Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies

Nicole A. Doria-Rose[1,*], **Chaim A. Schramm**[2,*], **Jason Gorman**[1,*], **Penny L. Moore**[3,4,5,*], **Jinal N. Bhiman**[3,4], **Brandon J. DeKosky**[6], **Michael J. Ernandes**[1], **Ivelin S. Georgiev**[1], **Helen J. Kim**[7,8,9], **Marie Pancera**[1], **Ryan P. Staupe**[1], **Han R. Altae-Tran**[1], **Robert T. Bailer**[1], **Ema T. Crooks**[10], **Albert Cupo**[11], **Aliaksandr Druz**[1], **Nigel J. Garrett**[5], **Kam H. Hoi**[12], **Rui Kong**[1], **Mark K. Louder**[1], **Nancy S. Longo**[1], **Krisha McKee**[1], **Molati Nonyane**[3], **Sijy O'Dell**[1], **Ryan S. Roark**[1], **Rebecca S. Rudicell**[1], **Stephen D. Schmidt**[1], **Daniel J. Sheward**[13], **Cinque Soto**[1], **Constantinos Kurt Wibmer**[3,4], **Yongping Yang**[1], **Zhenhai Zhang**[2], **NISC Comparative Sequencing Program**[14], **James C. Mullikin**[14,15], **James M. Binley**[10], **Rogier W. Sanders**[16], **Ian A. Wilson**[7,8,9,17], **John P. Moore**[11], **Andrew B. Ward**[7,8,9], **George Georgiou**[6,12,18], **Carolyn Williamson**[5,13], **Salim S. Abdool Karim**[5,19], **Lynn Morris**[3,4,5,#], **Peter D. Kwong**[1,#], **Lawrence Shapiro**[1,2,#], and **John R. Mascola**[1,#]

[1]Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892 USA [2]Department of Biochemistry, Columbia University, New York, NY 10032 USA [3]Center for HIV and STIs, National Institute for Communicable Diseases of the National Health Laboratory Service (NHLS), Johannesburg, South Africa [4]University of the Witwatersrand, Johannesburg, South Africa [5]Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Durban, South Africa [6]Department of Chemical Engineering, University of Texas at Austin, Austin, TX, USA [7]Department of Integrative Structural and Computational Biology, The Scripps Research Institute,

Readers are welcome to comment on the online version of this article at www.nature.com/nature.

La Jolla, California, 92037 USA [8]Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery, The Scripps Research Institute, La Jolla, California, 92037 USA [9]IAVI Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, California, 92037 USA [10]Torrey Pines Institute, San Diego, CA 92037 USA [11]Weill Medical College of Cornell University, New York, NY 10021 USA [12]Department of Biomedical Engineering, University of Texas at Austin, Austin, TX, USA [13]Institute of Infectious Diseases and Molecular Medicine, Division of Medical Virology, University of Cape Town and NHLS, Cape Town, South Africa [14]NISC Comparative Sequencing program, NIH, Bethesda, MD 20892 USA [15]NIH Intramural Sequencing Center (NISC), National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA [16]Department of Medical Microbiology, Academic Medical Center, Amsterdam Netherlands [17]Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, USA [18]Department of Molecular Biosciences, University of Texas at Austin, Austin, TX, USA [19]Department of Epidemiology, Columbia University, New York, NY, USA

Summary

Antibodies capable of neutralizing HIV-1 often target variable regions 1 and 2 (V1V2) of the HIV-1 envelope, but the mechanism of their elicitation has been unclear. Here we define the developmental pathway by which such antibodies are generated and acquire the requisite molecular characteristics for neutralization. Twelve somatically related neutralizing antibodies (CAP256-VRC26.01-12) were isolated from CAPRISA-donor CAP256; each antibody contained the protruding tyrosine-sulfated, anionic antigen-binding loop (CDR H3) characteristic of this category of antibodies. Their unmutated ancestor emerged between weeks 30–38 post-infection with a 35-residue CDR H3, and neutralized the virus that superinfected this individual 15 weeks after initial infection. Improved neutralization breadth occurred by week 59 with modest affinity maturation, and was preceded by extensive diversification of the virus population. HIV-1 V1V2-directed neutralizing antibodies can thus develop relatively rapidly through initial selection of B cells with a long CDR H3, and limited subsequent somatic hypermutation, an important vaccine insight.

