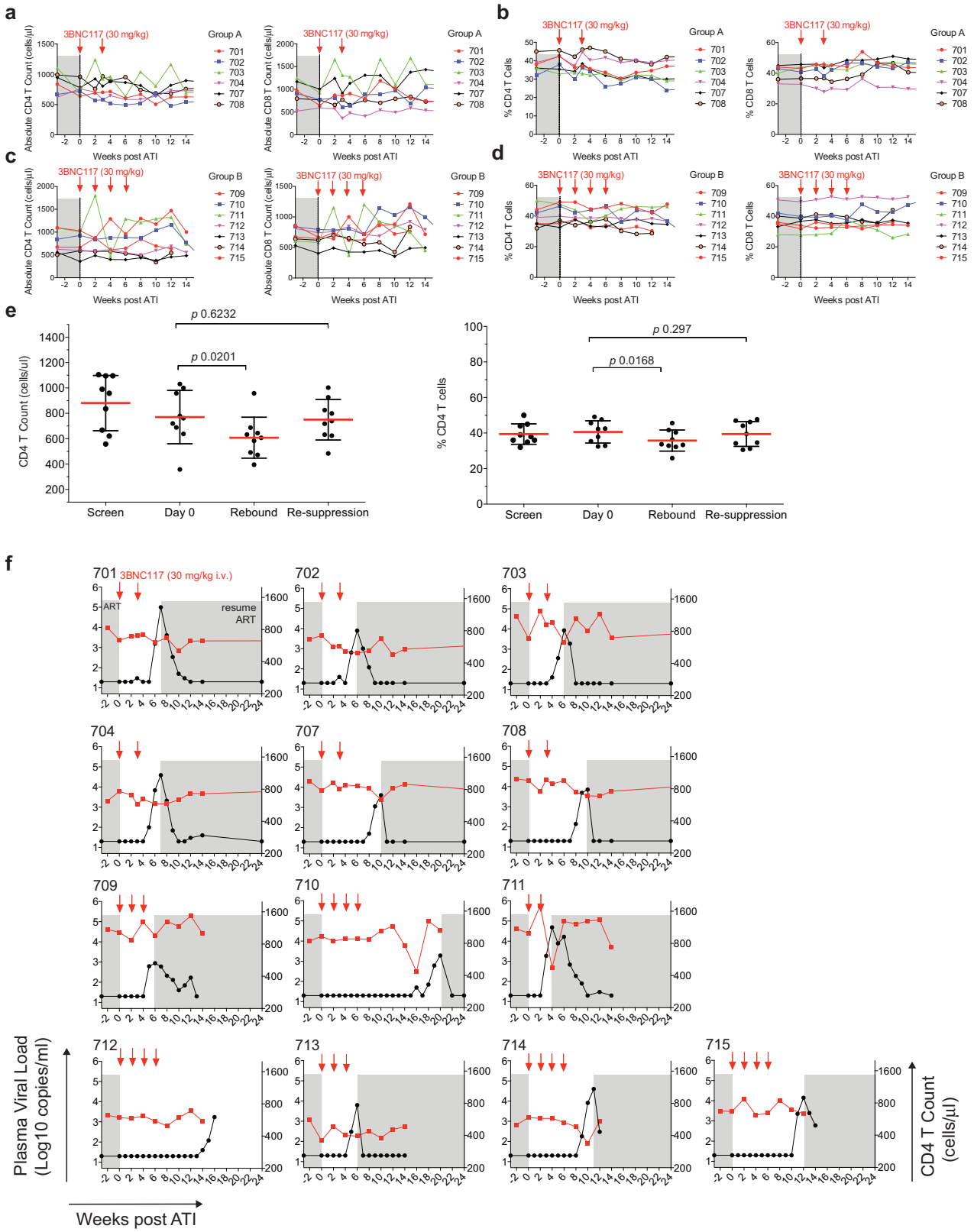


*Participant 705 received only 1 dose due to VL > 1,000 at day 0. Some participants experienced viral rebound prior to completion of all 4 scheduled infusions: 709 – received 3 doses; 711 - 2 doses; 713 - 3 doses.

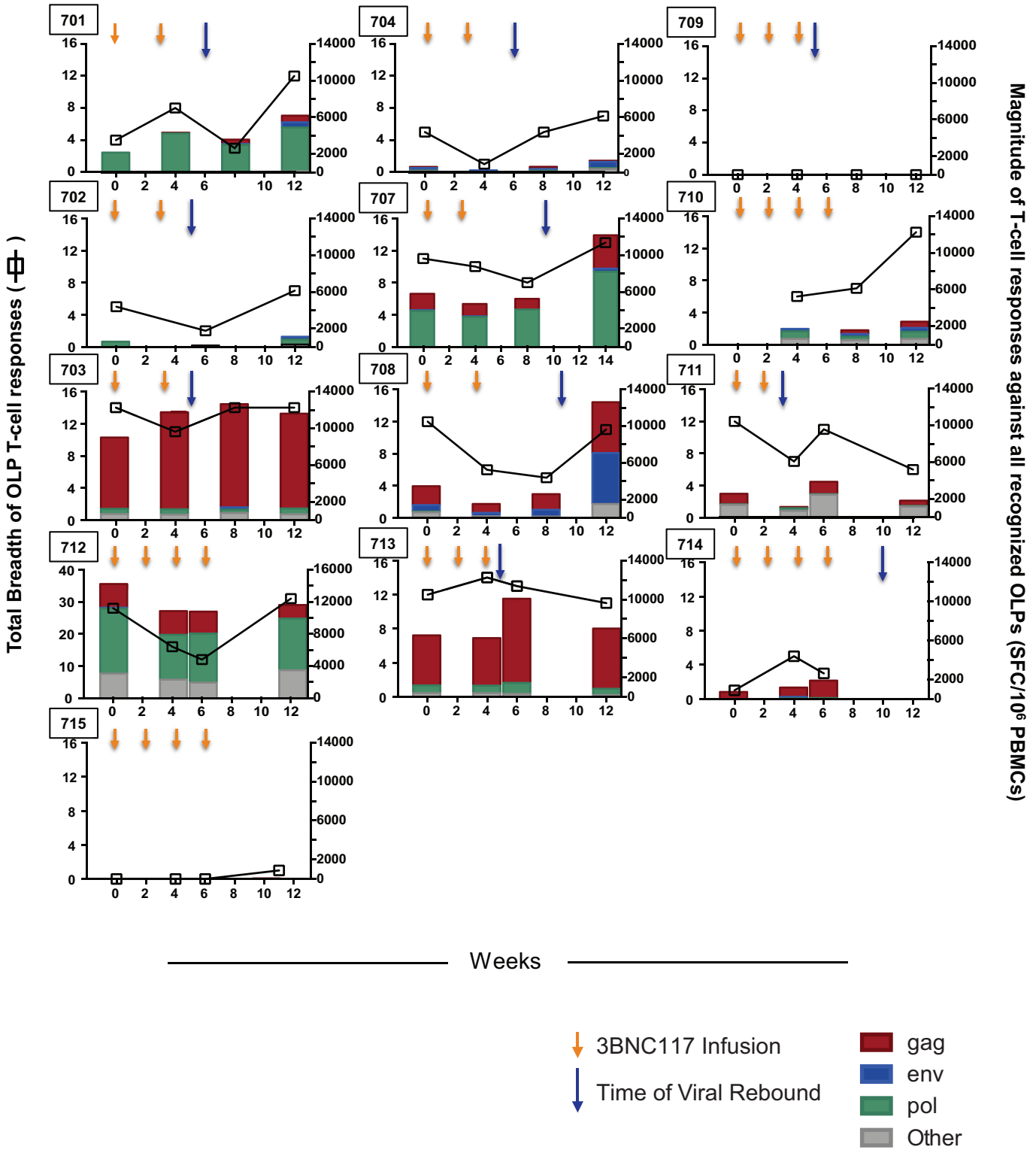
b



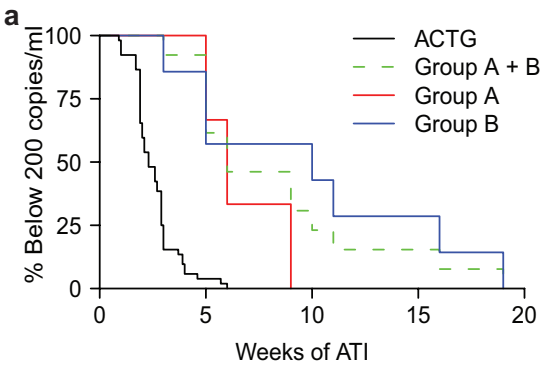
Extended Data Figure 1 | Study participant selection and neutralization of pre-infusion cultures by 3BNC117. a, Flow diagram showing the selection of study participants. **b**, Bar diagrams showing IC_{50} values (µg ml^{-1}) in TZM-bl assays for 3BNC117 against bulk virus outgrowth culture supernatants from the indicated time point pre-infusion for each participant (Supplementary Table 3). For some participants both screen and day 0 cultures were obtained and showed less than threefold variation in IC_{50} values. The red dotted line indicates an IC_{50} of 2 µg ml^{-1} which was used as a threshold for inclusion in the study.

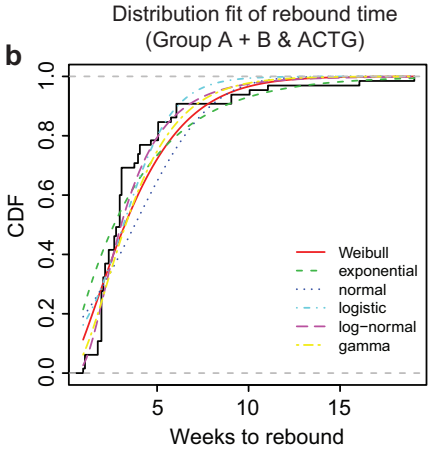


Extended Data Figure 2 | CD4$^+$ and CD8$^+$ T cells during study period in participants. a–d, Absolute T cell counts (**a**, **c**) and percentage of CD4$^+$ and CD8$^+$ T cells among CD3$^+$ T cells (**b**, **d**) for group A and B, respectively (Supplementary Table 4). 3BNC117 infusions are indicated with red arrows. **e**, Comparison of absolute CD4$^+$ T cell counts and percentage of CD4$^+$ T cells among CD3$^+$ T cells at screen, day 0, rebound and after re-suppression. Shown is the data from participants 701, 702, 703, 704, 707, 708, 709, 711 and 713 for whom re-suppression CD4 counts were available (Supplementary Table 4). The last available time point was used as re-suppression time point. Red lines indicate the mean value and error bars indicate standard deviation. *P* values were obtained using a paired *t*-test comparing the indicated time points. **f**, Plasma viral loads and CD4 counts in all study participants. 3BNC117 infusions are indicated with red arrows. The left *y*-axis shows plasma viral loads in RNA copies per ml (black curves), and right *y*-axis shows absolute CD4 counts in cells per µl (red curves). Grey areas indicate ART therapy.

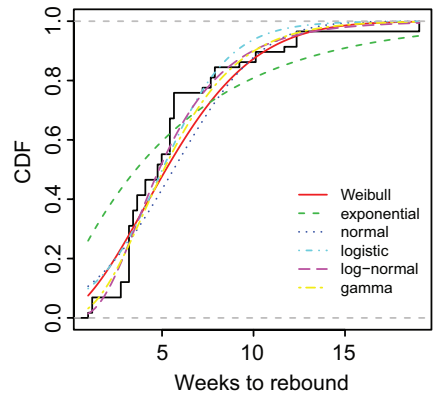


Extended Data Figure 3 | HIV-specific T-cell responses. Total breadth (open squares) and magnitude (bars) of T-cell responses against HIV-1 overlapping peptides (OLPs) at the designated time points following administration of 3BNC117 (yellow arrows indicate infusions of 3BNC117 at 30 mg kg^{-1}). For all study participants, antiretroviral therapy was discontinued on day 2 after the first 3BNC117 administration. Blue arrows indicate the time of viral rebound. For study participants 710, 712 and 715 rebound occurred at week 19, 16 and 11, respectively. Baseline samples for study participant 710 and week 12 samples for study participant 714 were not available for ELISpot analysis. Overall, breadth, magnitude and protein specificity were heterogeneous among the study participants.

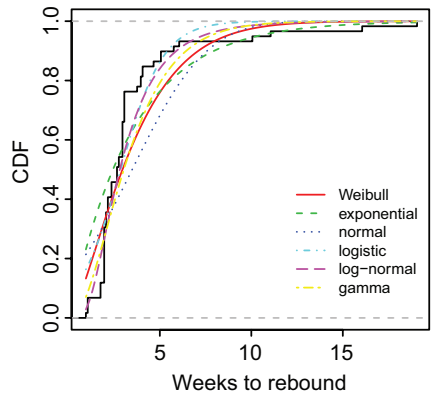




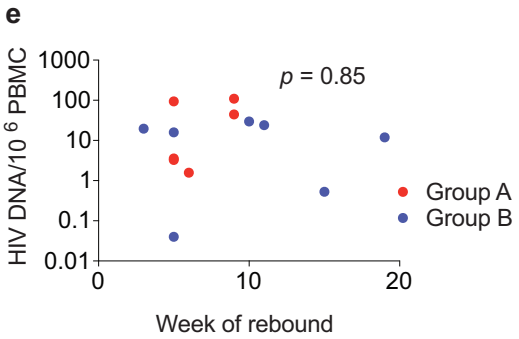
Distribution	AIC	BIC
Weibull	292.74	297.09
exponential	303.11	305.28
normal	340.57	344.91
logistic	314.41	318.76
log-normal	267.44	271.79
gamma	283.06	287.41



Distribution	AIC	BIC
Weibull	212.85	216.97
exponential	245.04	247.10
normal	228.66	232.78
logistic	218.50	222.62
log-normal	197.98	202.10
gamma	203.48	207.60

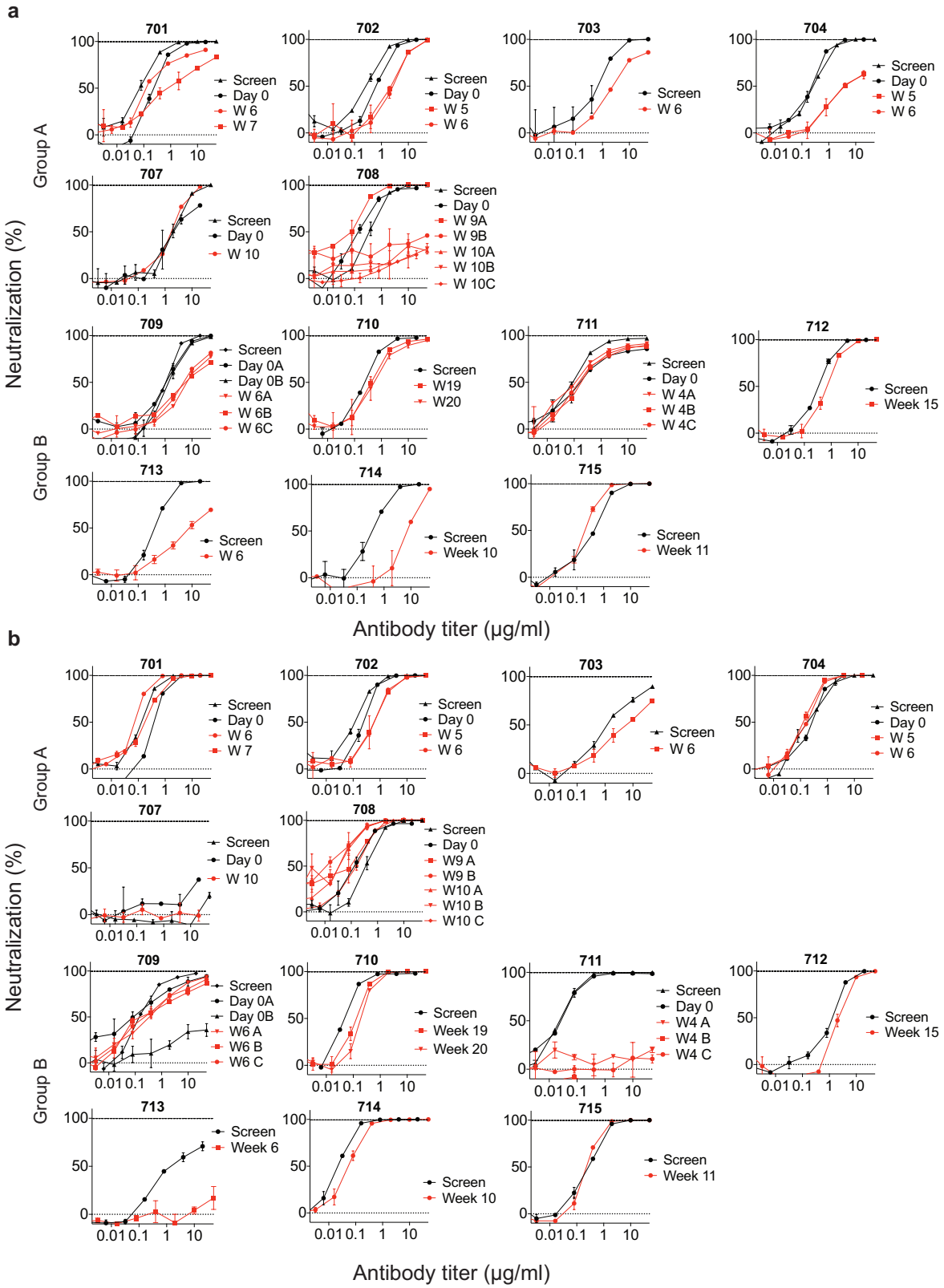


Distribution	AIC	BIC
Weibull	257.36	261.52
exponential	265.46	267.54
normal	308.32	312.48
logistic	273.96	278.12
log-normal	228.17	232.33
gamma	246.67	250.82

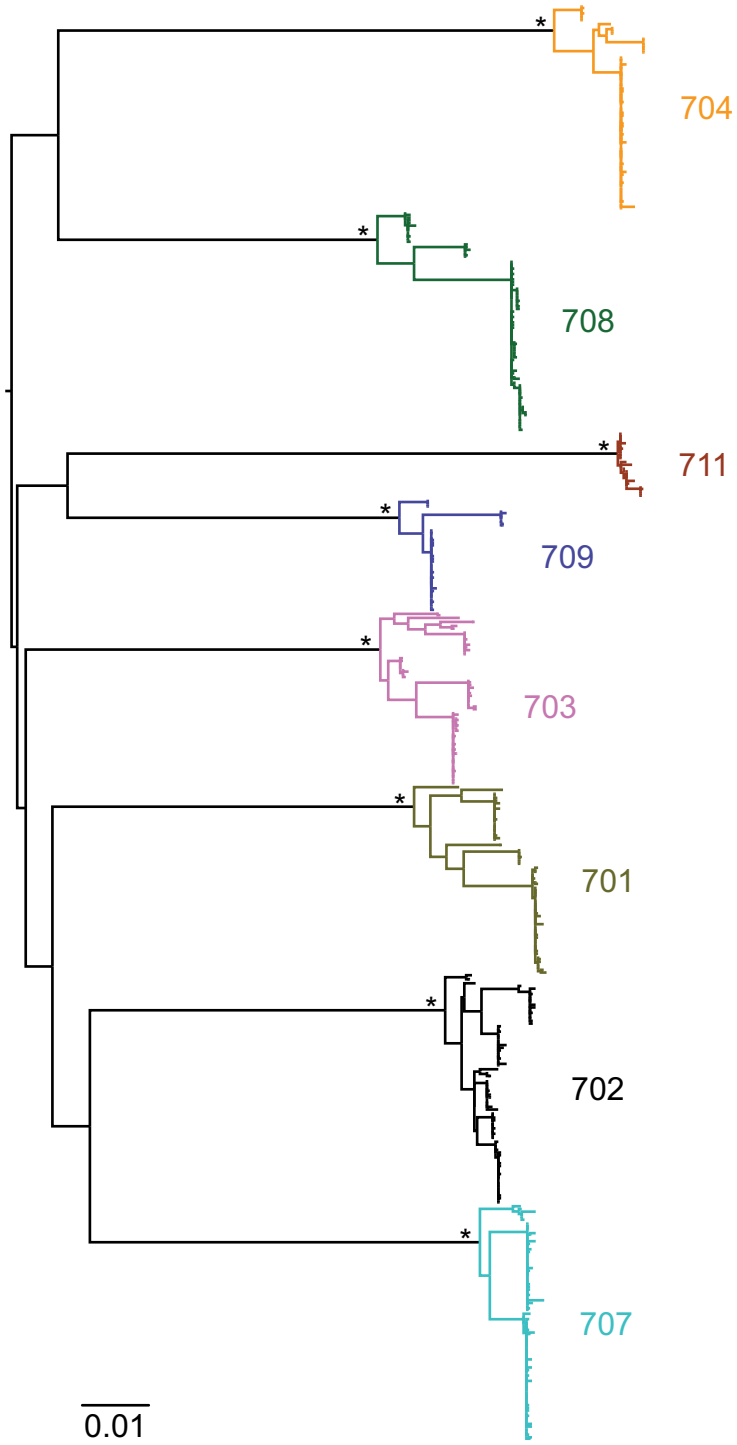


Extended Data Figure 4 | See next page for caption.