Developmental pathways of antibodies that neutralize HIV-1 represent potential templates to guide vaccine strategies, if their constituent molecular events were understood and could be reproduced[1–3]. Virtually all HIV-1 infected individuals mount a potent antibody response within months of infection, but this response preferentially neutralizes autologous virus, which rapidly escapes[4,5]. Cross-reactive antibodies capable of neutralizing most HIV-1 strains arise in only ~20% of donors after 2–3 years of infection[6–9]. An understanding of the development of broadly neutralizing antibody (NAb) lineages in such donors could provide a roadmap for vaccine design.

One means to obtain such a roadmap is through isolation of broadly cross-reactive NAbs, characterization of their genetic sequence and molecular properties, and examination of the B cell genetic record with next-generation sequencing (NGS)[10–14]. The greatest insight can be gained with longitudinal sampling from early after the time of HIV-1 infection[15]. This allows for a genetic delineation of the molecular evolution leading from an unmutated

ancestor antibody through affinity maturation to acquisition of neutralization breadth. In principle, such a roadmap should link antibody molecular characteristics to the genetic development that a successful vaccine would retrace.

NAbs to the V1V2 region of the HIV-1 viral spike are among the most prevalent cross-reactive antibodies elicited by natural infection[6,16–18] and have been isolated from several donors[19–21]. These antibodies have long heavy-chain complementarity-determining region 3 loops (CDR H3s) that are protruding, anionic, and often tyrosine sulfated[22,23]. These CDR H3s penetrate the HIV-1 glycan shield, recognizing a quaternary glycopeptide epitope at the apex of the HIV-1 spike that is formed by V1V2s from at least two gp120 protomers[22–24]. Here we employ antibody isolation, B-cell NGS, structural characterization, and viral single-genome analysis (SGA) to delineate longitudinal interactions between the developing antibody and autologous virus within donor CAP256, who showed evidence of V1V2-mediated neutralization breadth after one year[18,25,26]. Our results define the molecular requirements and genetic pathways that lead to V1V2-directed neutralization, providing a template for their vaccine elicitation.

Antibody isolation and characterization

Donor CAP256 peripheral blood mononuclear cells (PBMCs) sampled 59, 119 and 206 weeks post-infection were used to isolate 12 monoclonal antibodies by high-throughput B cell culture, functional screening by microneutralization, and reverse transcription-PCR of antibody variable regions[27,28] (Fig. 1a). All 12 were somatically related and distinguished by long CDR H3s of 35–37 amino acids (Kabat[29] numbering) (Fig. 1b, Extended Data Fig. 1a). The heavy and light chains exhibited somatic mutation of 4–15% from their germline-encoded V-genes, VH3-30 and Vλ1-51, respectively (Extended Data Fig. 1 and Extended Data Table 1). When these antibodies were reconstituted as IgG1s, they showed varying degrees of heterologous virus neutralization and were extremely potent against many subtype A and C strains (Fig. 1b–c, Extended Data Fig. 2, and Supplementary Fig. 1). Importantly, the combination of all 12 antibodies recapitulated plasma neutralization (Supplementary Fig. 2), indicating the CAP256-VRC26 antibody lineage to be responsible for the neutralization breadth and potency of donor CAP256.