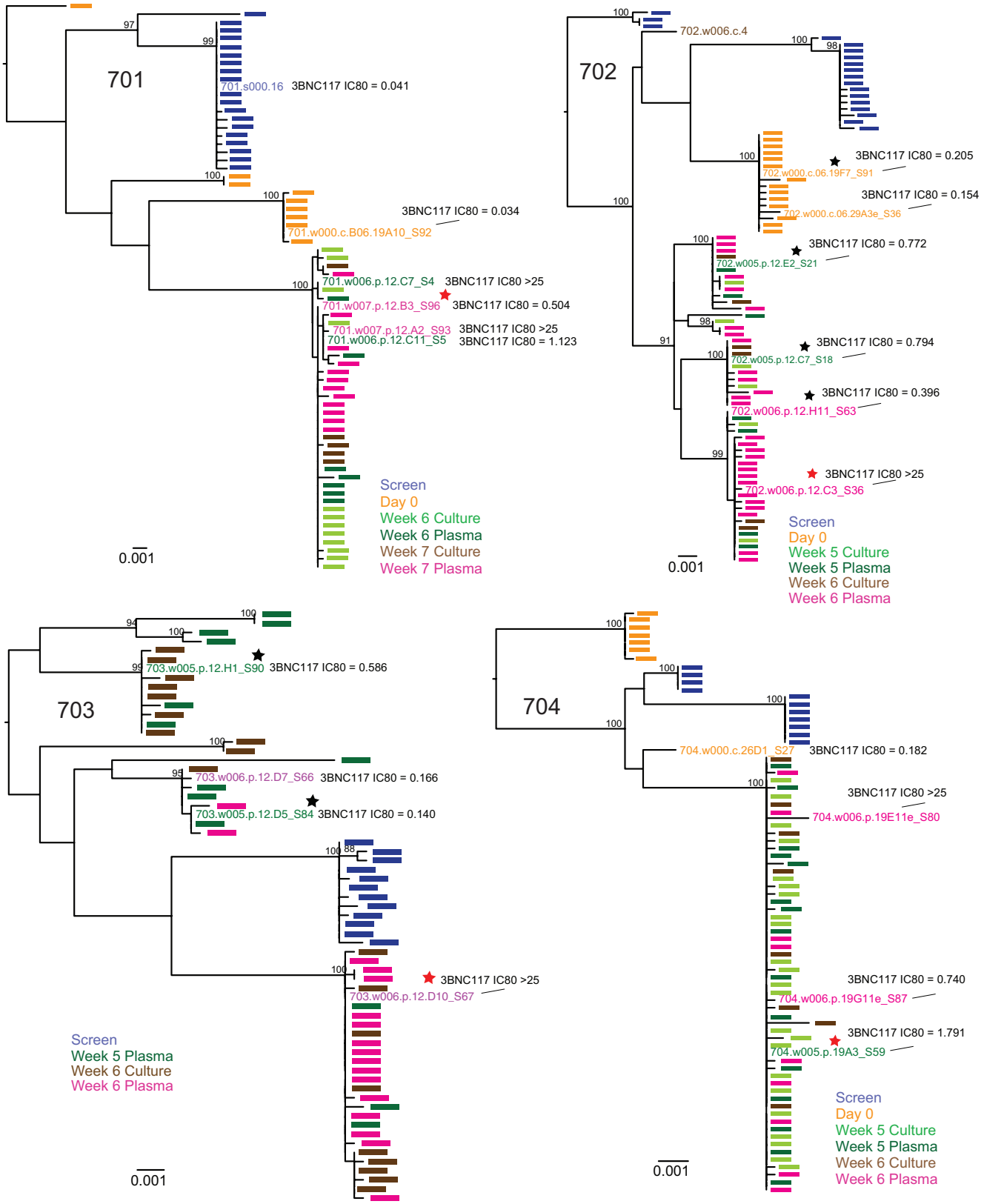
Extended Data Figure 4 | Viral rebound in ACTG control subjects and trial participants. a, Kaplan–Meier plot summarizing viral rebound in 52 ACTG trial participants who underwent ATI without antibody treatment (black, Supplementary Table 6) and trial participants (Fig. 2a, b, Supplementary Table 4). Six group A participants are shown in red, seven Group B participants in blue and the combination in green as indicated. The *y*-axis indicates the percentage of participants with viral levels below 200 RNA copies per ml, *x*-axis indicates weeks after ATI initiation. The survival curves of all considered partitions of the trial participants (group A, group B and group A + B) differed significantly at significance level $\alpha = 0.05$ from the survival curve of the ACTG trial participants. For the comparison of group A (group A + B) with the ACTG trial participants, we performed a weighted log-rank test adjusting for the clinical variables 'years on ART' and 'age' to correct for possible confounding factors (Supplementary Table 7, $P < 0.00001$). We identified those potential confounders by univariate parametric survival regression using a likelihood ratio test (Statistical Methods). Since we did not discover any confounders with the same analysis among all available clinical variables for the comparison between group B participants and the ACTG trial participants, we performed a standard log-rank test in that setting ($P < 0.0001$). **b**–**d**, In order to perform a survival regression, the distribution of the rebound times has to be determined. Therefore, we compared the empirical cumulative distribution function (CDF) of the rebound times (black, solid line) with the CDF of the rebound times to a fitted distribution (Weibull, exponential, normal, logistic, log-normal, and gamma) for each comparison group (combined trial participants, group A or group B with ACTG control patients). Since the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were smallest for the log-normal distribution (green), we have chosen to model the rebound times with the log-normal distribution. **e**, Dot plot indicating the relationship between cell associated HIV DNA in pre-infusion PBMCs (*y*-axis) and the week of rebound (*x*-axis). Group A and group B participants are coloured red and blue respectively. The *P* value was derived from calculating the Pearson correlation coefficient.



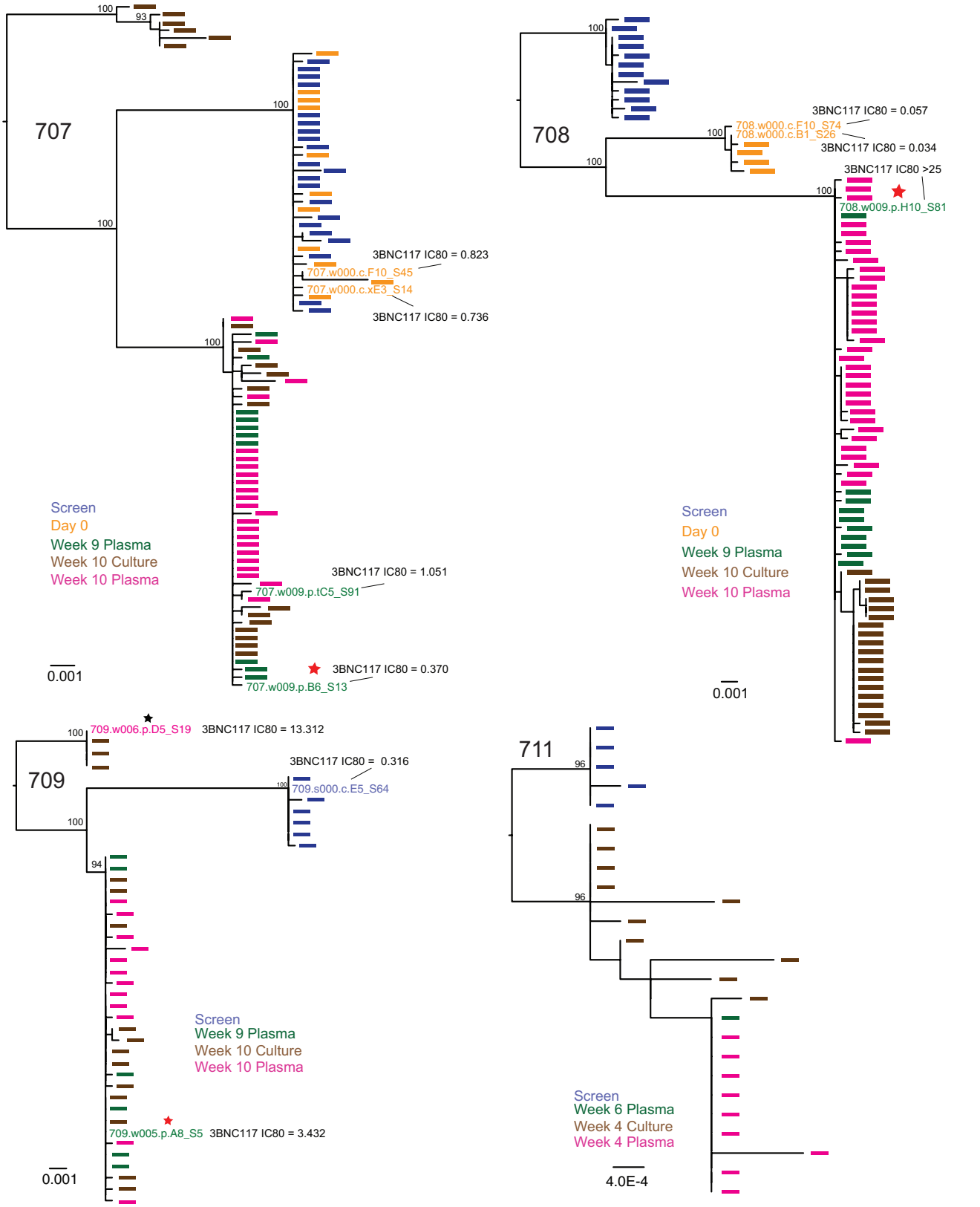
Extended Data Figure 5 | *In vitro* neutralization of pre-infusion and rebound virus outgrowth cultures by 3BNC117 or 10-1074. a, **b**, TZM-bl assay neutralization by 3BNC117 (**a**) and 10-1074 (**b**) are shown for individual virus outgrowth cultures derived from pre-infusion (black lines/symbols) or rebound (red lines/symbols) time points for each participant. In some cases, multiple independent cultures were grown from a single time point and assayed for neutralization (Supplementary Table 3). 'Screen' refers to cultures of PBMC samples taken weeks before infusion during screening, while 'Day 0' refers to cultures of PBMCs collected immediately before the first 3BNC117 infusion. Rebound culture time points are denoted by the week (W) at which the samples were collected. Symbols reflect the means of two technical replicates; error bars denote standard deviation.



Extended Data Figure 6 | Phylogenetic tree of *env* nucleotide sequences from trial participants. A maximum likelihood phylogenetic tree was constructed from single-genome-derived viral *env* sequences from outgrowth culture supernatants as well as plasma from participants 701 (olive), 702 (black), 703 (pink), 704 (yellow), 707 (light blue), 708 (green), 709 (dark blue) and 711 (brown). Hypervariable (as defined in http://www.hiv.lanl.gov/content/sequence/VAR_REG_CHAR/) and other poorly aligned regions were excluded from the analysis. The tree was constructed using PhyML with a GTR+I+G substitution model and midpoint rooted. Asterisks indicate 100% bootstrap support (only values for major nodes are shown). The scale bar indicates 0.01 substitutions per site.



Extended Data Figure 7 | Rebound virus clonality and neutralization sensitivity to 3BNC117. Maximum likelihood phylogenetic trees of plasma and culture-derived *env* sequences are shown for participants 701, 702, 703, 704. Sequences obtained at screening, on Day 0, and consecutive rebound time points (plasma and cultures) are colour coded as indicated. The trees were rooted based on the branch insertion identified in the between-subject tree (Extended Data Fig. 6). Bootstrap values ≥90% are shown. Names of *env* sequences used to generate pseudoviruses for 3BNC117 neutralization analysis are indicated along with the respective IC_{80} titres in μg ml^{-1}. Representative rebound viruses selected in Fig. 4b are marked with red stars (Fig. 4b, Supplementary Table 9). Zero branch length viruses in multi-rebounders 702 and 703 are marked with black stars.



Extended Data Figure 8 | Rebound virus clonality and neutralization sensitivity to 3BNC117. Maximum likelihood phylogenetic trees of plasma- and culture-derived *env* sequences are shown for participants 707, 708, 709 and 711. Sequences obtained at screening, on day 0, and consecutive rebound time points (plasma and cultures) are colour coded as indicated. The trees were rooted based on the branch insertion identified in the between-subject tree (Extended Data Fig. 6). Bootstrap values $\geq$90% are shown. Names of *env* sequences used to generate pseudoviruses for 3BNC117 neutralization analysis are indicated along with the respective IC_{80} titres in $\mu g\ ml^{-1}$. Representative rebound viruses selected in Fig. 4b are marked with red stars (Fig. 4b, Supplementary Table 9). Zero branch length virus in multi-rebounder 709 is marked with a black star.

A.4. Paper 4 - HIV immunoadaptation study

Copyright clearance

Insights to HIV-1 coreceptor usage by estimating HLA adaptation with Bayesian generalized linear mixed models

Anna Hake[1,2*], Anja Germann[3], Corena de Beer[4,5], Alexander Thielen[6], Martin Däumer[7], Wolfgang Preiser[4,5], Hagen von Briesen[3], Nico Pfeifer[8,9*]

1 Research Group Computational Biology, Max Planck Institute for Informatics, Saarbrücken, Germany
2 Saarbrücken Graduate School of Computer Science, Saarland University, Germany
3 Main Department Medical Biotechnology, Fraunhofer Institute for Biomedical Engineering, Sulzbach, Germany
4 Division of Medical Virology, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa
5 National Health Laboratory Service, Tygerberg Business Unit, Cape Town, South Africa
6 Seq IT GmbH & Co.KG, Kaiserslautern, Germany
7 Institute of Immunology and Genetics, Kaiserslautern, Germany
8 German Center for Infection Research, Partner Site Tübingen, Tübingen, Germany
9 Methods in Medical Informatics, Department of Computer Science, University of Tübingen, Germany

* nico.pfeifer@uni-tuebingen.de, ahake@mpi-inf.mpg.de

Abstract

The mechanisms triggering the human immunodeficiency virus type I (HIV-1) to switch the coreceptor usage from CCR5 to CXCR4 during the course of infection are not entirely understood. While low $CD4^+$ T cell counts are associated with CXCR4 usage, a predominance of CXCR4 usage with still high $CD4^+$ T cell counts remains puzzling. Here, we explore the hypothesis that viral adaptation to the human leukocyte antigen (HLA) complex, especially to the HLA class II alleles, contributes to the coreceptor switch. To this end, we sequence the viral *gag* and *env* protein with corresponding HLA class I and II alleles of a new cohort of 312 treatment-naive, subtype C, chronically-infected HIV-1 patients from South Africa. To estimate HLA adaptation, we develop a novel computational approach using Bayesian generalized linear mixed models (GLMMs). Our model allows to consider the entire HLA repertoire without restricting the model to pre-learned HLA-polymorphisms as well as to correct for phylogenetic relatedness of the viruses within the model itself to account for founder effects. Using our model, we observe that CXCR4-using variants are more adapted than CCR5-using variants (p-value =1.34e-2). Additionally, adapted CCR5-using variants have a significantly lower predicted false positive rate (FPR) by the geno2pheno[coreceptor] tool compared to the non-adapted CCR5-using variants (p-value =2.21e-2), where a low FPR is associated with CXCR4 usage. Consequently, estimating HLA adaptation can be an asset in predicting not only coreceptor usage, but also an approaching coreceptor switch in CCR5-using variants. We propose the usage of Bayesian GLMMs for modeling virus-host adaptation in general.

Author summary

Viral control is currently our only counter mechanism against HIV-1 with no practicable cure nor a vaccine at hand. In treatment-naive patients, HLA adaptation and coreceptor usage of HIV-1 play a major role in their capability to control the virus. The interplay between both factors, however, has remained unexplored so far. Assessing the degree of viral HLA adaptation is challenging due to the exceptional genetic diversity of both the HLA complex and HIV-1. Therefore, current approaches constrain the adaptation prediction to a set of p-value selected HLA-polymorphism candidates. The selection of these candidates, however, requires extensive external large-scale population-based experiments that are not always available for the population of interest, especially not for newly emerging viruses. In this work, we present a novel computational approach using Bayesian generalized linear mixed models (GLMMs) that enables not only to predict the adaptation to the complete HLA profile of a patient, but also to handle phylogenetic-dependencies of the variants within the model directly. Using this light-weight approach for modeling (any) virus-host adaptation, we show that HLA adaptation is associated with coreceptor usage.

Introduction

Without the prospect of a vaccine or a cure within reach, viral control is one of the major pillars for combating the HIV pandemic [1]. The coreceptor usage of HVI-1 affects the ability to control the virus. Apart from the CD4 receptor, HIV-1 needs a coreceptor for successful cell entry. Only two coreceptors have clinical relevance, CCR5 and CXCR4 [2]. Depending on their coreceptor usage, HIV-1 isolates are classified into R5-capable variants (CCR5 usage), X4-capable variants (CXCR4 usage), or R5X4-capable variants (CCR5 and CXCR4 usage) [3,4]. While R5 variants are known to dominate early infection [5], a switch to X4 at later stage of infection occurs in roughly 50% of patients infected with subtype B HIV-1 associated with increased depletion of $CD4^+$ T cells, faster progression to AIDS, and a higher mortality rate [4,6–8]. In patients infected with subtype C HIV-1, a switch to CXCR4 usage is observed less frequently compared to subtype B [9]. Recent studies suggest that an increase in subtype C X4 variants might emerge with the increasing access to antiretroviral drug treatment and the ongoing evolution of the subtype C HIV epidemic [10].

The importance of accurate determination of coreceptor usage has increased with the approval of entry inhibitor drugs that target the CCR5 coreceptor. A determinant of coreceptor usage is the *env* protein of HIV-1. Currently, phenotypic [11–14] and genotypic [15–22] tropism assays still have difficulties accurately detecting minority populations of X4-using variants, which might lead to a predominance of X4 usage after treatment with a CCR5 antagonist. In addition, it is not only important to predict the correct coreceptor usage, but it would be of advantage to predict how close the variant is to a coreceptor switch.

Though the clinical significance of the coreceptor usage is well studied, the trigger mechanisms behind the coreceptor switch from R5 to X4 variants remain unsolved. The emergence of X4-capable variants is associated with a decrease in N-linked glycosylation of the envelope glycoprotein *env* of HIV-1 [23]. Glycosylation is a viral mechanism to mask conserved amino acids from antibody recognition, such that X4-capable variants should be more prone to antibody neutralization in theory. For antibody development, B cells have to be activated by $CD4^+$ T cells. Thus, concurrent $CD4^+$ T cell depletion counteracts this mechanism. How X4 variants can emerge with still high $CD4^+$ T cells remains inconclusive. However, this is of great importance, since patients with intermediate to high $CD4^+$ T cells contradict the current typical clinical indicators for a

potential coreceptor switch such as low numbers of CD4+ T cells.

The potential interplay between HLA adaptation and coreceptor usage has not been explored so far. Viral adaptation to the immune system includes the emergence of viral escape mutations to the host's individual HLA profile. The central role of HLA molecules is to bind peptides and present them on the cell surface to compatible T cells, which are part of the adaptive immune response. T cells are HLA-restricted, meaning that they recognize only a specific HLA-antigen complex. There are two major HLA classes — HLA class I and HLA class II. HLA class I molecules exist on all nucleated cells and bind to (self and pathogen-derived) antigens degraded from synthesized proteins in the cytosol. The corresponding HLA-antigen complex is recognized by specific $CD8^+$ T cells. HLA class II molecules only occur on professional antigen-presenting cells that are able to uptake pathogens and proteins from extracellular fluid by phagocytosis or endocytosis. Thus, HLA class II molecules bind pathogen-derived antigens degraded from extracellular proteins in the vesicular compartment of the cell. The corresponding HLA:antigen complex is recognized by $CD4^+$ T cells. The emergence of a mutation that hinders the successful building of the HLA-antigen complex, a so-called escape mutation, allows HIV-1 to evade a T cell-mediated immune response [24, 25].

High-throughput technologies have enabled large-scale population studies to identify many HLA-restricted polymorphisms (HLA footprints) and their role on viral control [26–31]. A prominent example is the influence of the HLA-B*27 and the HLA-B*57:01 allele on disease progression [32, 33]. Determining virus-host adaptation experimentally and computationally on an individual level is challenging due to the extraordinary genetic diversity of both the HLA complex and HIV-1. HLA adaptation models usually focus on viral polymorphisms that likely emerged due to the patient's HLA profile. This approach requires the general consideration of the extreme large number of possible HLA alleles in the population and viral polymorphisms while modeling the fact that only few HLA alleles have a potential influence on a particular polymorphism. Current computational approaches [34, 35] tackle the complex modeling task by carrying out many rounds of preselection, including the identification of potential HLA-polymorphism candidates on large-scale cohort data and additional greedy feature selection steps to select the HLA alleles per polymorphism within the model, such that potential sites and HLA alleles might get disregarded based on significance threshold values. Since human populations and HIV subtypes display substantial genetic differences, such approaches require a large amount of data for every group of interest. Correcting for potential phylogenetic relatedness of the viral sequences used within the model as proposed by [36] is currently implemented by incorporating a transmission probability that has to be learned in a separate model. While HLA-1 restricted escape mechanisms to CTLs have been studied in detail, only few studies exist that have analyzed the impact of HLA-restricted $CD4^+$ T cell escape polymorphisms [35] on viral control. In total, there is currently no available approach to estimate viral adaptation jointly to HLA class I and class II. Moreover, the available approaches require rather complex training steps to be used on new data.

In this study, we investigate the hypothesis that coreceptor usage is associated with the adaptation of the virus to the host's HLA system, especially to the HLA class II alleles. We explore the novel possibility that viral adaptation to the HLA class II molecules would mask the virus from recognition by $CD4^+$ T cells, such that no B cells are activated, and, thus, no antibodies are developed despite still high numbers of $CD4^+$ T cells. Escape mutations in the rather conserved *p24* protein of HIV-1, which is involved in forming the viral capsid, emerge more likely under substantial fitness cost [37, 38]. Therefore, we estimate viral adaptation to the patient's HLA profile only based on the *p24* protein of the *gag* gene as done previously [39–44]. This study

requires a data set consisting of (1) the envelope protein sequences of the virus for determining the coreceptor usage, (2) the p24 protein for estimating the HLA adaptation, and (3) the HLA class I and II profile of the corresponding host. Chronically-infected HIV-1 patients are more likely to harbor viruses that have accumulated escape mutations to the HLA system due to the longer exposure to the human immune system. In treatment-naïve patients, the viral evolution is not restricted by selection pressure from drug exposure and is more able to mutate towards escape variants with respect to the immune system. Current available data sets often lack HLA class II allele information or have not sequenced the envelope sequence of the virus. Thus, we sequence the viral envelope gene *env* as well as the viral *gag* (*p24*) gene, and genotype the corresponding HLA class I (HLA-A, HLA-B, HLA-C) and II genes (HLA-DRB1, HLA-DQB1, HLA-DPB1) of the host in a new cohort of 312 treatment-naive, subtype C, chronically-infected HIV-1 patients from South Africa.

To jointly model HLA class I and class II adaptation, we develop a novel computational approach. In detail, the adaptation of a particular amino acid in a viral sequence to the host HLA profile is inferred using phylogeny-corrected, multinomial, Bayesian generalized linear mixed models (GLMMs). Without the need for an additional model, GLMMs allow to correct for phylogenetic relatedness of the variants directly by modeling the between-subject correlation as a group-level effect. Using a Bayesian setting allows to learn feature importance directly within the model by applying the horseshoe prior on all HLA class I and class II alleles of the data set and without the need for additional preselection steps or a large amount of data. The horseshoe prior is used in sparse model settings to shrink the majority of the coefficients to zero by having the point mass at zero and symmetric fat tails [45].

Materials and methods

Study cohort

Patients (male and female) who attended Wellness, Antenatal and HIV Clinics in the Durbanville and Stellenbosch regions of the Western Cape were recruited. Only patients older than 18 years were selected. Most of the patients were assumed to be in the chronic stage of the infection. Inclusion was based on recent diagnosis of HIV-1 infection (within the previous 6 months). In total, samples from 329 HIV-infected individuals were available. Subtype C was confirmed for 317 of the 329 samples using the COMET Tool [46]. Patients on antiretroviral were excluded from the analysis, resulting in a total of 312 patients. For each patient, clinical parameters such as sex, age, ethnicity, CD4 count, and viral load were collected. In addition, the HIV-1 genes *gag* (p24) and *env* were sequenced and the patients' HLA I and II genes were genotyped.

Ethical statement

PBMC and plasma samples from HIV-1 positive donors were provided by Stellenbosch University with the written informed consent of the donors. Sample collection was approved under the following ethical statement "VIROLOGICAL AND IMMUNOLOGICAL CHARACTERIZATION OF CRYOPRESERVED BLOOD AND VIRUS SAMPLES" PROJECT NUMBER: NO7/06/13

Molecular methods

HIV status was confirmed with a serological test (Architect HIV Ab/Ab Combo, 3rd generation) on serum according to the manufacturer protocol. After the surface staining

of PBMCs by incubation with a monoclonal mouse anti-human antibody coupled to fluorescent dyes, the quantification of cells expressing the CD4 antigen was measured by FACS analysis. Acquisition and analysis was performed on FACs flow cytometer using Cell Quest software.

HIV-1 deep sequencing was performed using previous described protocols [47]. Analysis of deep sequencing data was performed using an internally-developed analysis pipeline, where sequence reads in the form of FASTQ files were processed and aligned via a multi-step method.

HLA genotyping was performed using the following protocol. Genomic DNA was isolated from 200 μl of EDTA-anticoagulated blood using the QIAamp DNA Blood Mini Kit (QIAGEN, Hilden, Germany). Long-range PCR primers amplified the full-length of HLA class I genes (A, B, C) from 5'- to 3'-UTR. Class II genes (DPB1, DQB1, DRB1) were amplified from exon 2 to 3'-UTR. Fragment sizes were estimated to be around 3000 bp for Class I genes and 6000 bp for Class II genes, respectively. The PCR solution contained 1 x Phusion GC buffer (including 1.5 mM MgCl2), 200 μM dNTPs, 1 M Betaine, 8 μg Bovine Serum Albumin (BSA), 0.4 U Phusion Hot Start II High-Fidelity DNA Polymerase (Finnzymes, Vantaa, Finland), 0.5 μM of each primer and 90 ng of DNA in a total volume of 20 μl. After initial denaturation at 98°C for 1 minute, 35 cycles of 98°C for 10 seconds, 65°C for 20 seconds, and 72°C for 4 minutes were performed, followed by a final extension at 72°C for 20 minutes. Agarose gel electrophoresis was used to confirm amplification and correct fragment size as well as to check for non-specific product contamination. The 3 HLA class I and class II amplicons for each individual were pooled and afterwards purified with the Agencourt AMPure XP system (Agentcourt Bioscience, Beverly, MA, United States) according to the manufacturer's protocol to inactivate unconsumed dNTPs and to eliminate extraneous primers before library preparation. These pooled amplicons then comprised a single sample. Concentrations were measured on a FLUOstar OPTIMA microplate fluorimeter (BMP LABTECH, Ortenberg, Germany) using the Quant-iT PicoGreen assay (Invitrogen, Carlsbad, CA, United States). Sample libraries for NGS were then prepared with the Nextera XT DNA Sample Prep Kit (Illumina, San Diego, CA, United States) according to the manufacturer's protocol, including distinct DNA fragmentation, end-polishing, and adaptor-ligation steps. Through the adaptor, every sample was finally labeled with a unique identifier sequence. Sequencing was carried out then on the Illumina MiSeq Personal Sequencer (Illumina, San Diego, CA, United States) as described by the manufacturer.

Coreceptor prediction

Coreceptor usage is predicted using the well-established tool geno2pheno[coreceptor] [17] on the viral envelope sequences. The provided false-positive rate (FPR) corresponds to the confidence with which the sequence is classified as X4-capable. The higher the FPR, the more likely the sequence is not X4-capable, but R5. Viral strains with an FPR cutoff less than 20% are classified as X4-capable, otherwise as R5-capable according to the European Consensus Group on clinical management of HIV-1 tropism testing [48].

Estimating HLA adaptation

Assuming independence of all sites in the viral sequence, we define the adaptation of a sequence to its host HLA profile as the adaptation of each frequent single amino acid site in the sequence to the HLA profile. Moreover, though every patient is infected by a quasispecies of viruses, we only consider the consensus sequence as in previous approaches. In order to correct for potential phylogenetic relatedness of the viral sequences used within the model as proposed by [36], we also incorporate the phylogeny

of the viral sequences into the model. Thus, our model requires the amino acid sequences of the viral p24 protein, the corresponding host's HLA I and II alleles, and the phylogeny between the viral sequences for learning the HLA adaptation (training). For each frequent site, we infer a model (HLA model) to estimate the likelihood that the site is under HLA pressure as well as a hypothetical model (baseline model) that computes the likelihood that the site is not under HLA pressure. HLA adaptation of the complete protein is then defined as a function over the product of the per-site likelihood ratios of the HLA model against the baseline model. Each per-site model is built using multinomial Bayesian generalized linear mixed models (GLMMs).

In the following, we formalize the per-site model and the final adaptation score. Afterwards, we present the selection process of the frequent sites. Since each per-site model is built using Bayesian GLMMs, we provide a brief introduction to Bayesian GLMMs and their benefit over classical GLMMs and phylogeny-corrected LMMs. In addition, we provide a section on the model specification for each per-site model.

Notation

Let S be a random variable representing the set of all possible HIV-1 amino acid sequences of a particular protein of length L. A particular sequence s is a realization of S covering all sites $l = 1, \ldots, L$ of the protein. A particular site s_l can be realized by any amino acid (aa). Since we do not have enough power to find an HLA-restricted polymorphism at a very conserved site, we restrict the sites to m frequent single amino acid sites s_j with $j = 1, \ldots, m$, which are defined by sites that vary over the set of all HIV-1 sequences in their amino acid realization. A site is defined as frequent, if the particular amino acid is observed in at least 1% of the sequences.

The host immune system is represented by the HLA alleles of the HLA I and HLA II genes. The HLA profile of an individual consists in our case of six (homozygous in all genes) to 12 (heterozygous in all genes) different HLA alleles. Let H represent the set of all possible HLA I and HLA II alleles. A particular HLA profile h is encoded as a binary vector with zeros everywhere, apart from the positions corresponding to the HLA alleles of the HLA profile. Note, thereby homozygosity is not modeled.

We model adaptation as the conditional probability that a sequence s occurs under pressure from the host HLA profile similarly to [34] :

$$P(S = s|H = h). \tag{1}$$

Assuming independence among sites and relevance of only frequent sites, the conditional probability over the sequence s can be decomposed to the product over the conditional probabilities over all m frequent sites s_j (per-site model):

$$P(S = s|H = h) = \prod_{j=1}^{m} P(s_j|H = h). \tag{2}$$

Similarly, a hypothetical model estimating the likelihood of the sequence s without any HLA pressure is defined as:

$$P(S = s|H = \varnothing) = \prod_{j=1}^{m} P(s_j|H = \varnothing) \tag{3}$$

The conditional probabilities $P(s_j|H = h)$ and $P(s_j|H = \varnothing)$ for each site are referred to as the HLA model and the baseline model, respectively.

Identification and encoding of frequent sites

Since all patients in the study are infected with the subtype C variant of HIV-1, we align all nucleotide sequences to the subtype C consensus sequence using the alignment tool MAFFT (version 7.407) [49]. The subtype C consensus sequence is retrieved using the HIV Sequence Alignments tool from the Los Alamos HIV sequence database (www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html). We correct and translate the nucleotide alignment using the Codon Align Tool from the Los Alamos HIV sequence database (www.hiv.lanl.gov/content/sequence/CodonAlign/codonalign.html). The alignment positions are mapped to the corresponding HXB2 reference gene with Genbank accession 'AAB50258.1' (gag) using the alignment tool MAFFT (version 7.407) [49]. Ambiguous amino acids X are not considered and set to NA. Frameshifts and stop codons are disregarded and set to gaps. Each site in the sequence s with at least two frequent (1% prevalence) amino acid variants is selected as potential site s_j under HLA pressure. For each frequent site and each hypothesis (HLA and baseline model), a multinomial Bayesian generalized linear mixed model is built, where each frequent amino acid is considered a class, and all non-frequent amino acids are grouped together to an 'OTHER' class.

Bayesian generalized linear mixed models

We model the conditional probabilities for site adaptation (see Eq. 2 and Eq. 3) using separate multinomial Bayesian generalized linear mixed models (GLMMs). GLMMs are tailored for data with non-normal response distributions and dependency structures in the observations by combining the properties of generalized linear models (GLMs) [50, 51] and linear mixed models (LMMs). While GLMs model non-normal response distributions (such as binomial) via link functions of the means (e.g. logistic regression), LMMs enable to model not only population-level effects but also group-level effects assuming dependency structures in the samples. Mathematically, GLMMs have the following form excluding the residuals (ϵ) [52]:

$$g(E(Y|X,Z,u)) = \eta = X\beta + Zu, \tag{4}$$

where Y is the response variable, β and u the coefficients for the population and group-level effects, respectively, X and Z the corresponding design matrices and $g(x)$ a link function relating the response Y to the linear predictor η. Thus, between-subject correlations, like the phylogenetic relatedness of some viruses, can be modeled as a group-level effect.

While y, X and Z are given by the data, β and $u \sim \mathcal{N}(0, G)$ are unknown and have to be estimated. We use Markov chain Monte-Carlo (MCMC) based Bayesian GLMMs, since they are more robust and accurate in their parameter estimations of the group-level effects in contrast to classical maximum likelihood (ML) and restricted maximum likelihood (REML) methods [53]. In non-Bayesian frameworks, the group-level effect vector u is treated as part of the error term and thus likelihood computation requires the integration over the likelihood of all group-level effects, which might be analytically intractable for complex group-level structures [54]. In Bayesian settings where posterior distributions of the parameters are estimated by combining likelihood and prior distributions, both u and β are treated as parameters, allowing more accurate variance estimates for the group-level effects. We use the MCMC Bayesian GLMM implementation of the R [55] package brms [56] that provides an interface to the STAN software [57]. By implementing Hamiltonian Monte Carlo [58] and the No-U-Turn Sampler (NUTS) [59], Stan allows for faster convergence compared to conventional MCMC methods. Another advantage of Bayesian models is the possibility to include the prior information of the parameters into the model. The prior

knowledge that only few HLA alleles have potential influence on a variant site [31,34] is modeled using the horseshoe prior that has a global parameter τ shrinking most of the coefficients to zero and a local parameter λ, which is a heavy-tailed half-Cauchy ($C^+(0,1)$) prior, allowing some coefficients to escape the shrinkage [45]. Thus, the the horseshoe prior for the D population level coefficients $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_D)$ has the following form:

$$\begin{aligned} \beta_j|\lambda_j, \tau &\sim N(0, \lambda_j^2\tau^2), \\ \lambda_j &\sim C^+(0,1), \quad j = 1, \ldots, D. \end{aligned} \tag{5}$$

In addition, we regularize the horseshoe prior by setting the ratio of the expected number of non-zero coefficients to the expected number of zero coefficients to 10%. All other parameters of the horseshoe prior are set to default. For the remaining coefficients the default priors of the brm function are used (non or very weakly informative priors).

Estimating per-site adaptation

For each frequent site, we model an HLA model (see Eq. 2) and a baseline model (see Eq. 3) using multinomial Bayesian generalized linear mixed models (GLMMs) as implemented by the brms package [56] in R [55]. Both models estimate the probability distribution of each site s_j spanning over the space Y of all frequent amino acid variants (and 'OTHER' for the non-frequent variants') conditioned on the potential confounders age, sex, and ethnicity. Age is defined as the interval between sample extraction date and birthday and scaled to mean 0 and variance 1. If missing, months and days are set to the first day and month, respectively. Due to the ambiguous recording of ethnicity groups, samples are assigned to either African, Caucasian, or 'Other' ethnicity. Sex is modeled as a binary feature. Though deep sequencing has been performed, we use for this study only the consensus sequences derived using a 10% prevalence cutoff, which is commonly used in the research community [60]. The NGS reads were mapped with a customized version of MinVar [61].

Predicting if a polymorphism is under HLA pressure or not is confounded by the phylogenetic relatedness of the viral sequences. As proposed by [62], the phylogeny of the viral sequences of the subjects is incorporated as group-level effect $(1|subject)$ into the model using the option cov_ranef = $list(subject = A)$. Here, A denotes the computed covariance-matrix of the phylogenetic tree calculated using the vcv.phylo function from the ape package. A phylogenetic tree is constructed based on the nucleotide sequences of the p24 protein from the chronic_lowCD4 data set using the RAxML software (version 8.2.12) [63] under the GTRGAMMA model. Thus, the formula to compute the HLA model taking all HLA alleles H as potential covariates into the model has the following form:

$$Y \sim age + sex + ethnicity + (1|subject) + H, \tag{6}$$

in contrast to the baseline model, which estimates the probability that the frequent site is not under HLA pressure:

$$Y \sim age + sex + ethnicity + (1|subject). \tag{7}$$

The logistic function is used as a link function. As described in previous sections, the horseshoe prior is used on all population-level effects [45]. Alleles in H are converted to four digit resolution. Alleles with alternative expression (suffix 'L', 'S', 'C', 'A', or 'Q') are treated separately from the normally expressed allele. The complete call to compute the per-site models using the brms package is provided in the code repository.