To map the epitope of the CAP256-VRC26 antibodies, we employed neutralization fingerprints[18]; binding assays for HIV-1 Envelope (Env) in soluble, cell surface[30], and viral particle[31] contexts; and negative stain electron microscopy (EM) of Fab CAP256-VRC26.09 bound to a soluble cleaved version of the HIV-1 spike[24,32,33] (Fig. 2a–c, Supplementary Fig. 3, Extended Data Figs. 3–4). Recognition of Env by CAP256-VRC26 antibodies was similar to PG9-class NAbs that recognize the trimeric V1V2 cap[24], with high specificity for the Env native quaternary conformation and one Fab bound per trimer (Fig. 2c, left and Extended Data Fig. 4). Neutralization activity of CAP256-VRC26 antibodies was reduced or knocked out by Env mutations in V1V2 strands B and C (Fig. 2d), much like the CAP256 plasma[25,26] and PG9-class NAbs[22,23,34]; though unlike PG9, the CAP256-VRC26 antibodies were only partially and variably sensitive to loss of glycans at N160 and N156 (Fig. 2d and Extended Data Fig. 5). Overall, these data indicated the epitope to be at the membrane-distal apex of

the HIV-1 spike close to the trimer axis (Fig. 2e), providing a structural explanation for the observed quaternary specificity.

Origin and development of the lineage

To obtain a genetic record of the CAP256-VRC26 antibody lineage, we analyzed B cell-immunoglobulin transcripts at eight time points between 15 and 206 weeks post-infection by 454 pyrosequencing. Although no CAP256-VRC26 lineage-related transcripts were detected at 15 and 30 weeks, related heavy chain and light chain transcripts were found at all later time points (Fig. 3a). To track longitudinal prevalence, we utilized identity-divergence plots of all heavy chain reads assigned to the same VH3-30 germline gene as the isolated antibodies. Using CAP256-VRC26.01 or CAP256-VRC26.08 as the identity referents, segregated islands of related heavy chain sequences first appeared at week 38 (Fig. 3b). For all 12 antibodies, the prevalence and identity of related sequences peaked close to the time of the antibody isolation (Supplementary Fig. 4). To obtain additional antibody lineage data, we performed linked VH:VL paired sequencing[35] at five time points (Fig. 3a and Supplementary Table 1). Of 157 unique CAP256-VRC26 pairs, 7 matched either heavy or light chain sequences present in the 454 pyrosequencing data, including 2 for which both heavy and light chain sequences had previously been captured (Fig. 3c).

Maximum-likelihood phylogenetic trees were constructed using the isolated antibodies and the 454 data (Fig. 3c). The lineage bifurcates early, with one branch leading to CAP256-VRC26.01 and a second developing into CAP256-VRC26.02-12. The unmutated common ancestors (UCAs) for the heavy and light chain were inferred from the phylogenetic trees (Fig. 3c). For the light chain, the UCA had a 12-residue CDR L3, as in CAP256-VRC26.01, and for the heavy chain, the inferred UCA had a 35-residue CDR H3 (Extended Data Fig. 6), likely the result of VDJ recombination with a single D-gene, IgHD3-3*01 and N-nucleotide insertions of 34 and 31 nucleotides at each junction (Supplementary Fig. 5). This inferred UCA was further supported by the identification of many low-divergence sequences in week 38 heavy chain data, five of which had only two amino acid differences from the inferred UCA (Extended Data Fig. 6). Thus, the longitudinal NGS analysis established the first appearance of the CAP256-VRC26 lineage; defined the UCA, the product of gene recombination in the ancestor B cell of the lineage; and provided a genetic record of the development of this lineage over four years.

Structures of CAP256-VRC26 antibodies

To define the structural characteristics of CAP256-VRC26 lineage development, we determined crystal structures for Fabs of the UCA and six antibodies from weeks 59, 119 and 206 (Fig. 4, Supplementary Table 2, and Supplementary Fig. 6a). The mature CDR H3s protruded ~20Å above the antigen-combining surface of the heavy chain and contained a 2-stranded β-sheet, *O*-sulfated tyrosines, and an intra-CDR H3 disulfide bond (Fig. 4a–b). The CDR H3s of the UCA and CAP256-VRC26.01 lacked a CDR H3 disulfide bond, exhibited greater disorder, and were positioned more proximal to the light chain (Fig 4c); the appearance of the disulfide bond correlated with adoption of the mature CDR H3 orientation (Fig 4c, Supplementary Fig. 6b, and Extended Data Fig. 7a). Mutation to remove the