Calculation of adaptation score

We define the adaptation score, as proposed by [34], as:

$$\begin{aligned} adapt(s,h) &= \frac{P(S=s|H=h)}{P(S=s|H=\varnothing)} \\ &\approx \frac{\prod_{j=1}^{m} P(s_j|H=h)}{\prod_{j=1}^{m} P(s_j|H=\varnothing)} \\ &= \prod_{j=1}^{m} \frac{P(s_j|H=h)}{P(s_j|H=\varnothing)}, \end{aligned} \tag{8}$$

where the per-site likelihood $P(s_j|H=h)$ and $P(s_j|H=\varnothing)$ are defined by Eq.2 and Eq. 3, respectively. For better interpretation, we also transform the estimated adaptation $adapt(s,h)$ using a sigmoidal function $g(x)$ to a range of -1 to 1 [34]:

$$g(x) = \frac{2}{\pi}\arctan(\ln(x))$$

Thereby, a positive adaptation score denotes that the sequence has more likely occurred under HLA pressure than without, and vice versa.

Logo computation

The adaptation score can be decomposed into the likelihood ratios per frequent variant sites (see Eq. 8). Odds ratios above or below 1 indicate that either the polymorphism at the site s_j is more likely to be under HLA pressure, or vice versa. We use this information to provide a visual logo depicting the amino acids that contributed most to the adaptation score. Therefore, only sites with odds ratios differing from 1 (and an offset of 0.01 to account for the variance) are considered. The contribution is scaled by the maximum contribution. In order to use the existing Weblogo 3.0 software to produce the logos [64], we create a pseudo-alignment with 100 sequences with length of the number of important sites. Each position in the alignment represents a polymorphism site. The sequences contain the polymorphism at this position with a frequency equal to the scaled contribution and a gap for the remaining sequences. Thus, the logo is a consensus logo for the pseudo-alignment.

Data sets

We divide the newly sequenced study cohort based on a $CD4^+$ T cell count cutoff of 500 cells/mm^3 into a chronic_highCD4 data set and a chronic_lowCD4 count data set. High $CD4^+$ T cell count indicates a stronger immune system. Since infection duration is not known for the patients, a high $CD4^+$ T cell count might indicate that the patients have been infected for a shorter time (less chronic). Moreover, a virus is assumed to be less adapted to a host with a strong immune system compared to a host with a weak immune system. Thus, the adaptation model is only trained on the chronic_lowCD4 data set. In addition, we create an artificial data set (random) based on the chronic_lowCD4 data set, where the HLA alleles per HLA gene and haplotype have been randomized 100 times. HLA adaptation for this random data set is predicted with models based on the chronic_lowCD4 data set as well. For further validation of the adaptation model, we estimate HIV-1 adaptation of publicly available cohort of

acutely-infected HIV-1 patients (n = 23) from the Los Alamos HIV sequence database (http://www.hiv.lanl.gov). The acute data set comprised the p24 sequence as well as the HLA I information of 23 patients with the following accession numbers GQ275453, GQ275750, GQ275852, GQ275894, KM192425, KM192440, KM192471, KM192536, KM192566, KM192640, KM192653, KM192674, KM192686, KM192702, KM192762, KM192844, KM192856, KM192870, KM192884, KM192912, KM192942, KM192970, KM192998. Since only the HLA I profile was available, we build an adaptation model based only on the HLA I profile for this purpose.

Statistical analyses

We perform a one-sided Wilcoxon rank-sum test to compare the adaptation scores (i) between different data sets and (ii) with respect to different clinical characteristics. For settings, where the data is paired (random data set - same subjects, R5-FPR analysis - matched CD4 count, heterologous - autologous viruses), a one-sided Wilcoxon signed-rank test is performed. A significance threshold of 0.05 is set for all hypothesis tests.

Data and code availability

The NGS sequences from the study cohort are available under the BioProject ID PRJNA810303 (reviewer link, see submission). The corresponding BioSample Accession IDs are SAMN26241863:26242168 and SAMN28728524:SAMN28728529. The generated consensus nucleotide sequences are provided on Zenodo at link 10.5281/zenodo.6797532. Due to privacy reasons, the HLA information cannot be published. Consequently, we cannot publish the trained models as the HLA information can be exposed thereby. A minimal data set including the estimated adaptation scores for all presented data sets is available on Zenodo at link 10.5281/zenodo.6797722. All code not compromising the privacy concerns, including the complete call to train and build the multinomial Bayesian generalized linear mixed models, is provided at GitHub at link https://github.com/annahake/HIVIA_TOOL.git to be used as template. To ensure reproducibility, we have used the workflow manager Snakemake 5.4.5 [65] and the Anaconda Software Distribution [66] for the training and prediction pipeline. We have used the R Language and Environment for Statistical Computing, Version 3.5.1 [55] for modeling and analyses.

Results and Discussion

Validation of the adaptation score

We trained our adaptation model on data from a cohort consisting of 274 chronically-infected, untreated, subtype C, HIV-1 patients, all having a $CD4^+$ T cell count less than 500 cells/mm^3 and on average a log viral load of 4.87 ('chronic_lowCD4' data set). In addition, 38 samples from the same study cohort with a $CD4^+$ T cell count above 500 cells/mm^3 ('chronic_highCD4' data set) were available. Apart from the $CD4^+$ T cell count, the two data sets are comparable with regard to potential confounders and clinical variables (see Table 1).

Performing several runs of 10-fold cross-validation revealed that the predicted adaptation score is quite robust, changing with an average standard deviation around 0.1. Though there exists no ground truth for HLA adaptation, we set some requirements that a valid adaptation score should fulfill, which can be seen in the following.

Table 1. Summary statistics for the variables of interest for both data sets.

variable	chronic_highCD4	chronic_lowCD4
age	35.24±10.03	35.38±9.99
$CD4^+$ T cell count	690.47±188.39	199.28±117.27
VL	4.22±0.73	4.87±0.85
adapt(x)	0.05±0.38	0.19±0.31
sex:F	0.58	0.61
sex:M	0.42	0.39
ethnicity:AFRICAN	0.79	0.64
ethnicity:CAUCASIAN	0.03	0
ethnicity:OTHER	0.18	0.36
coreceptor:R5	0.82	0.71
coreceptor:X4	0.18	0.29

Study cohort contains HLA adapted sequences

We assume that by construction the study cohort should harbor some HLA adapted sequences. 62% of the samples from the chronic_lowCD4 data set (n = 274) are estimated to be adapted (adaptation score >0.1), compared to 47% of the chronic_highCD4 data set (n = 38). The adaptation scores of the chronic_lowCD4 data set are taken from a 10-fold cross-validation, while the adaptation scores of the chronic_highCD4 data set are predicted using the full chronic_lowCD4 data set for training. Fig. 1 shows the distribution of the adaptation score in the chronic_lowCD4 data set and the chronic_lowCD4 data set. Statistically, HIV-1 isolates of patients with $CD4^+$ T cell count below 500 (chronic_lowCD4 data set) are significantly more adapted than patients with higher $CD4^+$ T cell count (one-sided, unpaired Wilcoxon rank-sum test, p-value = 1.97e-2). The comparison of the chronic_lowCD4 data set with the chronic_highCD4 data set is however not straightforward. On the one hand, the size of the chronic_highCD4 data set is quite small compared to the low $CD4^+$ T cell. On the other hand, while we exclude the patients with the higher $CD4^+$ T cell count from the training process as a precaution because they might be less chronic, this assumption does not have to be true and the samples cannot be treated to test the hypothesis that chronically-infected patients have more adapted viruses compared to patients with shorter infection duration. Last but not least, HLA-1 adapted viruses are assumed to escape the CTL response, resulting in fewer infected $CD4^+$ T cells being killed. As a consequence, it is not necessarily the case that patients with higher $CD4^+$ T cell count have less adapted viruses compared to patients with a lower $CD4^+$ T cell count.

Random HLA profile leads to non-adaptedness

We expect that viruses in the study cohort are more adapted to the host's HLA profile than to a random HLA profile. Therefore, we predicted the HLA adaptation of the viral sequences of the cohort to a random HLA profile (100 times). Adaptation scores in the random data set are averaged per patient over 100 draws. Only 10% of the random samples (n = 274) are predicted to be adapted. As expected, the adaptation of the same virus to a randomized HLA profile is significantly lower than to its host HLA profile (one-sided, paired Wilcoxon signed-rank test, p-value = 1.50e-44). Fig. 2 shows the distribution of the estimated adaptation scores for the random data set compared to the chronic_lowCD4 data set.

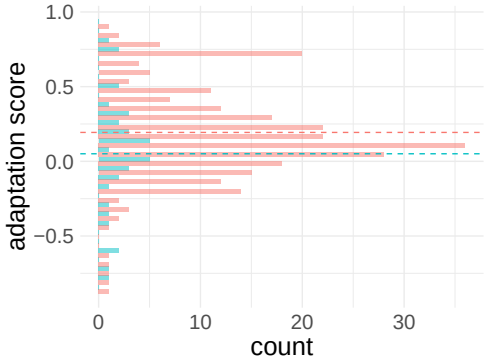


Fig 1. Histogram of the adaptation scores of the chronic_lowCD4 data set (red) and the chronic_highCD4 data set (turquoise). Dashed line represents the mean adaptation score per data set. The mean adaptation score is 0.19 for the chronic_lowCD4 data set and 0.05 for the chronic_highCD4 data set.

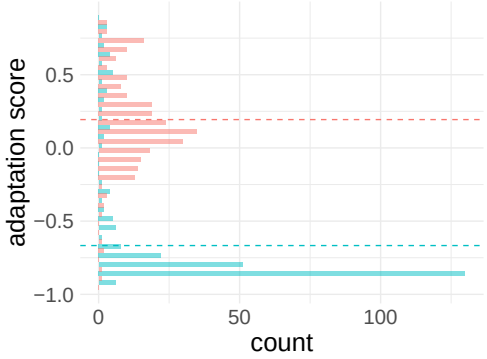


Fig 2. Histogram of the adaptation scores of the chronic_lowCD4 data set (red) and the the averaged adaptation scores of the random data set (turquoise). Dashed line represents the mean adaptation score per data set. The mean adaptation score is 0.19 for the chronic_lowCD4 data set and -0.67 for the random data set.

Autologous viruses more adapted than heterologous viruses

We observed that the adaptation score of the harbored virus to its host (autologous virus) is higher (p-value = 1.48e-31) in contrast to the adaptation of the other viruses in the cohort to the same HLA profile (heterologous virus). This meets our expectation, since we define the adaptation score to reflect how likely the virus acquired escape mutations specific to the host HLA profile. Fig. 3 shows the adaptation scores of the autologous virus and the averaged heterologous viruses for each subject (HLA profile).

Viruses in acute phase less adapted than in chronic phase

We expect that viruses from acutely-infected HIV-1 patients should be less adapted than from chronically-infected HIV-1 patients due to the shorter exposure to the immune system. Fig. 4 shows a histogram of the estimated adaptation scores for the acute and the chronic data sets. Since only the HLA I profile was available for the acute data set, we built an adaptation model based only on the HLA I profile for this purpose. We observed that viral strains from acutely infected patients have significantly lower estimated adaptation scores compared to the chronically-infected HIV-1 patients from

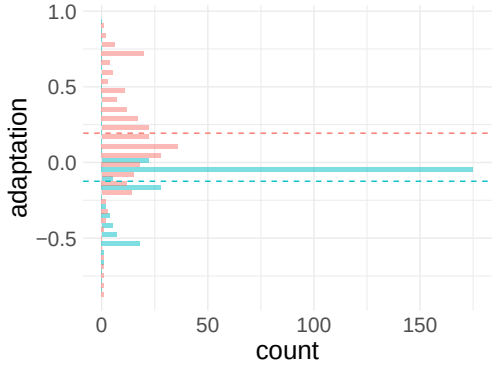


Fig 3. Histogram of estimated adapatation score for each HLA profile and autologous and heterologous viruses. Estimated adaptation scores for each HLA profile and its autologous virus (red) and heterologous viruses of the cohort (turquoise). The adaptation scores of the heterologous viruses are averaged. Dashed line represents the mean adaptation score per data set. The mean adaptation score for autologous viruses is 0.19 and -0.12 for heterologous viruses.

our cohort (one-sided, unpaired Wilcoxon rank-sum test, p-value = 4.17e-5). Note that viruses from acutely-infected patients might also carry HLA-related escape mutations due to transmission.

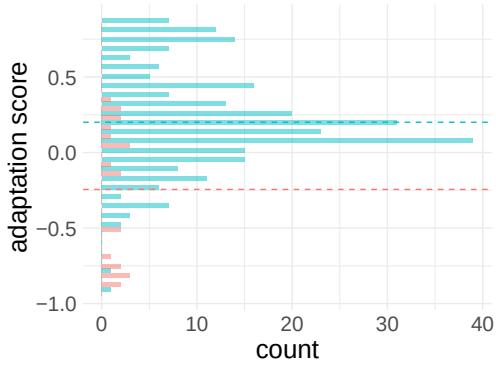


Fig 4. Comparison of HLA adaptation in acutely- and chronically-infected HIV-1 patients. Histogram of the estimated adaptation scores for the chronically-infected data set (turquoise) and the acutely infected data set (red). Dashed line represents the mean adaptation score per data set. The mean adaptation score is 0.20 for the chronic_lowCD4 data set based only on HLA I alleles and -0.24 for the acute data set, respectively.

Validation of the per-site models

Non-informative per-site models have no influence on the adaptation score

In contrast to the overall adaptation score, it is possible to evaluate the performance of the per-site models. This is useful for the interpretation and validation of the model but irrelevant for the quality of the adaptation score. For each frequent site, we compute the likelihood ratio of a model that estimates the likelihood that the site is under HLA pressure (HLA model) and a hypothetical model that assumes no HLA pressure

(baseline). Thereby, the estimated per-site adaptations are directly adjusted by a baseline model and calibrated among all sites. Thus, including sites which are not under HLA pressure will more likely contribute with a factor of 1 to the overall adaptation score and, consequently, have no influence. This allows to take all frequent sites into consideration without any preselection or apriori knowledge. Note, by definition of the adaptation score, the adaptation of each frequent site contributes with the same weight to the overall adaptation score. All per-site models reached the Gelman-Rubin convergence criteria by having an Rhat value less than or equal to 1.

Informative models learn HLA footprints

While it is not the focus of the study, we can identify sites with a likelihood ratio over 1, indicating a potential association between the frequent site and the HLA profile. In the study cohort, we identified 68 frequent sites in the p24 protein. Out of the corresponding 68 per-site HLA models, 21 had an averaged AUC under the precision-recall ROC curve higher than the averaged precision-recall baseline, where the precision-recall baseline is computed as the ratio of positive samples in the data set. Precision-recall was computed for each possible amino acid at a frequent site via 10-fold cross-validation. If models are evaluated by the performance to predict each frequent single amino acid polymorphism (SAP) separately, 52 models out of 210 perform better than the precision-recall baseline. Table 2 shows the top 10 polymorphisms with precision-recall AUC exceeding the baseline. Further analyzing the learned coefficients of the per-site models with high performance revealed that the models learned known footprints for subtype C such as the association between the T242N mutation and the HLA alleles HLA-B*57:01/02/03 or HLA-B*58:01 as well as the T186S escape mutation associated with HLA-B*81:01 [67–69].

Table 2. Top ten potential HLA-restricted sites and single amino acid polymorphisms (SAPs) with respect to precision-recall baseline performance The performance of the HLA model at a specific site and for a specific SAP is computed as the AUC under the precision-recall curve (PRROC).

site	SAP	PRROC	baseline
242	n	0.81	0.12
186	s	0.45	0.04
163	g	0.26	0.06
309	c	0.20	0.01
146	p	0.38	0.24
357	g	0.65	0.53
242	t	0.98	0.87
146	s	0.21	0.11
312	e	0.39	0.30
230	d	0.14	0.05

Interpretable adaptation score by providing logos for each virus

For each frequent variant site, an odds ratio above or below 1 (with an offset of 0.1) indicates whether the amino acid at this site is more likely under HLA pressure or not. This information can be used to compute a logo revealing the amino acids that contributed the most to the adaptation score. This information helps the user to understand the results for different inputs. Fig. 5 shows the logo for the patient with the highest adaptation score in the cohort. The known HLA escape mutation T186S [70] has the highest contribution to the predicted adaptation score.

Fig 5. Logo for the patient with the highest adaptation score. The logo shows the viral polymorphisms that have the highest contribution to the adaptation score of this patient. Blue capital letters indicate adapted amino acids, while orange lowercase letters reflect non-adapted amino acids. The height of the letters reflects the contribution to the adaptation score and is scaled by the maximum contribution. The x-axis denotes the corresponding sites in the HXB2 virus.

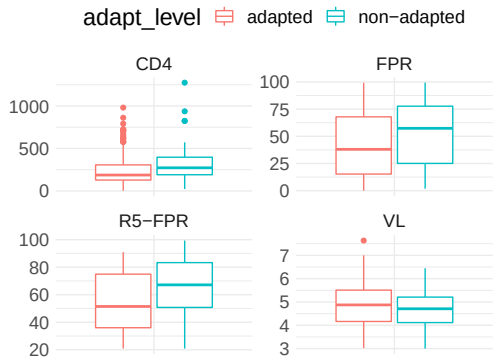


Fig 6. Difference in clinical variables based on HLA adaption. Measurement of CD4 $^+$ T cell (CD4), logarithmized viral load (VL), FPR, and the FPR of R5 viruses matched based on their CD4 count (R5-FPR) stratified among adapted (red) and non-adapted(turquoise) viruses.

HLA adaptation associated with CD4$^+$ T cell count but not viral load

We analyzed the estimated adaptation score with respect to viral load, CD4$^+$ T cell count and coreceptor usage. On the one hand, we tested whether patients with adapted and non-adapted viruses differ in these variables, where adapted is defined as an adaptation score > 0.1 and non-adapted as an adaptation score < -0.1, based on the expected variance of 0.1 (see Fig. 6). On the other hand, we analyzed whether viruses of patients with different known levels of these variables differ in their adaptation (see Fig. 7).

Though HLA class-I restricted polymorphism are known to be predictive for viral load and CD4$^+$ T cell count in general [26, 71], we observed only a correlation between the estimated adaptation scores (based on HLA I and HLA II alleles) and the CD4$^+$ count (Pearson correlation coefficient -0.16, p-value=0.02) but not with viral load (0.04, p-value=0.88). Note, however, that the study cohort consists of rather chronically-infected patients at a later stage of infection, where other factors more likely affect fluctuations in the viral load than the HLA adaptation, and a difference between controllers and non-controllers, for example, is not expected to be seen as in the beginning of the infection. We also observed that adapted viruses do not have

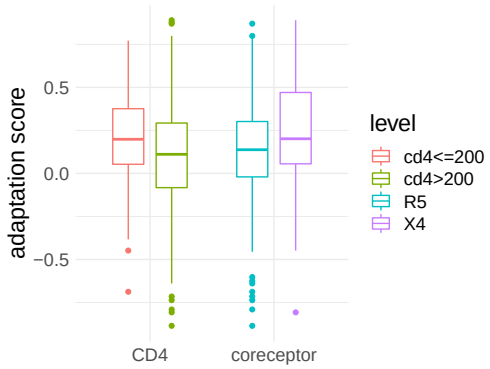


Fig 7. Adaptation score for different levels of CD4$^+$ T cell count (CD4) and coreceptor usage (coreceptor)

Table 3. Averaged and maximum adaptation score stratified on the coreceptor usage and two data sets.

data set	coreceptor	adapt(x)	max(adap(x))
chronic_highCD4	R5	0.03±0.40	0.80
chronic_highCD4	X4	0.14±0.28	0.51
chronic_lowCD4	R5	0.17±0.30	0.87
chronic_lowCD4	X4	0.25±0.33	0.89

statistically significant higher viral loads than non-adapted viruses (one-sided, unpaired Wilcoxon rank-sum test, p-value = 1.86e-1), and that patients with low viral load have not less adapted viruses (one-sided, unpaired Wilcoxon rank-sum test, p-value=8.54e-2). In addition to the significant correlation between the CD4$^+$ T cell count and adaptation score, we observed that patients with AIDS (CD4$^+$ T cell count < 200) have more adapted viruses than patients with higher CD4$^+$ T cell counts (one-sided, unpaired, Wilcoxon rank-sum test, p-value = 3.20e-3). CD4$^+$ T cell count was also lower in patients with adapted viruses compared to non-adapted (Wilcoxon rank-sum test, p-value = 1.27e-3).

Adaptation associated with coreceptor usage

Using our adaptation score, we investigated the relationship between HLA adaptation and coreceptor usage. More precisely, we analyzed the hypothesis that high HLA adaptation might trigger the coreceptor switch in a similar way as a weak immune system (measured by a low number of CD4$^+$ T cell counts). Coreceptor usage was determined with the false positive rate (FPR) of the coreceptor prediction tool geno2pheno[coreceptor] [17]. The provided FPR corresponds to the confidence with which the sequence is classified as X4-capable. The higher the FPR, the more likely the sequence is not X4-capable, but R5-capable. Table 3 shows the average adaptation scores stratified for coreceptor usage. We observed a negative correlation between estimated adaptation score and corresponding FPR (Pearson correlation coefficient of -0.15, p-value = 0.03). This means that the more adapted the virus, the higher the likelihood that the virus is classified as X4-capable. This was further confirmed by the observation that X4-capable viruses are more adapted compared to R5 viruses (Wilcoxon rank-sum test, p-value = 1.34e-2, see Fig. 7) and that, in general, adapted viruses have a lower FPR (rather X4 variants) compared to non-adapted viruses (Wilcoxon rank-sum test, p-value = 6.76e-3, see Fig. 6). Note, since the variants are already determined as $X4$-capable, it is impossible to show if the emergence of

X4-variants is driven by HLA adaptation. This analysis would require longitudinal data where the emergence of the coreceptor switch is captured. To rule out the possibility that higher adaptation of the X4 variants occurs due to longer exposure to the host immune system in contrast to R5 variants, the exact duration of infection is required. However, we observed that even among all R5 viruses, higher adaptation is associated with lower FPR, indicating that more adapted R5 samples might be closer to the coreceptor switch compared to non-adapted samples (one-sided paired Wilcoxon signed-rank test, p-value = 2.21e-2). Since the $CD4^+$ T cell count is a major confounder for the coreceptor usage, we have matched for this test adapted and non-adapted R5 samples with similar $CD4^+$ count ($\pm$ 50 cells/mm^3). Note, high adaptation of an R5 variant in a chronically-infected patient can also occur due to the long exposure to the immune system, since a coreceptor switch is only observed in 50% of the patients.

Conclusion

Here, we introduced a novel computational approach to jointly estimate HLA I and HLA II adaptation of HIV-1 using Bayesian generalized linear mixed models. In addition, we presented a new study cohort of 312 treatment-naive, subtype C, chronically-infected HIV-1 patients from South Africa, where we sequenced the viral *gag* (and *env*) protein with corresponding HLA class I and II alleles for the training of our models. Apart from validating that our adaptation score inherited appropriate characteristics, we showed that the models underlying the adaptation score are biologically meaningful by learning well-known HLA-restricted polymorphisms. Using our approach and the data, we investigated the relationship between HLA adaptation and coreceptor usage of HIV-1, which had been unexplored up to now. We observed that X4-capable viruses are more adapted compared to R5-capable viruses (Wilcoxon rank-sum test, p-value =1.34e-2). Moreover, even among all R5-capable viruses, higher adaptation is associated with lower FPR, indicating that more adapted R5 variants might be closer to the coreceptor switch compared to non-adapted variants (Wilcoxon signed-rank test, p-value = 2.21e-2). Thus, HLA adaptation might be another factor that should be considered prior to the administration of CCR5 coreceptor antagonists. It might also be useful in predicting how imminent the coreceptor switch is.

In general, the estimated adaptation score allows to measure and understand HIV-1's adaptation to the immune system. The adaptation score can be used to guide the design of suitable immunogens as vaccine targets by selecting sites that are more likely to be non-adapted to the immune system. Since the approach itself is not HIV-1 specific, the presented method can be also applied to study any virus-host adaptation. We encourage the usage of Bayesian GLMMs for modeling virus-host adaptation due to their ability to adjust for phylogenetic dependencies in the data and to handle highly overparameterized settings within the model. In light of current and potential future viral threats to mankind, such as SARS-CoV-2 or Ebola or MERS-CoV, this flexible, data non-intensive method can be useful to reveal and analyze virus-host dynamics of new viruses where little data is available.

Future studies of the study cohort are required to further evaluate how the adaptation score is coherent with CTL escape experiments. While the study cohort was appropriate to learn HLA adaptation, it only allows to study the coreceptor switch and the role of HLA adaptation on it from a retrospective angle. Moreover, the number of CXCR4-using variants with intermediate to high $CD4^+$ T cell count was very low. Consequently, a study cohort with longitudinal data on coreceptor usage and intermediate to high $CD4^+$ T cell count would give additional insights. This would not only allow to investigate if HLA II adaptation occurs prior to the coreceptor switch, but also if the degree of adaptation is associated with the time until the coreceptor switch

occurs.

Note that the presented adaptation score here is a simple approach that can be easily optimized and extended in different ways. Given the available data, we restricted the analysis to subtype C infected patients and the p24 protein. However, our approach can also be applied to other subtypes and/or combined over different viral proteins. Further, we used the viral consensus sequences instead of the NGS sequences, since we aimed at predicting the adaptation per sequence. Still, the within-subject relatedness of different reads per virus could be easily incorporated into the Bayesian models. A larger data set might improve the current adaptation score by better representing the population with respect to the HLA repertoire and potential frequent variant sites, resulting in more informative per-site models.

Furthermore, it is possible to make the proposed models more complex by incorporating more dependency structures such as HLA linkage, or by relaxing the assumption of independence among all sites to capture compensatory mutations. Another assumption is that each frequent variant site has the same probability to be under HLA pressure. Prior knowledge about common HLA epitopes can be added to the model by weighting the per-site likelihood odds ratios accordingly. However, if a site is more likely to be under HLA pressure, given by the underlying data, by construction of the adaptation score, the likelihood odds ratio should contribute with a higher factor to the overall adaptation score.

Decomposing the adaptation score based on the potential adaptation of each frequent variant site is very advantageous with respect to model explainability and to settings, where little prior information exists. However, it requires the computation of two models per frequent variant sites, leading to a high number of models. Currently, the per-site models are not optimized with regard to parameter and predictor selection. To avoid overfitting and p-value based selection, we forced each model to capture our prior beliefs that the model should be unbiased with regard to sex, age, ethnicity and not be hampered by phylogenetic relatedness. While we ensured that the models are all converging according to the Rubin-Gelman criterion, we do not perform visual checks of the Bayesian GLMMs, such as prior and posterior predictive checks. Though this is a standard procedure for Bayesian GLMMs, it was not feasible in our case due to the high number of models. In our case, this was also not mandatory. Setting the horseshoe prior for the beta coefficients was based on our apriori knowledge that only few HLA alleles and clinical variables should have influence on a site. Setting potential non-optimal parameters might lead to non-informative per-site models. While we might lose some potential information for these sites, the quality of the overall adaptation score remains guaranteed by calibrating the per-site models with a baseline model. For computational reasons, it might be also beneficial to reduce the computation of the adaptation score based on only the informative per-site models.

Funding

This study was funded by the annual donation in 2016 of the Supporting Members of the Max Planck Society for the project 'How is the immune system tricked by the HI-virus?'. N. P. was supported by infrastructural funding from the Deutsche Forschungsgemeinschaft (DFG), Cluster of Excellence EXC 2124 Controlling Microbes to Fight Infections and the DFG Cluster of Excellence 'Machine Learning—New Perspectives for Science' (EXC 2064/1, project no. 390727645). Sample collection was funded by the Bill & Melinda Gates Foundation (Grant 38580 and *OPP*38580_01) for the project "Global HIV Vaccine Research Cryorepository – GHRC".

References

1. on HIV/AIDS JUNP, et al. Fast-track: ending the AIDS epidemic by 2030. Geneva: UNAIDS. 2014;.

2. Lusso P. HIV and the chemokine system: 10 years later. The EMBO Journal. 2006;25(3):447–456. doi:10.1038/sj.emboj.7600947.

3. Berger EA, Doms R, Fenyö EM, Korber B, Littman D, Moore J, et al. A new classification for HIV-1. Nature. 1998;391(6664):240–240.

4. Moore JP, Kitchen SG, Pugach P, Zack JA. The CCR5 and CXCR4 Coreceptors—Central to Understanding the Transmission and Pathogenesis of Human Immunodeficiency Virus Type 1 Infection. AIDS Research and Human Retroviruses. 2004;20(1):111–126. doi:10.1089/088922204322749567.

5. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proceedings of the National Academy of Sciences. 2008;105(21):7552–7557. doi:10.1073/pnas.0802203105.

6. Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR. Change in Coreceptor Use Correlates with Disease Progression in HIV-1–Infected Individuals. Journal of Experimental Medicine. 1997;185(4):621–628. doi:10.1084/jem.185.4.621.

7. Regoes RR, Bonhoeffer S. The HIV coreceptor switch: a population dynamical perspective. Trends in microbiology. 2005;13(6):269–277.

8. Schuitemaker H, Koot M, Kootstra NA, Dercksen MW, de Goede RE, van Steenwijk RP, et al. Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population. Journal of Virology. 1992;66(3):1354–1360.

9. Tscherning C, Alaeus A, Fredriksson R, Åsa Björndal, Deng H, Littman DR, et al. Differences in Chemokine Coreceptor Usage between Genetic Subtypes of HIV-1. Virology. 1998;241(2):181–188. doi:https://doi.org/10.1006/viro.1997.8980.

10. Connell BJ, Michler K, Capovilla A, Venter WDF, Stevens WS, Papathanasopoulos MA. Emergence of X4 usage among HIV-1 subtype C: evidence for an evolving epidemic in South Africa. AIDS. 2008;22:896–899.

11. Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, et al. Development and Characterization of a Novel Single-Cycle Recombinant-Virus Assay To Determine Human Immunodeficiency Virus Type 1 Coreceptor Tropism. Antimicrobial Agents and Chemotherapy. 2006;51(2):566–575. doi:10.1128/aac.00853-06.

12. Low AJ, McGovern RA, Harrigan PR. Trofile HIV co-receptor usage assay. Expert opinion on medical diagnostics. 2009;3(2):181–191.

13. Reeves J, Coakley E, Petropoulos C, Whitcomb J. An enhanced sensitivity Trofile HIV coreceptor tropism assay for selecting patients for therapy with entry inhibitors targeting CCR5: a review of analytical and clinical studies. J Viral Entry. 2009;3(3):94–102.

14. Gonzalez-Serna A, Leal M, Genebat M, Abad MA, Garcia-Perganeda A, Ferrando-Martinez S, et al. TROCAI (Tropism Coreceptor Assay Information): a New Phenotypic Tropism Test and Its Correlation with Trofile Enhanced Sensitivity and Genotypic Approaches. Journal of Clinical Microbiology. 2010;48(12):4453–4458. doi:10.1128/JCM.00953-10.

15. Fouchier R, Groenink M, Kootstra NA, Tersmette M, Huisman H, Miedema F, et al. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. Journal of virology. 1992;66(5):3183–3187.

16. Jensen MA, Li FS, van't Wout AB, Nickle DC, Shriner D, He HX, et al. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. Journal of virology. 2003;77(24):13376–13388.

17. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. Nature biotechnology. 2007;25(12):1407–1410.

18. Pfeifer N, Lengauer T. Improving HIV coreceptor usage prediction in the clinic using hints from next-generation sequencing data. Bioinformatics. 2012;28(18):i589–i595. doi:10.1093/bioinformatics/bts373.

19. Cashin K, Gray LR, Jakobsen MR, Sterjovski J, Churchill MJ, Gorry PR. CoRSeqV3-C: a novel HIV-1 subtype C specific V3 sequence based coreceptor usage prediction algorithm. Retrovirology. 2013;10(1):24. doi:10.1186/1742-4690-10-24.

20. Sander O, Sing T, Sommer I, Low AJ, Cheung PK, Harrigan PR, et al. Structural descriptors of gp120 V3 loop for the prediction of HIV-1 coreceptor usage. PLoS computational biology. 2007;3(3).

21. Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues FS, et al. Predicting HIV coreceptor usage on the basis of genetic and clinical covariates. Antiviral therapy. 2007;12(7):1097.

22. Swenson LC, Mo T, Dong WW, Zhong X, Woods CK, Jensen MA, et al. Deep sequencing to infer HIV-1 co-receptor usage: application to three clinical trials of maraviroc in treatment-experienced patients. Journal of Infectious Diseases. 2011;203(2):237–245.

23. Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J, Paxton WA. N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. Journal of Biological Chemistry. 2001;276(16):13433–13441.

24. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, Ogunlesi AO, et al. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature. 1991;354(6353):453–459.

25. Goulder PJR, Watkins DI. HIV and SIV CTL escape: implications for vaccine design. Nature Reviews Immunology. 2004;4(8):630–640. doi:10.1038/nri1417.

26. The International HIV Controllers Study. The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation. Science. 2010;330(6010):1551–1557. doi:10.1126/science.1195271.

27. Rousseau CM, Daniels MG, Carlson JM, Kadie C, Crawford H, Prendergast A, et al. HLA Class I-Driven Evolution of Human Immunodeficiency Virus Type 1 Subtype C Proteome: Immune Escape and Viral Load. Journal of Virology. 2008;82(13):6434–6446. doi:10.1128/jvi.02455-07.

28. Carlson JM, Brumme ZL, Rousseau CM, Brumme CJ, Matthews P, Kadie C, et al. Phylogenetic Dependency Networks: Inferring Patterns of CTL Escape and Codon Covariation in HIV-1 Gag. PLoS Computational Biology. 2008;4(11):e1000225. doi:10.1371/journal.pcbi.1000225.

29. Moore C, John M, James I, Christiansen F, Witt C, Mallal S. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. Science (New York, NY). 2002;296:1439–43.

30. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, et al. A Whole-Genome Association Study of Major Determinants for Host Control of HIV-1. Science. 2007;317(5840):944–947. doi:10.1126/science.1143767.

31. Carlson JM, Brumme CJ, Martin E, Listgarten J, Brockman MA, Le AQ, et al. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. Journal of virology. 2012;86(24):13202—13216. doi:10.1128/JVI.01998-12.

32. Migueles SA, Sabbaghian MS, Shupert WL, Bettinotti MP, Marincola FM, Martino L, et al. HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. Proceedings of the National Academy of Sciences. 2000;97(6):2709–2714. doi:10.1073/pnas.050567397.

33. Altfeld M, Addo MM, Rosenberg ES, Hecht FM, Lee PK, Vogel M, et al. Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. AIDS. 2003;17(18).

34. Carlson JM, Du VY, Pfeifer N, Bansal A, Tan VYF, Power K, et al. Impact of pre-adapted HIV transmission. Nature Medicine. 2016;22(6):606–613. doi:10.1038/nm.4100.

35. Erdmann N, Du VY, Carlson J, Schaefer M, Jureka A, Sterrett S, et al. HLA Class-II Associated HIV Polymorphisms Predict Escape from CD4+ T Cell Responses. PLOS Pathogens. 2015;11(8):e1005111. doi:10.1371/journal.ppat.1005111.

36. Bhattacharya T, Daniels M, Heckerman D, Foley B, Frahm N, Kadie C, et al. Founder Effects in the Assessment of HIV Polymorphisms and HLA Allele Associations. Science. 2007;315(5818):1583–1586. doi:10.1126/science.1131528.

37. Troyer RM, McNevin J, Liu Y, Zhang SC, Krizan RW, Abraha A, et al. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. PLoS pathogens. 2009;5(4).

38. Payne R, Branch S, Kløverpris H, Matthews P, Koofhethile C, Strong T, et al. Differential escape patterns within the dominant HLA-B* 57: 03-restricted HIV Gag epitope reflect distinct clade-specific functional constraints. Journal of virology. 2014;88(9):4668–4678.

39. Prince JL, Claiborne DT, Carlson JM, Schaefer M, Yu T, Lahki S, et al. Role of transmitted Gag CTL polymorphisms in defining replicative capacity and early HIV-1 pathogenesis. PLoS pathogens. 2012;8(11).

40. Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, Moodley E, et al. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nature medicine. 2007;13(1):46–53.

41. Edwards BH, Bansal A, Sabbaj S, Bakari J, Mulligan MJ, Goepfert PA. Magnitude of Functional CD8+ T-Cell Responses to the Gag Protein of Human Immunodeficiency Virus Type 1 Correlates Inversely with Viral Load in Plasma. Journal of Virology. 2002;76(5):2298–2305. doi:10.1128/jvi.76.5.2298-2305.2002.

42. Geldmacher C, Currier JR, Herrmann E, Haule A, Kuta E, McCutchan F, et al. CD8 T-Cell Recognition of Multiple Epitopes within Specific Gag Regions Is Associated with Maintenance of a Low Steady-State Viremia in Human Immunodeficiency Virus Type 1-Seropositive Patients. Journal of Virology. 2006;81(5):2440–2448. doi:10.1128/jvi.01847-06.

43. Zuniga R, Lucchetti A, Galvan P, Sanchez S, Sanchez C, Hernandez A, et al. Relative Dominance of Gag p24-Specific Cytotoxic T Lymphocytes Is Associated with Human Immunodeficiency Virus Control. Journal of Virology. 2006;80(6):3122–3125. doi:10.1128/jvi.80.6.3122-3125.2006.

44. Kløverpris HN, Leslie A, Goulder P. Role of HLA Adaptation in HIV Evolution. Frontiers in Immunology. 2016;6. doi:10.3389/fimmu.2015.00665.

45. Carvalho CM, Polson NG, Scott JG. Handling sparsity via the horseshoe. Journal of Machine Learning Research W&CP. 2009;.

46. Struck D, Lawyer G, Ternes AM, Schmit JC, Bercoff DP. COMET: adaptive context-based modeling for ultrafast HIV-1 subtype identification. Nucleic acids research. 2014;42(18):e144. doi:10.1093/nar/gku739.

47. Porter D, Daeumer M, Thielen A, Chang S, Martin R, Cohen C, et al. Emergent HIV-1 Drug Resistance Mutations Were Not Present at Low-Frequency at Baseline in Non-Nucleoside Reverse Transcriptase Inhibitor-Treated Subjects in the STaR Study. Viruses. 2015;7(12):6360–6370. doi:10.3390/v7122943.

48. Vandekerckhove L, Wensing A, Kaiser R, Brun-Vézinet F, Clotet B, De Luca A, et al. European guidelines on the clinical management of HIV-1 tropism testing. The Lancet infectious diseases. 2011;11(5):394–407.

49. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution. 2013;30(4):772–80. doi:10.1093/molbev/mst010.

50. Nelder JA, Wedderburn RW. Generalized linear models. Journal of the Royal Statistical Society: Series A (General). 1972;135(3):370–384.

51. McCullagh P. Generalized linear models. Routledge; 2018.

52. Agresti A. An Introduction to Categorical Data Analysis. John Wiley & Sons, Inc.; 2007. Available from: `https://doi.org/10.1002/0470114754`.

53. Fong Y, Rue H, Wakefield J. Bayesian inference for generalized linear mixed models. Biostatistics. 2009;11(3):397–412. doi:10.1093/biostatistics/kxp053.