relevant Cys residues in VRC26.03 resulted in loss of neutralization potency and breadth (Extended Data Fig. 7b–c). Additionally, the appearance of CDR H3 cysteines coincided with a glycine to arginine mutation at the base of the CDR H3, possibly limiting flexibility of the mature antibodies (Extended Data Fig. 7a–b and Supplementary Fig. 7). Overall, the CAP256-VRC26 lineage begins with an anionic protruding CDR H3 with structural properties similar to previously determined V1V2-directed broadly NAbs. Development over four years involves the introduction of almost 20 light chain and over 30 heavy chain mutations, including a disulfide bond. The CDR H3 changes its overall orientation while losing negative charge and maintaining tyrosine sulfation (Fig. 4b–c, right).

HIV Env evolution during NAb development

To gain insight into the temporal HIV-1 Env changes driving the development of the CAP256-VRC26 lineage, we used SGA to determine viral sequences over ~3 years. CAP256 Env sequences showed high levels of diversity driven, in part, by recombination between the superinfecting virus (SU) that was first detected 15 weeks post-infection, and the primary infecting virus (PI)[26] (Fig. 5a, Supplementary Figs. 8, 9). Differences between the PI and SU Env sequences included V2 residues 165 and 169, and an N160 glycan in the SU that was not present in the PI (Fig 5b, Extended Data Fig. 8a–b). Notably, compared to the PI, the SU contained V2 residues that are more commonly found among circulating viruses (Extended Data Fig 8a). All 12 antibodies neutralized the SU, and, with the exception of CAP256-VRC26.06, failed to neutralize the PI, suggesting the SU V1V2 initially engaged the naive B cell of the CAP256-VRC26 lineage (Fig. 5d, Extended Data Fig. 8c, and Supplementary Fig. 10).

Before the CAP256-VRC26 antibodies developed, most Env sequences had V1V2 regions derived from the PI (Fig. 5a–c and Supplementary Fig. 8–9) and were therefore largely neutralization resistant (Fig. 5d and Supplementary Fig. 10). Among SU-like sequences, a rare K169I mutation arose under strong directional selection (Supplementary Table 3) as the CAP256-VRC26 lineage emerged and rendered the SU resistant to only the earliest antibody (Extended Data Fig. 8d–e), suggesting that CAP256-VRC26.01-like antibodies drove this viral escape, followed by maturation of the lineage to tolerate I169. At 48 weeks, the viral population underwent a significant shift (Fig 5a and Supplementary Figs. 8–9), with the SU-like V1V2 dominating just prior to the development of neutralization breadth. Neutralization of Env clones by later antibodies (CAP256-VRC26.02-12) tracked with the presence of SU-like V1V2 sequences (black bar, Fig 5c) until escape occurred through mutations at positions 166 or 169 (Fig 5c–d, Extended Data Figs. 8d). These mutations resulted in a net charge change in the V2 epitope (+3 to 0, Fig. 5c, Extended Data Fig. 8b) concomitant with the antibody CDR H3s becoming less acidic over time (−10 to −4, Fig. 4 and Extended Data Fig. 9) suggesting co-evolution of the viral epitope and the antibody paratope. Overall, these results highlight the interplay between virus and antibody, with the SU-like V1V2 epitope stimulating expansion of the CAP256-VRC26 lineage.

Rapid development of CAP256-VRC26.01

To gain insight into the development of V1V2-directed neutralization, we focused on the early antibody CAP256-VRC26.01, isolated at week 59, which neutralized 30% of clade C viruses and showed cross-clade neutralization of nearly 20% (Supplementary Fig 1). Notably, this week 59 time point was 44 weeks after superinfection and only 21 weeks after the CAP256-VRC26 lineage was first detected by NGS. We also inferred heavy and light chains for two developmental intermediates (VRC26-I1 and VRC26-I2) (Fig. 6a and Extended Data Fig. 1) and characterized their function along with the UCA (Fig. 6b–e): The UCA bound and neutralized the SU weakly, but did not bind nor neutralize heterologous viruses. VRC26-I1, VRC26-I2 and CAP256-VRC26.01 demonstrated progressively greater binding and neutralization, with VRC26-I1 neutralizing 2 of 7 strains and VRC26-I2 neutralizing 6 of 7 strains (Fig 6e), with dependence on residues in V2 (Fig 6c). Interestingly, the PI was neither bound nor neutralized by the UCA, intermediates, or CAP256-VRC26.01 (Fig 6c and Supplementary Fig 11). These data provide further evidence that the CAP256-VRC26 lineage was initiated by interaction with an SU-like V1V2. Subsequent affinity maturation, focused in CDR H3 (Fig. 6f and Extended Data Table 1), allowed for progressively greater binding and neutralization with increased viral diversity preceding the emergence of neutralization breadth. Based on the inferred UCA, CAP256-VRC26.01 diverged 11% from germline heavy chain and 7% from germline light chain (Fig. 6f). Thus, once an appropriate gene recombination allows for B-cell receptor recognition of the trimeric V1V2 epitope, development of cross-reactive neutralization can be achieved with moderate somatic mutation in a matter of months.

Vaccine implications

The V1V2 region of HIV-1 is a common target of serum NAbs[6,16–18]. In the RV144 Thai vaccine trial, an increased level of binding antibodies to the V1V2 region was associated with a reduced risk of infection[36] and viral sieve analysis showed immune pressure in the same region[37]. While the vaccine in the RV144 trial did not elicit broadly neutralizing V1V2-directed antibodies similar to those described here and elsewhere[19–21], a more effective vaccine would ideally elicit cross-reactive NAbs [1–3,38]. Previously described V1V2 antibodies, and the CAP256-VRC26 lineage, all have long CDR H3 regions that are necessary to penetrate the glycan shield and engage a V1V2 epitope (Extended Data Table 1). An important unanswered question has been whether these long CDR H3s are fully formed by VDJ recombination, as has been seen in HIV-uninfected donors[39], or emerge by insertions during the process of affinity maturation. We show here that the 35-residue CDR H3 of the CAP256-VRC26 UCA was produced during initial gene rearrangement and therefore existed at the level of the naive B cell receptor.

A potential rate-limiting developmental step in the CAP256-VRC26 lineage is the gene rearrangement that generated its UCA. By one estimate, human B cells with recombined antibody genes encoding long (24aa, IMGT[40] definition) or very long (28aa) CDR H3s constitute ~3.5% and 0.4%, respectively, of naïve B cells[39]. These long B cell receptors have been associated with autoreactivity, and are subject to both central and peripheral deletion, resulting in an even smaller population of IgG[+] memory B cells[39,41]. We therefore

tested the UCA and all 12 CAP256-VRC26 cloned antibodies for autoreactivity[42]. The UCA and mature CAP256-VRC26 antibodies demonstrated little or no reactivity on Hep2 cells or with cardiolipin (Extended Data Fig. 6b–c). In addition, NGS of CAP256 peripheral B cells indicated that <0.4% of sequences had CDR H3s of 28aa (Extended Data Fig. 6d) suggesting that this donor did not have an unusually high frequency of clonal lineages with long CDR H3 regions.