54. Fox J, Weisberg S. An R Companion to Applied Regression. 3rd ed. Thousand Oaks CA: Sage; 2019. Available from: `http://tinyurl.com/carbook`.

55. R Core Team. R: A Language and Environment for Statistical Computing; 2017. Available from: `https://www.R-project.org/`.

56. Bürkner PC. brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, Articles. 2017;80(1):1–28. doi:10.18637/jss.v080.i01.

57. Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, et al. Stan: A Probabilistic Programming Language. Journal of Statistical Software, Articles. 2017;76(1):1–32. doi:10.18637/jss.v076.i01.

58. Duane S, Kennedy AD, Pendleton BJ, Roweth D. Hybrid Monte Carlo. Physics Letters B. 1987;195(2):216–222. doi:10.1016/0370-2693(87)91197-x.

59. Homan MD, Gelman A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. J Mach Learn Res. 2014;15(1):1593–1623.

60. Döring M, Büch J, Friedrich G, Pironti A, Kalaghatgi P, Knops E, et al. geno2pheno [ngs-freq]: a genotypic interpretation system for identifying viral drug resistance using next-generation sequencing data. Nucleic Acids Research. 2018;.

61. Huber M, Metzner KJ, Geissberger FD, Shah C, Leemann C, Klimkait T, et al. MinVar: A rapid and versatile tool for HIV-1 drug resistance genotyping by deep sequencing. Journal of Virological Methods. 2017;240:7–13. doi:https://doi.org/10.1016/j.jviromet.2016.11.008.

62. Hadfield JD, Nakagawa S. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. Journal of Evolutionary Biology. 2010;23(3):494–508. doi:10.1111/j.1420-9101.2009.01915.x.

63. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312–1313. doi:10.1093/bioinformatics/btu033.

64. Schneider TD, Stephens RM. Sequence Logos: A New Way to Display Consensus Sequences. Nucleic Acids Res. 1990;18:6097–6100.

65. Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-Tinch CH, Sochat VV, et al. Sustainable data analysis with Snakemake. F1000Research. 2021;10.

66. Anaconda Software Distribution; 2020. Available from: `https://docs.anaconda.com/`.

67. Wright JK, Naidoo VL, Brumme ZL, Prince JL, Claiborne DT, Goulder PJR, et al. Impact of HLA-B*81-Associated Mutations in HIV-1 Gag on Viral Replication Capacity. Journal of Virology. 2012;86(6):3193–3199. doi:10.1128/jvi.06682-11.

68. Leslie A, Matthews PC, Listgarten J, Carlson JM, Kadie C, Ndung'u T, et al. Additive Contribution of HLA Class I Alleles in the Immune Control of HIV-1 Infection. Journal of Virology. 2010;84(19):9879–9888. doi:10.1128/jvi.00320-10.

69. Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, Chetty S, et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature. 2004;432(7018):769–775.

70. Kløverpris HN, Leslie A, Goulder P. Role of HLA Adaptation in HIV Evolution. Frontiers in Immunology. 2016;6. doi:10.3389/fimmu.2015.00665.

71. Carrington M, O'Brien SJ. The Influence of HLA Genotype on AIDS. Annual Review of Medicine. 2003;54(1):535–551. doi:10.1146/annurev.med.54.101601.152346.

B. Glossaries

List of abbreviations

AI	artificial intelligence
AIC	Akaike information criterion
AIDS	acquired immunodeficiency syndrome
APC	antigen presenting cell
ATI	analytical treatment interruption
AUC	area under the receiver operating characteristic curve
BIC	Bayesian information criterion
bNAb	broadly neutralizing antibody
CCR5	C-C chemokine receptor type 5
CD4	cluster of differentiation 4
CTL	cytotoxic T cells
CXCR4	C-X-C chemokine receptor type 4
DNA	deoxyribonucleic acid
DP	differential privacy
Env	glycoprotein gp160
Fc	crystallizable fragment
FDR	false discovery rate
FHE	fully homomorphic encryption
FL	federated learning
FN	false negatives
FP	false positives
FPR	false positive rate
GBM	gradient boosting machine
HIV-1	human immunodeficiency virus type 1
HLA	human leukocyte antigen
IC50	half maximum inhibitory concentration
Ig	immunoglobulin
MHC	major histocompatibility complex

NGS	next generation sequencing
PNGS	potential N-linked glycosylation sites
PR-AUC	area under the precision-recall curve
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
scFv	single-chain variable fragment
SMPC	secure multi-party computation
SVM	support vector machine
TP	true positives
TPR	true positive rate

Glossary

HLA footprints

HLA-restricted viral escape mutations towards the T-cell mediated immune response.

accuracy

Evaluation measure for predictions defined as

$$acc(y_{true}, y_{pred}) = \frac{(TP + TN)}{(TP + FP + TN + FN)} = \frac{TP + TN}{N},$$

where y_{true} refers to the ground truth labels, y_{pred} refers to the predicted labels, TP to the true positive predictions, TN to the true negative predictions, FP to the false positive predictions, FN to the false negatives predictions, and N stands for the total amount of samples.

adaptive immune system

Second line of defensive mechanism by the human immune system, which is initiated if the innate immune system is not able to handle the pathogen. The adaptative immune system produces a pathogen-specific immune response and is able to maintain a memory of the pathogen.

antiretroviral

Effective against retroviruses.

area under the receiver operating characteristic curve

The area under the curve of TPR and FPR for varying decision thresholds.

false discovery rate

Measure of the expected proportion of discoveries that are false with q controlling the number of false discoveries.

$$FDR = E(q) = E\left(\frac{FP}{TP + FP}\right)$$

false positive rate

The false positive rate describes the ratio of false positives to the total number of negative samples as calculated by the following formula:

$$FPR = \frac{FP}{FP + TN}$$

fitness

In evolution theory, this term refers to surviving. In particular, it refers to the reproducing capacity of an organism. The fitness of a genotype depends on the environment and can confer an advantage or disadvantage in comparison to other genotypes. A genotype conferring an advantage will likely increase in frequency within the (quasispecies) population.

heterodimer

a protein composed of two non-identical monomers (polypetide chains)

HLA supertype

Classification of HLA alleles into groups with shared peptide binding specificities.

innate immune system

First line of defensive mechanisms by the human immune system, which is initiated fast but is non-specific to the pathogen.

linkage disequilibrium

Non-random association between different genetic loci.

mutation

A change in the genetic sequence compared to the reference genome.

polymorphism

The existence of two or more forms of a genotype.

precision

Precision describes the positive predictive value, i.e. the ratio of true positive predicted samples over all positive predicted samples:

$$precision = \frac{TP}{TP + FP}$$

pseudovirus

Artificially created virus, where a non-infective virion lacking the Env protein is combined with a foreign viral envelope protein capable of a single round of infection.

recall

Recall is another term for true positive rate.

recombination

Viral genetic recombination may occur in retroviruses during the reverse transcriptation step if a cell is coinfected with two genetically distant variants. The newly formed virion contains afterwards a mixture of the the genetic information of the two distinct variants. This process is also termed *viral sex*.

selection pressure

Factors in the environment of a pathogen like the host immune system response or drugs provide some characteristic traits of the pathogen an advantage or disadvantage with respect to survival fitness - the ability to reproduce.

tier

A classification system of HIV-1 isolates in neutralization assays based on their phentotypical neutralization sensitivity, frequency, and conformational state of the envelope trimer [189, 190]. Tier 1a and 1b refers to a very small fraction of viruses which are very sensitive to antibodies - mostly in open or intermediate trimer conformation. Tier 2 defines the most prevalent phenotype of HIV-1 isolates with rather closed trimer conformation and modest neutralization. Tier 3 refers to the least sensitive HIV-1 isolates having the highest frequency of a closed Env conformation.

true positive rate

The true positive rate describes the ratio of true positives to the total number of positive samples as calculated by the following formula

$$TPR = \frac{TP}{TP + FN}$$

turnover rate

The yield of virions per replication time.

viremic

Existence of detectable levels of the virus in the blood.

C. References

References

[1] A. **Hake** and N. Pfeifer. „Prediction of HIV-1 sensitivity to broadly neutralizing antibodies shows a trend towards resistance over time.“ In: *PLOS Computational Biology* 13.10 (Oct. 2017), pp. 1–23. DOI: 10.1371/journal.pcbi.1005789. URL: https://doi.org/10.1371/journal.pcbi.1005789.

[2] T. Schoofs, F. Klein, M. Braunschweig, E. F. Kreider, A. **Feldmann**, L. Nogueira, T. Oliveira, J. C. C. Lorenzi, E. H. Parrish, G. H. Learn, A. P. West, P. J. Bjorkman, S. J. Schlesinger, M. S. Seaman, J. Czartoski, M. J. McElrath, N. Pfeifer, B. H. Hahn, M. Caskey, and M. C. Nussenzweig. „HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1.“ In: *Science* 352.6288 (2016), pp. 997–1001. DOI: 10.1126/science.aaf0972. eprint: https://www.science.org/doi/pdf/10.1126/science.aaf0972. URL: https://www.science.org/doi/abs/10.1126/science.aaf0972.

[3] J. Scheid, J. Horwitz, Y. Bar-On, E. Kreider, C.-L. Lu, J. C. Cetrulo Lorenzi, A. **Feldmann**, M. Braunschweig, L. Nogueira, T. Oliveira, I. Shimeliovich, R. Patel, L. Burke, Y. Cohen, S. Hadrigan, A. Settler, M. Witmer-Pack, J. West, B. Juelg, and M. Caskey. „HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption.“ In: *Nature* 535 (June 2016). DOI: 10.1038/nature18929.

[4] A. **Hake**, A. Germann, C. de Beer, A. Thielen, M. Däumer, W. Preiser, H. von Briesen, and N. Pfeifer. „Insights to HIV-1 coreceptor usage by estimating HLA adaptation with Bayesian generalized linear mixed models.“ In: *bioRxiv* (2022). DOI: 10.1101/2022.07.06.498925. eprint: https://www.biorxiv.org/content/early/2022/07/06/2022.07.06.498925.full.pdf. URL: https://www.biorxiv.org/content/early/2022/07/06/2022.07.06.498925.

[5] J. M. Cuevas, R. Geller, R. Garijo, J. López-Aldeguer, and R. Sanjuán. „Extremely High Mutation Rate of HIV-1 In Vivo.“ In: *PLOS Biology* 13.9 (Sept. 2015), pp. 1–19. DOI: 10.1371/journal.pbio.1002251. URL: https://doi.org/10.1371/journal.pbio.1002251.

[6] D. D. Ho, A. U. Neumann, A. S. Perelson, W. H. Chen, J. M. Leonard, and M. Markowitz. „Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection.“ In: *Nature* 373 (1995), pp. 123–126.

[7] W.-S. Hu and H. M. Temin. „Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination.“ In: *Proceedings of the National Academy of Sciences* 87.4 (1990), pp. 1556–1560.

[8] C. Jiang, X. Lian, C. Gao, X. Sun, K. B. Einkauf, J. M. Chevalier, S. M. Chen, S. Hua, B. Rhee, K. Chang, et al. „Distinct viral reservoirs in individuals with spontaneous control of HIV-1.“ In: *Nature* 585.7824 (2020), pp. 261–267.

[9] „A Possible Sterilizing Cure of HIV-1 Infection Without Stem Cell Transplantation.“ In: *Annals of Internal Medicine* 175.1 (2022). PMID: 34781719, pp. 95–100. DOI: 10.7326/L21-0297. eprint: https://doi.org/10.7326/L21-0297. URL: https://doi.org/10.7326/L21-0297.

[10] J. Hemelaar et al. „Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis.“ In: *The Lancet Infectious Diseases* 19.2 (2019), pp. 143–155. ISSN: 1473-3099. DOI: https://doi.org/10.1016/S1473-3099(18)30647-9. URL: https://www.sciencedirect.com/science/article/pii/S1473309918306479.

[11] H. W. Sheppard and M. S. Ascher. „The natural history and pathogenesis of HIV infection.“ In: *Annual Review of Microbiology* 46.1 (1992). PMID: 1444266, pp. 533–564. DOI: 10.1146/annurev.mi.46.100192.002533. eprint: https://doi.org/10.1146/annurev.mi.46.100192.002533. URL: https://doi.org/10.1146/annurev.mi.46.100192.002533.

[12] UNAIDS. *Global HIV & AIDS statistics — Fact sheet*. 2021. URL: https://www.unaids.org/en/resources/fact-sheet.

[13] J. U. N. P. on HIV/AIDS (UNAIDS) et al. *United Nations Political Declaration on Ending AIDS sets world on the Fast-Track to end the epidemic by 2030. 2016*. 2016.

[14] B. S. Taylor, M. E. Sobieszczyk, F. E. McCutchan, and S. M. Hammer. „The challenge of HIV-1 subtype diversity.“ In: *New England Journal of Medicine* 358.15 (2008), pp. 1590–1602.

[15] K. T. Arrildt, S. B. Joseph, and R. Swanstrom. „The HIV-1 env protein: a coat of many colors.“ In: *Current Hiv/Aids Reports* 9 (2012), pp. 52–63.

[16] R. Pantophlet and D. R. Burton. „GP120: target for neutralizing HIV-1 antibodies.“ In: *Annual review of immunology* 24.1 (2006), pp. 739–769.

[17] D. R. Burton, R. L. Stanfield, and I. A. Wilson. „Antibody vs. HIV in a clash of evolutionary titans.“ In: *Proceedings of the National Academy of Sciences* 102.42 (2005), pp. 14943–14948.

[18] P. D. Kwong, M. L. Doyle, D. J. Casper, C. Cicala, S. A. Leavitt, S. Majeed, T. D. Steenbeke, M. Venturi, I. Chaiken, M. Fung, et al. „HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites.“ In: *Nature* 420.6916 (2002), pp. 678–682.

[19] M. Casadellà, A. Cozzi-Lepri, A. Phillips, M. Noguera-Julian, M. Bickel, D. Sedlacek, K. Zilmer, B. Clotet, J. D. Lundgren, R. Paredes, et al. „Plasma HIV-1 tropism and the risk of short-term clinical progression to AIDS or death.“ In: *PloS one* 12.1 (2017), e0166613.

[20] R. Payne, M. Muenchhoff, J. Mann, H. E. Roberts, P. Matthews, E. Adland, A. Hempenstall, K.-H. Huang, M. Brockman, Z. Brumme, et al. „Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence.“ In: *Proceedings of the National Academy of Sciences* 111.50 (2014), E5393–E5400.

[21] H. Schuitemaker, A. B. van't Wout, and P. Lusso. „Clinical significance of HIV-1 coreceptor usage." In: *Journal of translational medicine* 9 (2011), pp. 1–17.

[22] E. A. Berger, P. M. Murphy, and J. M. Farber. „Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease." In: *Annual review of immunology* 17.1 (1999), pp. 657–700.

[23] E. Coakley, C. J. Petropoulos, and J. M. Whitcomb. „Assessing chemokine co-receptor usage in HIV." In: *Current opinion in infectious diseases* 18.1 (2005), pp. 9–15.

[24] P. Dorr, M. Westby, S. Dobbs, P. Griffin, B. Irvine, M. Macartney, J. Mori, G. Rickett, C. Smith-Burchnell, C. Napier, et al. „Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity." In: *Antimicrobial agents and chemotherapy* 49.11 (2005), pp. 4721–4732.

[25] P. E. Sax. „FDA approval: maraviroc." In: *AIDS clinical care* 19.9 (2007), pp. 75–75.

[26] G. Pollakis, S. Kang, A. Kliphuis, M. I. Chalaby, J. Goudsmit, and W. A. Paxton. „N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization." In: *Journal of Biological Chemistry* 276.16 (2001), pp. 13433–13441.

[27] J. D. Siliciano, J. Kajdas, D. Finzi, T. C. Quinn, K. Chadwick, J. B. Margolick, C. Kovacs, S. J. Gange, and R. F. Siliciano. „Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells." In: *Nature medicine* 9.6 (2003), pp. 727–728.

[28] J. A. Zack, S. J. Arrigo, S. R. Weitsman, A. S. Go, A. Haislip, and I. S. Chen. „HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure." In: *Cell* 61.2 (1990), pp. 213–222.

[29] D. R. Burton. „A vaccine for HIV type 1: the antibody perspective." In: *Proceedings of the National Academy of Sciences* 94.19 (1997), pp. 10018–10023.

[30] E. Gray, P. Moore, I. Choge, J. Decker, F. Bibollet-Ruche, H. Li, N. Leseka, F. Treurnicht, K. Mlisana, G. Shaw, et al. „Neutralizing antibody responses in acute human immunodeficiency virus type 1 subtype C infection." In: *Journal of virology* 81.12 (2007), pp. 6187–6196.

[31] T. Muster, F. Steindl, M. Purtscher, A. Trkola, A. Klima, G. Himmler, F. Rüker, and H. Katinger. „A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1." In: *Journal of virology* 67.11 (1993), pp. 6642–6647.

[32] A. Trkola, A. B. Pomales, H. Yuan, B. Korber, P. J. Maddon, G. P. Allaway, H. Katinger, C. Barbas 3rd, D. R. Burton, and D. D. Ho. „Cross-clade neutralization of primary isolates of human immunodeficiency virus type 1 by human monoclonal antibodies and tetrameric CD4-IgG." In: *Journal of virology* 69.11 (1995), pp. 6609–6617.

[33] D. R. Burton, J. Pyati, R. Koduri, S. J. Sharp, G. B. Thornton, P. W. Parren, L. S. Sawyer, R. M. Hendry, N. Dunlop, P. L. Nara, et al. „Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody.“ In: *Science* 266.5187 (1994), pp. 1024–1027.

[34] M. Thali, J. Moore, C. Furman, M. Charles, D. Ho, J. Robinson, and J. Sodroski. „Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding.“ In: *Journal of virology* 67.7 (1993), pp. 3978–3988.

[35] X. Wei, J. M. Decker, S. Wang, H. Hui, J. C. Kappes, X. Wu, J. F. Salazar-Gonzalez, M. G. Salazar, J. M. Kilby, M. S. Saag, et al. „Antibody neutralization and escape by HIV-1.“ In: *Nature* 422.6929 (2003), pp. 307–312.

[36] D. R. Burton, R. C. Desrosiers, R. W. Doms, W. C. Koff, P. D. Kwong, J. P. Moore, G. J. Nabel, J. Sodroski, I. A. Wilson, and R. T. Wyatt. „HIV vaccine design and the neutralizing antibody problem.“ In: *Nature immunology* 5.3 (2004), pp. 233–236.

[37] C. Armbruster, G. M. Stiegler, B. A. Vcelar, W. Jäger, U. Köller, R. Jilch, C. G. Ammann, M. Pruenster, H. Stoiber, and H. W. Katinger. „Passive immunization with the anti-HIV-1 human monoclonal antibody (hMAb) 4E10 and the hMAb combination 4E10/2F5/2G12.“ In: *Journal of Antimicrobial chemotherapy* 54.5 (2004), pp. 915–920.

[38] J. R. Mascola, P. D'Souza, P. Gilbert, B. H. Hahn, N. L. Haigwood, L. Morris, C. J. Petropoulos, V. R. Polonis, M. Sarzotti, and D. C. Montefiori. „Recommendations for the design and use of standard virus panels to assess neutralizing antibody responses elicited by candidate human immunodeficiency virus type 1 vaccines.“ In: *Journal of virology* 79.16 (2005), pp. 10103–10107.

[39] J. F. Scheid, H. Mouquet, N. Feldhahn, B. D. Walker, F. Pereyra, E. Cutrell, M. S. Seaman, J. R. Mascola, R. T. Wyatt, H. Wardemann, et al. „A method for identification of HIV gp140 binding memory B cells in human blood.“ In: *Journal of immunological methods* 343.2 (2009), pp. 65–67.

[40] L. M. Walker, S. K. Phogat, P.-Y. Chan-Hui, D. Wagner, P. Phung, J. L. Goss, T. Wrin, M. D. Simek, S. Fling, J. L. Mitcham, et al. „Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target.“ In: *Science* 326.5950 (2009), pp. 285–289.

[41] J. F. Scheid, H. Mouquet, N. Feldhahn, M. S. Seaman, K. Velinzon, J. Pietzsch, R. G. Ott, R. M. Anthony, H. Zebroski, A. Hurley, et al. „Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals.“ In: *Nature* 458.7238 (2009), pp. 636–640.

[42] L. M. Walker, M. Huber, K. J. Doores, E. Falkowska, R. Pejchal, J.-P. Julien, S.-K. Wang, A. Ramos, P.-Y. Chan-Hui, M. Moyle, et al. „Broad neutralization coverage of HIV by multiple highly potent antibodies.“ In: *Nature* 477.7365 (2011), pp. 466–470.

[43] M. D. Simek, W. Rida, F. H. Priddy, P. Pung, E. Carrow, D. S. Laufer, J. K. Lehrman, M. Boaz, T. Tarragona-Fiol, G. Miiro, et al. „Human immunodeficiency virus type 1 elite neutralizers: individuals with broad and potent neutralizing activity identified by using a high-throughput neutralization assay together with an analytical selection algorithm.“ In: *Journal of virology* 83.14 (2009), pp. 7337–7348.

[44] D. N. Sather, S. Carbonetti, J. Kehayia, Z. Kraft, I. Mikell, J. F. Scheid, F. Klein, and L. Stamatatos. „Broadly neutralizing antibodies developed by an HIV-positive elite neutralizer exact a replication fitness cost on the contemporaneous virus.“ In: *Journal of virology* 86.23 (2012), pp. 12676–12685.

[45] E. S. Gray, M. C. Madiga, T. Hermanus, P. L. Moore, C. K. Wibmer, N. L. Tumba, L. Werner, K. Mlisana, S. Sibeko, C. Williamson, et al. „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.“ In: *Journal of virology* 85.10 (2011), pp. 4828–4840.

[46] N. A. Doria-Rose, R. M. Klein, M. M. Manion, S. O'Dell, A. Phogat, B. Chakrabarti, C. W. Hallahan, S. A. Migueles, J. Wrammert, R. Ahmed, et al. „Frequency and phenotype of human immunodeficiency virus envelope-specific B cells from patients with broadly cross-neutralizing antibodies.“ In: *Journal of virology* 83.1 (2009), pp. 188–199.

[47] D. Sok and D. R. Burton. „Recent progress in broadly neutralizing antibodies to HIV.“ In: *Nature immunology* 19.11 (2018), pp. 1179–1188.

[48] D. R. Burton and J. R. Mascola. „Antibody responses to envelope glycoproteins in HIV-1 infection.“ In: *Nature immunology* 16.6 (2015), pp. 571–576.

[49] D. H. Barouch, J. B. Whitney, B. Moldt, F. Klein, T. Y. Oliveira, J. Liu, K. E. Stephenson, H.-W. Chang, K. Shekhar, S. Gupta, et al. „Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys.“ In: *Nature* 503.7475 (2013), pp. 224–228.

[50] F. Klein, A. Halper-Stromberg, J. A. Horwitz, H. Gruell, J. F. Scheid, S. Bournazos, H. Mouquet, L. A. Spatz, R. Diskin, A. Abadir, et al. „HIV therapy by a combination of broadly neutralizing antibodies in humanized mice.“ In: *Nature* 492.7427 (2012), pp. 118–122.

[51] „Safety, pharmacokinetics and neutralization of the broadly neutralizing HIV-1 human monoclonal antibody VRC01 in healthy adults.“ In: *Clinical & Experimental Immunology* 182.3 (2015), pp. 289–301.

[52] R. M. Lynch, E. Boritz, E. E. Coates, A. DeZure, P. Madden, P. Costner, M. E. Enama, S. Plummer, L. Holman, C. S. Hendel, et al. „Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection.“ In: *Science translational medicine* 7.319 (2015), 319ra206–319ra206.

[53] R. Kong, M. K. Louder, K. Wagh, R. T. Bailer, A. deCamp, K. Greene, H. Gao, J. D. Taft, A. Gazumyan, C. Liu, et al. „Improving neutralization potency and breadth by combining broadly reactive HIV-1 antibodies targeting major neutralization epitopes.“ In: *Journal of virology* 89.5 (2015), pp. 2659–2671.

[54] M. Grobben, R. A. Stuart, and M. J. van Gils. „The potential of engineered antibodies for HIV-1 therapy and cure.“ In: *Current Opinion in Virology* 38 (2019), pp. 70–80.

[55] S. Eyuboglu, B. Karlaš, C. Ré, C. Zhang, and J. Zou. „dcbench: a benchmark for data-centric AI systems.“ In: *Proceedings of the Sixth Workshop on Data Management for End-To-End Machine Learning*. 2022, pp. 1–4.

[56] H. AI. *High-level expert group on artificial intelligence*. 2019.

[57] P. Voigt and A. Von dem Bussche. „The eu general data protection regulation (gdpr).“ In: *A Practical Guide, 1st Ed., Cham: Springer International Publishing* 10.3152676 (2017), pp. 10–5555.

[58] European Commision. *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on harmonised rules on fair access to and use of data*. © European Union 1998-2022, Accessed 2022-08-23. 2022. URL: http://eur-lex.europa.eu.

[59] A. Englander, M. Gabriel, J. Sicking, J. Volmer, J. Voosholz, A. Voss, S. Wrobel, D. Hecker, M. Poretschkin, and A. Cremers. „TRUSTWORTHY USE OF ARTIFICIAL INTELLIGENCE.“ In: (July 2019).

[60] S. S. Wilks. „The large-sample distribution of the likelihood ratio for testing composite hypotheses.“ In: *The annals of mathematical statistics* 9.1 (1938), pp. 60–62.

[61] R. E. Kass and A. E. Raftery. „Bayes factors.“ In: *Journal of the american statistical association* 90.430 (1995), pp. 773–795.

[62] J. Berger and L. Pericchi. „Bayes factors.“ In: *Wiley StatsRef: statistics reference online* (2014), pp. 1–14.

[63] D. Pfeffermann. „The role of sampling weights when modeling survey data.“ In: *International Statistical Review/Revue Internationale de Statistique* (1993), pp. 317–337.

[64] B. E. Boser, I. M. Guyon, and V. N. Vapnik. „A training algorithm for optimal margin classifiers.“ In: *Proceedings of the fifth annual workshop on Computational learning theory*. 1992, pp. 144–152.

[65] C. Cortes and V. Vapnik. „Support-vector networks.“ In: *Machine learning* 20.3 (1995), pp. 273–297.

[66] P. Meinicke, M. Tech, B. Morgenstern, and R. Merkl. „Oligo kernels for datamining on biological sequences: a case study on prokaryotic translation initiation sites.“ In: *BMC bioinformatics* 5.1 (2004), pp. 1–14.

[67] M. Bouvin-Pley, M. Morgand, A. Moreau, P. Jestin, C. Simonnet, L. Tran, C. Goujard, L. Meyer, F. Barin, and M. Braibant. „Evidence for a continuous drift of the HIV-1 species towards higher resistance to neutralizing antibodies over the course of the epidemic.“ In: *PLoS pathogens* 9.7 (2013), e1003477.

[68] M. Bouvin-Pley, M. Morgand, L. Meyer, C. Goujard, A. Moreau, H. Mouquet, M. Nussenzweig, C. Pace, D. Ho, P. Bjorkman, et al. „Drift of the HIV-1 envelope glycoprotein gp120 toward increased neutralization resistance over the course of the epidemic: a comprehensive study using the most potent and broadly neutralizing monoclonal antibodies.“ In: *Journal of virology* 88.23 (2014), pp. 13910–13917.

[69] E. M. Bunnik, L. Pisas, A. C. Van Nuenen, and H. Schuitemaker. „Autologous neutralizing humoral immunity and evolution of the viral envelope in the course of subtype B human immunodeficiency virus type 1 infection.“ In: *Journal of virology* 82.16 (2008), pp. 7932–7941.

[70] F. Nimmerjahn and J. V. Ravetch. „Antibody-mediated modulation of immune responses.“ In: *Immunological reviews* 236.1 (2010), pp. 265–275.

[71] S. Bournazos, F. Klein, J. Pietzsch, M. S. Seaman, M. C. Nussenzweig, and J. V. Ravetch. „Broadly neutralizing anti-HIV-1 antibodies require Fc effector functions for in vivo activity.“ In: *Cell* 158.6 (2014), pp. 1243–1253.

[72] X. Yu, P. B. Gilbert, C. E. Hioe, S. Zolla-Pazner, and S. G. Self. „Statistical approaches to analyzing HIV-1 neutralizing antibody assay data.“ In: *Statistics in biopharmaceutical research* 4.1 (2012), pp. 1–13.

[73] J. M. Carlson, V. Y. Du, N. Pfeifer, A. Bansal, V. Y. F. Tan, K. Power, C. J. Brumme, A. Kreimer, C. E. DeZiel, N. Fusi, M. Schaefer, M. A. Brockman, J. Gilmour, M. A. Price, W. Kilembe, R. Haubrich, M. John, S. Mallal, R. Shapiro, J. Frater, P. R. Harrigan, T. Ndung'u, S. Allen, D. Heckerman, J. Sidney, T. M. Allen, P. J. R. Goulder, Z. L. Brumme, E. Hunter, and P. A. Goepfert. „Impact of pre-adapted HIV transmission.“ In: *Nature Medicine* 22.6 (2016), pp. 606–613. ISSN: 1546-170X. DOI: 10.1038/nm.4100. URL: https://doi.org/10.1038/nm.4100.

[74] N. Erdmann, V. Y. Du, J. Carlson, M. Schaefer, A. Jureka, S. Sterrett, L. Yue, D. Dilernia, S. Lakhi, J. Tang, J. Sidney, J. Gilmour, S. Allen, E. Hunter, S. Heath, A. Bansal, and P. A. Goepfert. „HLA Class-II Associated HIV Polymorphisms Predict Escape from CD4+ T Cell Responses.“ In: *PLOS Pathogens* 11.8 (Aug. 2015). Ed. by R. Swanstrom, e1005111. DOI: 10.1371/journal.ppat.1005111. URL: https://doi.org/10.1371/journal.ppat.1005111.

[75] N. Pfeifer, H. Walter, and T. Lengauer. „Association between HIV-1 coreceptor usage and resistance to broadly neutralizing antibodies.“ In: *Journal of acquired immune deficiency syndromes (1999)* 67.2 (2014), p. 107.

[76] J. Xie and C. Liu. „Adjusted Kaplan–Meier estimator and log-rank test with inverse probability of treatment weighting for survival data.“ In: *Statistics in medicine* 24.20 (2005), pp. 3089–3110.

[77] C. M. Carvalho, N. G. Polson, and J. G. Scott. „Handling sparsity via the horseshoe.“ In: *Artificial Intelligence and Statistics*. PMLR. 2009, pp. 73–80.

[78] J. Piironen and A. Vehtari. „Sparsity information and regularization in the horseshoe and other shrinkage priors.“ In: *Electronic Journal of Statistics* 11.2 (2017), pp. 5018–5051.

[79] URL: https://datacentricai.org/neurips21/.

[80] URL: https://https-deeplearning-ai.github.io/data-centric-comp/.

[81] P. Schommers, H. Gruell, M. E. Abernathy, M.-K. Tran, A. S. Dingens, H. B. Gristick, C. O. Barnes, T. Schoofs, M. Schlotz, K. Vanshylla, et al. „Restriction of HIV-1 escape by a highly broad and potent neutralizing antibody.“ In: *Cell* 180.3 (2020), pp. 471–489.

[82] M. S. Seaman, M. Bilska, F. Ghantous, A. Eaton, C. C. LaBranche, K. Greene, H. Gao, J. A. Weiner, M. E. Ackerman, D. A. Garber, et al. „Optimization and qualification of a functional anti-drug antibody assay for HIV-1 bnAbs.“ In: *Journal of immunological methods* 479 (2020), p. 112736.

[83] H. Yoon, J. Macke, A. P. West Jr, B. Foley, P. J. Bjorkman, B. Korber, and K. Yusim. „CATNAP: a tool to compile, analyze and tally neutralizing antibody panels.“ In: *Nucleic acids research* 43.W1 (2015), W213–W219.

[84] M. J. Van der Laan, E. C. Polley, and A. E. Hubbard. „Super learner.“ In: *Statistical applications in genetics and molecular biology* 6.1 (2007).

[85] V. Lakshmanan, S. Robinson, and M. Munn. *Machine learning design patterns*. O'Reilly Media, 2020.

[86] L. E. McCoy, E. Falkowska, K. J. Doores, K. Le, D. Sok, M. J. van Gils, Z. Euler, J. A. Burger, M. S. Seaman, R. W. Sanders, et al. „Incomplete neutralization and deviation from sigmoidal neutralization curves for HIV broadly neutralizing monoclonal antibodies.“ In: *PLoS pathogens* 11.8 (2015), e1005110.

[87] N. E. Webb, D. C. Montefiori, and B. Lee. „Dose–response curve slope helps predict therapeutic potency and breadth of HIV broadly neutralizing antibodies.“ In: *Nature communications* 6.1 (2015), pp. 1–10.

[88] W.-H. Yu, D. Su, J. Torabi, C. M. Fennessey, A. Shiakolas, R. Lynch, T.-W. Chun, N. Doria-Rose, G. Alter, M. S. Seaman, et al. „Predicting the broadly neutralizing antibody susceptibility of the HIV reservoir.“ In: *JCI insight* 4.17 (2019).

[89] C. Buiu, M. V. Putz, and S. Avram. „Learning the relationship between the primary structure of HIV envelope glycoproteins and neutralization activity of particular antibodies by using artificial neural networks.“ In: *International Journal of Molecular Sciences* 17.10 (2016), p. 1710.

[90] V.-R. Dănăilă and C. Buiu. „Prediction of HIV Sensitivity to Monoclonal Antibodies Using Aminoacid Sequences and Deep Learning.“ In: *Bioinformatics* (2022).

[91] J. Davis and M. Goadrich. „The relationship between Precision-Recall and ROC curves.“ In: *Proceedings of the 23rd international conference on Machine learning*. 2006, pp. 233–240.

[92] K. Boyd, K. H. Eng, and C. D. Page. „Area under the precision-recall curve: point estimates and confidence intervals.“ In: *Joint European conference on machine learning and knowledge discovery in databases*. Springer. 2013, pp. 451–466.

[93] B. Ozenne, F. Subtil, and D. Maucort-Boulch. „The precision–recall curve overcame the optimism of the receiver operating characteristic curve in rare diseases.“ In: *Journal of clinical epidemiology* 68.8 (2015), pp. 855–859.

[94] H. R. Sofaer, J. A. Hoeting, and C. S. Jarnevich. „The area under the precision-recall curve as a performance metric for rare binary events.“ In: *Methods in Ecology and Evolution* 10.4 (2019), pp. 565–577.

[95] P. Flach and M. Kull. „Precision-recall-gain curves: PR analysis done right.“ In: *Advances in neural information processing systems* 28 (2015).

[96] L. Breiman. „Random forests.“ In: *Machine learning* 45.1 (2001), pp. 5–32.

[97] C. Chen, A. Liaw, L. Breiman, et al. „Using random forest to learn imbalanced data.“ In: *University of California, Berkeley* 110.1-12 (2004), p. 24.

[98] X. Zeng and T. R. Martinez. „Distribution-balanced stratified cross-validation for accuracy estimation.“ In: *Journal of Experimental & Theoretical Artificial Intelligence* 12.1 (2000), pp. 1–12.

[99] Y.-P. Zhang, L.-N. Zhang, and Y.-C. Wang. „Cluster-based majority under-sampling approaches for class imbalance learning.“ In: *2010 2nd IEEE International Conference on Information and Financial Engineering*. 2010, pp. 400–404. DOI: 10.1109/ICIFE.2010.5609385.

[100] W.-C. Lin, C.-F. Tsai, Y.-H. Hu, and J.-S. Jhang. „Clustering-based undersampling in class-imbalanced data.“ In: *Information Sciences* 409 (2017), pp. 17–26.

[101] W. W. Y. Ng, S. Xu, J. Zhang, X. Tian, T. Rong, and S. Kwong. „Hashing-Based Undersampling Ensemble for Imbalanced Pattern Classification Problems.“ In: *IEEE Transactions on Cybernetics* 52.2 (2022), pp. 1269–1279. DOI: 10.1109/TCYB.2020.3000754.

[102] I. Tomek. „Two Modifications of CNN.“ In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-6.11 (1976), pp. 769–772. DOI: 10.1109/TSMC.1976.4309452.

[103] I. Mani and I. Zhang. „kNN approach to unbalanced data distributions: a case study involving information extraction.“ In: *Proceedings of workshop on learning from imbalanced datasets*. Vol. 126. ICML. 2003, pp. 1–7.

[104] P. Sobhani, H. Viktor, and S. Matwin. „Learning from imbalanced data using ensemble methods and cluster-based undersampling.“ In: *International Workshop on New Frontiers in Mining Complex Patterns*. Springer. 2014, pp. 69–83.

[105] L. Breiman. „Bagging predictors.“ In: *Machine learning* 24.2 (1996), pp. 123–140.

[106] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, et al. „Xgboost: extreme gradient boosting.“ In: *R package version 0.4-2* 1.4 (2015), pp. 1–4.

[107] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. „SMOTE: synthetic minority over-sampling technique.“ In: *Journal of artificial intelligence research* 16 (2002), pp. 321–357.

[108] N. Beerenwinkel, B. Schmidt, H. Walter, R. Kaiser, T. Lengauer, D. Hoffmann, K. Korn, and J. Selbig. „Diversity and complexity of HIV-1 drug resistance: a bioinformatics approach to predicting phenotype from genotype.“ In: *Proceedings of the National Academy of Sciences* 99.12 (2002), pp. 8271–8276.

[109] N. Beerenwinkel, M. Daumer, M. Oette, K. Korn, D. Hoffmann, R. Kaiser, T. Lengauer, J. Selbig, and H. Walter. „Geno2pheno: estimating phenotypic drug resistance from HIV-1 genotypes.“ In: *Nucleic acids research* 31.13 (2003), pp. 3850–3855.

[110] T. Lengauer and T. Sing. „Bioinformatics-assisted anti-HIV therapy.“ In: *Nature Reviews Microbiology* 4.10 (2006), pp. 790–797.

[111] L. Blassel, A. Zhukova, C. J. Villabona-Arenas, K. E. Atkins, S. Hué, and O. Gascuel. „Drug resistance mutations in HIV: new bioinformatics approaches and challenges.“ In: *Current Opinion in Virology* 51 (2021), pp. 56–64. ISSN: 1879-6257. DOI: {https://doi.org/10.1016/j.coviro.2021.09.009}. URL: %7Bhttps://www.sciencedirect.com/science/article/pii/S1879625721001073%7D.

[112] T. Lengauer, O. Sander, S. Sierra, A. Thielen, and R. Kaiser. „Bioinformatics prediction of HIV coreceptor usage.“ In: *Nature biotechnology* 25.12 (2007), pp. 1407–1410.

[113] T. Sing, A. J. Low, N. Beerenwinkel, O. Sander, P. K. Cheung, F. S. Domingues, J. Büch, M. Däumer, R. Kaiser, T. Lengauer, et al. „Predicting HIV coreceptor usage on the basis of genetic and clinical covariates.“ In: *Antiviral therapy* 12.7 (2007), pp. 1097–1106.