We also inferred the virological events leading to the stimulation and evolution of the CAP256-VRC26 lineage by the superinfecting virus. Similar to the CH103 CD4-binding site lineage in donor CH505[15], the autologous virus in CAP256 showed extensive diversification prior to the development of breadth. Subsequent antibody-virus interactions appeared to drive somatic mutation and development of cross-reactive neutralization. Finally, the ontogeny of V1V2-directed NAbs revealed by the CAP256-VRC26 lineage suggests that neutralization potency and breadth can be achieved without extraordinary levels of somatic hypermutation. While some NAbs appear to require years of maturation[1,3,43,44], we show that a V1V2-directed B cell lineage can acquire HIV-1 neutralization breadth within months rather than years. The critical event appears to be an uncommon gene rearrangement that produces a B-cell receptor with protruding, tyrosine-sulfated, anionic CDR H3. Identifying features of antigens able to engage naive B cells with such CDR H3s is a critical step in design of vaccines targeting V1V2. Such antigens could be screened for binding to the UCA versions of NAbs as an indicator of the ability to engage an appropriate naïve B cell receptor. This work also suggests that although an appropriate trimeric V1V2 construct may elicit neutralizing V1V2 antibodies, sequential immunogens that mirror viral evolution may be needed to drive the development of breadth. Overall, the precise delineation of the developmental pathway for the CAP256-VRC26 lineage should provide a basis for attempts to elicit broad V1V2-directed HIV-1-neutralizing antibodies.

Full Methods

Study subject

CAPRISA participant CAP256 was enrolled into the CAPRISA Acute Infection study [51] that was established in 2004 in KwaZulu-Natal, South Africa for follow-up and subsequent identification of HIV seroconversion. CAP256 was one of the 7 women in this cohort who developed neutralization breadth [6]. The CAPRISA 002 Acute Infection study was reviewed and approved by the research ethics committees of the University of KwaZulu-Natal (E013/04), the University of Cape Town (025/2004), and the University of the Witwatersrand (MM040202). CAP256 provided written informed consent for study participation. Samples were drawn between 2005–2009.

Isolation and expression of CAP256-VRC26 family genes

PBMC isolated from CAP256 blood draws at weeks 59, 119, and 206 were stained and sorted for IgG+ B cells on a FACS Aria II as described in [18]. Cells were plated at 2 B cells/well in 384 well plates and cultured for 14 days in the presence of IL-2, IL-21, and CD40L-expressing irradiated feeder cells, as described in[27,45]. Culture supernatants were screened by microneutralization as described in[52] against HIV-1 ZM53.12 and CAP45.G3 Env-

pseudoviruses. Kappa and lambda light chain gene and IgG heavy chain gene variable regions were amplified from neutralization-positive wells, subcloned, expressed, and purified as described in [18]. Heavy chains were reconstituted as IgG1. The efficiency of cloning was as follows. For week 59, 15000 B cells (7500 wells) were plated, 8.3% of wells produced IgG, 4 were positive in microneutralization, and one heavy-light chain pair was recovered. For week 119, 45000 B cells were plated, 48% of wells produced IgG, 49 wells were positive in microneutralization, and 8 heavy-light chain pairs were recovered. For week 206, 42000 B cells were plated, 29% of wells produced IgG, 34 wells were positive in microneutralization, and 3 heavy-light chain pairs were recovered.

The antibodies are numbered CAP256-VRC26.01-.12 in order of the timepoint of the sample from which they were isolated, and then the degree of heavy chain somatic mutation.

Neutralization assays

Single round of replication Env-pseudoviruses were prepared, titered, and used to infect TZM-bl target cells as described previously[46,47]. Neutralization breadth of CAP256-VRC26.01, 03, .06, and .08 were determined using a previously described [18,53] panel of 194 geographically and genetically diverse Env-pseudoviruses representing the major subtypes and circulating recombinant forms. The remaining antibodies were assayed on a subset of this panel. The data were calculated as a reduction in luminescence units compared with control wells, and reported as 50% inhibitory concentration (IC50) in in micrograms per microlitre for monoclonal antibodies, or reciprocal dilution (ID50) for plasma samples.

Neutralization fingerprints

Due to the high sequence variability of HIV-1 Env, different viral strains may exhibit different neutralization sensitivities to the same antibody, and this pattern of neutralization variation can be used to define the neutralization fingerprint for a given antibody. Namely, the neutralization fingerprint of an antibody is defined as the rank-order of neutralization potencies for the antibody against a set of diverse viral strains[18].