[114] A. P. West Jr, L. Scharf, J. Horwitz, F. Klein, M. C. Nussenzweig, and P. J. Bjorkman. „Computational analysis of anti–HIV-1 antibody neutralization panel data to identify potential functional epitope residues.“ In: *Proceedings of the National Academy of Sciences* 110.26 (2013), pp. 10598–10603.

[115] M. Lacerda, P. L. Moore, N. K. Ngandu, M. Seaman, E. S. Gray, B. Murrell, M. Krishnamoorthy, M. Nonyane, M. Madiga, C. K. Wibmer, et al. „Identification of broadly neutralizing antibody epitopes in the HIV-1 envelope glycoprotein using evolutionary models.“ In: *Virology journal* 10.1 (2013), pp. 1–18.

[116] A. L. Ferguson, E. Falkowska, L. M. Walker, M. S. Seaman, D. R. Burton, and A. K. Chakraborty. „Computational prediction of broadly neutralizing HIV-1 antibody epitopes from neutralization activity data.“ In: *PLoS One* 8.12 (2013), e80562.

[117] G.-Y. Chuang, P. Acharya, S. D. Schmidt, Y. Yang, M. K. Louder, T. Zhou, Y. D. Kwon, M. Pancera, R. T. Bailer, N. A. Doria-Rose, et al. „Residue-level prediction of HIV-1 antibody epitopes based on neutralization of diverse viral strains.“ In: *Journal of virology* 87.18 (2013), pp. 10047–10058.

[118] Y. Cai, S. Karaca-Griffin, J. Chen, S. Tian, N. Fredette, C. E. Linton, S. Rits-Volloch, J. Lu, K. Wagh, J. Theiler, et al. „Antigenicity-defined conformations of an extremely neutralization-resistant HIV-1 envelope spike.“ In: *Proceedings of the National Academy of Sciences* 114.17 (2017), pp. 4477–4482.

[119] M. C. Evans, P. Phung, A. C. Paquet, A. Parikh, C. J. Petropoulos, T. Wrin, and M. Haddad. „Predicting HIV-1 broadly neutralizing antibody epitope networks using neutralization titers and a novel computational method.“ In: *BMC bioinformatics* 15.1 (2014), pp. 1–15.

[120] S. Gnanakaran, M. G. Daniels, T. Bhattacharya, A. S. Lapedes, A. Sethi, M. Li, H. Tang, K. Greene, H. Gao, B. F. Haynes, et al. „Genetic signatures in the envelope glycoproteins of HIV-1 that associate with broadly neutralizing antibodies.“ In: *PLoS computational biology* 6.10 (2010), e1000955.

[121] L. He and J. Zhu. „Computational tools for epitope vaccine design and evaluation.“ In: *Current opinion in virology* 11 (2015), pp. 103–112.

[122] N. L. Hepler, K. Scheffler, S. Weaver, B. Murrell, D. D. Richman, D. R. Burton, P. Poignard, D. M. Smith, and S. L. Kosakovsky Pond. „IDEPI: rapid prediction of HIV-1 antibody epitopes and other phenotypic features from sequence data using a flexible machine learning platform.“ In: *PLoS computational biology* 10.9 (2014), e1003842.

[123] R. Rawi, R. Mall, C.-H. Shen, S. K. Farney, A. Shiakolas, J. Zhou, H. Bensmail, T.-W. Chun, N. A. Doria-Rose, R. M. Lynch, et al. „Accurate prediction for antibody resistance of clinical HIV-1 isolates.“ In: *Scientific reports* 9.1 (2019), pp. 1–12.

[124] C. A. Magaret, D. C. Benkeser, B. D. Williamson, B. R. Borate, L. N. Carpp, I. S. Georgiev, I. Setliff, A. S. Dingens, N. Simon, M. Carone, et al. „Prediction of VRC01 neutralization sensitivity by HIV-1 gp160 sequence features.“ In: *PLoS computational biology* 15.4 (2019), e1006952.

[125] S. Conti and M. Karplus. „Estimation of the breadth of CD4bs targeting HIV antibodies by molecular modeling and machine learning.“ In: *PLoS computational biology* 15.4 (2019), e1006954.

[126] B. D. Williamson, C. A. Magaret, P. B. Gilbert, S. Nizam, C. Simmons, and D. Benkeser. „Super LeArner Prediction of NAb Panels (SLAPNAP): a containerized tool for predicting combination monoclonal broadly neutralizing antibody sensitivity.“ In: *Bioinformatics* 37.22 (2021), pp. 4187–4192.

[127] S. t. Loewe and H. Muischnek. „Über kombinationswirkungen.“ In: *Naunyn-Schmiedebergs Archiv für experimentelle Pathologie und Pharmakologie* 114.5 (1926), pp. 313–326.

[128] S. Loewe. „Die quantitativen probleme der pharmakologie.“ In: *Ergebnisse der Physiologie* 27.1 (1928), pp. 47–187.

[129] S. Loewe. „The problem of synergism and antagonism of combined drugs.“ In: *Arzneimittelforschung* 3 (1953), pp. 285–290.

[130] W. R. Greco. „The search for synergy: a critical review from a response surface perspective.“ In: *Pharmacol Rev* 47 (1995), pp. 331–385.

[131] K. Wagh, T. Bhattacharya, C. Williamson, A. Robles, M. Bayne, J. Garrity, M. Rist, C. Rademeyer, H. Yoon, A. Lapedes, et al. „Optimal combinations of broadly neutralizing antibodies for prevention and treatment of HIV-1 clade C infection.“ In: *PLoS pathogens* 12.3 (2016), e1005520.

[132] H. Zhang, J. Holden-Wiltse, J. Wang, and H. Liang. „A strategy to model nonmonotonic dose-response curve and estimate IC50.“ In: *PLoS One* 8.8 (2013), e69301.

[133] D. A. Grimes and K. F. Schulz. „Bias and causal associations in observational research.“ In: *The lancet* 359.9302 (2002), pp. 248–252.

[134] P. R. Rosenbaum and D. B. Rubin. „The central role of the propensity score in observational studies for causal effects.“ In: *Biometrika* 70.1 (1983), pp. 41–55.

[135] N. P. Jewell. *Statistics for epidemiology*. chapman and hall/CRC, 2003.

[136] R. McNamee. „Confounding and confounders.“ In: *Occupational and environmental medicine* 60.3 (2003), pp. 227–234.

[137] J. Pearl and D. Mackenzie. *The book of why: the new science of cause and effect*. Basic books, 2018.

[138] H. B. Mann and D. R. Whitney. „On a test of whether one of two random variables is stochastically larger than the other.“ In: *The annals of mathematical statistics* (1947), pp. 50–60.

[139] F. Wilcoxon. „Individual comparisons of grouped data by ranking methods.“ In: *Journal of economic entomology* 39.2 (1946), pp. 269–270.

[140] F. Wilcoxon. „Individual comparisons by ranking methods.“ In: *Breakthroughs in statistics*. Springer, 1992, pp. 196–202.

[141] R. F. Woolson. „Wilcoxon signed-rank test.“ In: *Wiley encyclopedia of clinical trials* (2007), pp. 1–3.

[142] Student. „The probable error of a mean.“ In: *Biometrika* (1908), pp. 1–25.

[143] E. Brunner, A. C. Bathke, and F. Konietschke. *Rank and pseudo-rank procedures for independent observations in factorial designs*. Springer, 2018.

[144] J. D. Karch. „Psychologists should use Brunner-Munzel's instead of Mann-Whitney's U test as the default nonparametric procedure.“ In: *Advances in Methods and Practices in Psychological Science* 4.2 (2021). DOI: 10.1177/2515245921999602. eprint: https://doi.org/10.1177/2515245921999602. URL: https://doi.org/10.1177/2515245921999602.

[145] G. W. Brown, A. M. Mood, et al. „On median tests for linear hypotheses.“ In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Vol. 2. University of California Press Berkeley. 1951, pp. 159–166.

[146] B. Efron and R. J. Tibshirani. *An introduction to the bootstrap*. CRC press, 1994.

[147] M. Oestreich, D. Chen, J. L. Schultze, M. Fritz, and M. Becker. „Privacy considerations for sharing genomics data.“ In: *EXCLI journal* 20 (2021), p. 1243.

[148] M. Gymrek, A. L. McGuire, D. Golan, E. Halperin, and Y. Erlich. „Identifying personal genomes by surname inference.“ In: *Science* 339.6117 (2013), pp. 321–324.

[149] L. Sweeney, A. Abu, and J. Winn. „Identifying participants in the personal genome project by name (a re-identification experiment).“ In: *arXiv preprint arXiv:1304.7605* (2013).

[150] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al. „The FAIR Guiding Principles for scientific data management and stewardship.“ In: *Scientific data* 3.1 (2016), pp. 1–9.

[151] M. L. Metzker, D. P. Mindell, X.-M. Liu, R. G. Ptak, R. A. Gibbs, and D. M. Hillis. „Molecular evidence of HIV-1 transmission in a criminal case.“ In: *Proceedings of the National Academy of Sciences* 99.22 (2002), pp. 14292–14297.

[152] S. R. Mehta, S. A. Vinterbo, and S. J. Little. „Ensuring privacy in the study of pathogen genetics.“ In: *The Lancet infectious diseases* 14.8 (2014), pp. 773–777.

[153] J. Levine. *HIV and the Law: Risks, Rights & Health*. 2012.

[154] J. U. N. P. on HIV/AIDS (UNAIDS) et al. *Ending overly broad criminalization of HIV non-disclosure, exposure and transmission: Critical scientific, medical and legal considerations. Geneva: UNAIDS; 2013*. 2013.

[155] C. Dwork, A. Roth, et al. „The algorithmic foundations of differential privacy.“ In: *Foundations and Trends® in Theoretical Computer Science* 9.3–4 (2014), pp. 211–407.

[156] Y. Erlich and A. Narayanan. „Routes for breaching and protecting genetic privacy.“ In: *Nature Reviews Genetics* 15.6 (2014), pp. 409–421.

[157] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas. „Communication-efficient learning of deep networks from decentralized data.“ In: *Artificial intelligence and statistics*. PMLR. 2017, pp. 1273–1282.

[158] N. Rieke, J. Hancox, W. Li, F. Milletari, H. R. Roth, S. Albarqouni, S. Bakas, M. N. Galtier, B. A. Landman, K. Maier-Hein, et al. „The future of digital health with federated learning.“ In: *NPJ digital medicine* 3.1 (2020), pp. 1–7.

[159] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, et al. „Advances and open problems in federated learning.“ In: *Foundations and Trends® in Machine Learning* 14.1–2 (2021), pp. 1–210.

[160] S. Warnat-Herresthal, H. Schultze, K. L. Shastry, S. Manamohan, S. Mukherjee, V. Garg, R. Sarveswara, K. Händler, P. Pickkers, N. A. Aziz, et al. „Swarm learning for decentralized and confidential clinical machine learning.“ In: *Nature* 594.7862 (2021), pp. 265–270.

[161] A. Alaa, B. Van Breugel, E. S. Saveliev, and M. van der Schaar. „How faithful is your synthetic data? sample-level metrics for evaluating and auditing generative models.“ In: *International Conference on Machine Learning*. PMLR. 2022, pp. 290–306.

[162] R. C. Shean and A. L. Greninger. „Private collection: high correlation of sample collection and patient admission date in clinical microbiological testing complicates sharing of phylodynamic metadata.“ In: *Virus evolution* 4.1 (2018), vey005.

[163] L. Bonomi, Y. Huang, and L. Ohno-Machado. „Privacy challenges and research opportunities for genomic data sharing.“ In: *Nature genetics* 52.7 (2020), pp. 646–654.

[164] D. Butler. „Data sharing threatens privacy.“ In: *Nature* 449.7163 (2007), pp. 644–646.

[165] A. Bernier and B. M. Knoppers. „Pandemics, privacy, and public health research.“ In: *Canadian Journal of Public Health* 111.4 (2020), pp. 454–457.

[166] L. Song, H. Liu, F. S. Brinkman, E. Gill, E. J. Griffiths, W. W. Hsiao, S. Savić-Kallesøe, S. Moreira, G. Van Domselaar, H. Z. Ma'n, et al. „Addressing privacy concerns in sharing viral sequences and minimum contextual data in a public repository during the COVID-19 pandemic.“ In: *Frontiers in genetics* 12 (2021).

[167] J. Köster and S. Rahmann. „Snakemake—a scalable bioinformatics workflow engine.“ In: *Bioinformatics* 28.19 (2012), pp. 2520–2522.

[168] *Anaconda Software Distribution*. Version Vers. 2-2.4.0. 2020. URL: https://docs.anaconda.com/.

[169] W. Chen and D. S. Dimitrov. „Human monoclonal antibodies and engineered antibody domains as HIV-1 entry inhibitors.“ In: *Current opinion in HIV and AIDS* 4.2 (2009), p. 112.

[170] Z. A. Ahmad, S. K. Yeap, A. M. Ali, W. Y. Ho, N. B. M. Alitheen, and M. Hamid. „scFv antibody: principles and clinical application.“ In: *Clinical and developmental immunology* 2012 (2012).

[171] A. P. West Jr, R. P. Galimidi, P. N. Gnanapragasam, and P. J. Bjorkman. „Single-chain Fv-based anti-HIV proteins: potential and limitations.“ In: *Journal of virology* 86.1 (2012), pp. 195–202.

[172] R. T. van Dorsten, B. E. Lambson, C. K. Wibmer, M. S. Weinberg, P. L. Moore, and L. Morris. „Neutralization breadth and potency of single-chain variable fragments derived from broadly neutralizing antibodies targeting multiple epitopes on the HIV-1 envelope.“ In: *Journal of Virology* 94.2 (2020), e01533–19.

[173] R. T. van Dorsten, L. Reh, A. Trkola, L. Morris, and P. L. Moore. „Single-Chain Variable Fragments of Broadly Neutralizing Antibodies Prevent HIV Cell-Cell Transmission.“ In: *Journal of Virology* 96.4 (2022), e01934–21.

[174] M. Asokan, R. Rudicell, M. Louder, K. McKee, S. O'dell, G. Stewart-Jones, K. Wang, L. Xu, X. Chen, M. Choe, et al. „Bispecific antibodies targeting different epitopes on the HIV-1 envelope exhibit broad and potent neutralization.“ In: *Journal of virology* 89.24 (2015), pp. 12501–12512.

[175] L. Xu, A. Pegu, E. Rao, N. Doria-Rose, J. Beninga, K. McKee, D. M. Lord, R. R. Wei, G. Deng, M. Louder, et al. „Trispecific broadly neutralizing HIV antibodies mediate potent SHIV protection in macaques.“ In: *Science* 358.6359 (2017), pp. 85–90.

[176] K. Wagh, M. S. Seaman, M. Zingg, T. Fitzsimons, D. H. Barouch, D. R. Burton, M. Connors, D. D. Ho, J. R. Mascola, M. C. Nussenzweig, et al. „Potential of conventional & bispecific broadly neutralizing antibodies for prevention of HIV-1 subtype A, C & D infections.“ In: *PLoS pathogens* 14.3 (2018), e1006860.

[177] S. N. Khan, D. Sok, K. Tran, A. Movsesyan, V. Dubrovskaya, D. R. Burton, and R. T. Wyatt. „Targeting the HIV-1 spike and coreceptor with bi- and trispecific antibodies for single-component broad inhibition of entry.“ In: *Journal of virology* 92.18 (2018), e00384–18.

[178] J. J. Steinhardt, J. Guenaga, H. L. Turner, K. McKee, M. K. Louder, S. O'Dell, C.-I. Chiang, L. Lei, A. Galkin, A. K. Andrianov, et al. „Rational design of a trispecific antibody targeting the HIV-1 Env with elevated antiviral activity.“ In: *Nature communications* 9.1 (2018), pp. 1–12.

[179] J. Z. Li and D. R. Kuritzkes. „Clinical implications of HIV-1 minority variants.“ In: *Clinical infectious diseases* 56.11 (2013), pp. 1667–1674.

[180] K. Van Laethem, K. Theys, and A.-M. Vandamme. „HIV-1 genotypic drug resistance testing: digging deep, reaching wide?“ In: *Current opinion in virology* 14 (2015), pp. 16–23.

[181] H. Ji, P. Sandstrom, R. Paredes, P. R. Harrigan, C. J. Brumme, S. Avila Rios, M. Noguera-Julian, N. Parkin, and R. Kantor. „Are we ready for NGS HIV drug resistance testing? The second “Winnipeg Consensus” symposium.“ In: *Viruses* 12.6 (2020), p. 586.

[182] O. Zagordi, R. Klein, M. Däumer, and N. Beerenwinkel. „Error correction of next-generation sequencing data and reliable estimation of HIV quasispecies.“ In: *Nucleic acids research* 38.21 (2010), pp. 7400–7409.

[183] L. Fuhrmann, K. P. Jablonski, and N. Beerenwinkel. „Quantitative measures of within-host viral genetic diversity.“ In: *Current Opinion in Virology* 49 (2021), pp. 157–163. ISSN: 1879-6257. DOI: https://doi.org/10.1016/j.coviro.2021.06.002. URL: https://www.sciencedirect.com/science/article/pii/S1879625721000651.

[184] S. Posada-Céspedes, D. Seifert, I. Topolsky, K. P. Jablonski, K. J. Metzner, and N. Beerenwinkel. „V-pipe: a computational pipeline for assessing viral genetic diversity from high-throughput data.“ In: *Bioinformatics* 37.12 (Jan. 2021), pp. 1673–1680. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btab015. eprint: https://academic.oup.com/bioinformatics/article-pdf/37/12/1673/39119206/btab015.pdf. URL: https://doi.org/10.1093/bioinformatics/btab015.

[185] M. Döring, J. Büch, G. Friedrich, A. Pironti, P. Kalaghatgi, E. Knops, E. Heger, M. Obermeier, M. Däumer, A. Thielen, et al. „geno2pheno [ngs-freq]: a genotypic interpretation system for identifying viral drug resistance using next-generation sequencing data.“ In: *Nucleic acids research* 46.W1 (2018), W271–W277.

[186] M. C. Sneller, J. Blazkova, J. S. Justement, V. Shi, B. D. Kennedy, K. Gittens, J. Tolstenko, G. McCormack, E. J. Whitehead, R. F. Schneck, et al. „Combination anti-HIV antibodies provide sustained virological suppression.“ In: *Nature* (2022), pp. 1–7.

[187] B. Schubert, M. Walzer, H.-P. Brachvogel, A. Szolek, C. Mohr, and Kohlbacher, Oliver. „FRED 2: an immunoinformatics framework for Python.“ In: *Bioinformatics* 32.13 (Feb. 2016), pp. 2044–2046. ISSN: 1367-4803. DOI: {10.1093/bioinformatics/btw113}. eprint: https://academic.oup.com/bioinformatics/article-pdf/32/13/2044/16920742/btw113.pdf. URL: https://doi.org/10.1093/bioinformatics/btw113.

[188] *IEDB Analysis Resource - T Cell Prediction Tools*. URL: http://tools.iedb.org/main/tcell/.

[189] M. S. Seaman, H. Janes, N. Hawkins, L. E. Grandpre, C. Devoy, A. Giri, R. T. Coffey, L. Harris, B. Wood, M. G. Daniels, et al. „Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies.“ In: *Journal of virology* 84.3 (2010), pp. 1439–1452.

[190] D. C. Montefiori, M. Roederer, L. Morris, and M. S. Seaman. „Neutralization tiers of HIV-1.“ In: *Current Opinion in HIV and AIDS* 13.2 (2018), p. 128.