The correlations between the neutralization fingerprints of the CAP256-VRC26 antibodies and the neutralization patterns of four longitudinal serum timepoints (at 59, 106, 159, and 220 weeks post infection) were computed over a set of 29 HIV-1 strains (6535.3, AC10.29, CAAN.A2, CAP210.E8, CAP244.D3, CAP45.G3, DU156.12, DU172.17, DU422.01, PVO. 04, Q168.a2, Q23.17, Q259.d2.17, Q461.e2, Q769.d22, Q842.d12, QH0692.42, REJO.67, RHPA.7, SC422.8, THRO.18, TRJO.58, TRO.11, WITO.33, ZM109.4, ZM135.10a, ZM197.7, ZM233.6, ZM53.12) [18]. The correlations between the neutralization potencies of the CAP256-VRC26 antibodies and a reference set of antibodies targeting the four major sites of vulnerability, with at most two antibodies per unique donor, were computed over a set of 41 HIV-1 strains (6535.3, 0260.v5.c36, 6405.v4.c34, AC10.29, C1080.c3, CAAN.A2, CAP210.E8, CAP244.D3, CAP45.G3, CNE3, DU156.12, DU172.17, DU422.01, KER2008.12, KER2018.11, MB201.A1, MB539.2B7, PVO.04, Q168.a2, Q23.17, Q259.17, Q461.e2, Q769.d22, Q842.d12, QH0692.42, REJO.67, RHPA.7, RW020.2, SC422.8, TH976.17, THRO.18, TRJO.58, TRO.11, UG037.8, WITO.33, ZM109.4, ZM135.10a, ZM197.7, ZM214.15, ZM249.1, ZM53.12). The correlations between the neutralization

patterns of the four longitudinal serum timepoints and the neutralization fingerprints of the reference antibodies were computed over a set of 28 HIV-1 strains (6535.3, AC10.29, CAAN.A2, CAP210.E8, CAP244.D3, CAP45.G3, DU156.12, DU172.17, DU422.01, PVO.04, Q168.a2, Q23.17, Q259.17, Q461.e2, Q769.d22, Q842.d12, QH0692.42, REJO.67, RHPA.7, SC422.8, THRO.18, TRJO.58, TRO.11, WITO.33, ZM109.4, ZM135.10a, ZM197.7, ZM53.12). For the reference antibodies, data from multiple neutralization experiments was averaged and consolidated. All correlations are based on the Spearman rank coefficient.

Virus-like particle ELISA

VLP ELISAs were performed as described previously [31]. Briefly, VLPs were produced by PEI-based cotransfection of 293T cells with a pCAGGS-based, Env-expressing plasmid and the Env-deficient HIV-1 genomic backbone plasmid pNL-LucR-E. VLPs were coated on ELISA wells at 20x the concentration in transfection supernatants. MAb binding was then assessed by ELISA, omitting detergent in PBS wash buffers and probing with an anti-human Fc alkaline phosphatase conjugate (Accurate, Westbury, NY) and SigmaFAST p-nitrophenyl phosphate tablets (Sigma). Plates were read at 405 nm.

Cell-surface Env Binding

293T cells were transiently transfected with plasmids encoding Env ZM53.12 or CAP256-SU with deletions of the cytoplasmic tail [30]. For binding experiments, after 2 days, the cells were stained with ViVid viability dye (Invitrogen) followed by serial dilutions of antibodies, two washes, then R-PE-conjugated F(ab) goat anti-human IgG specific for the Fc fragment (Jackson ImmunoResearch) at a 1:200 dilution[54]. For competition assays, the cells were stained with ViVid viability dye followed by biotinylated CAP256-VRC26.01 (10 μg/ml) or CAP256-VRC26.08 (0.8 μg/ml) premixed with serially diluted unlabeled competitor antibodies. After incubation and 2 washes, cells were stained with streptavidin-PE (Invitrogen) at 1:200 dilution. Cells were analyzed on a BD LSRII (Becton Dickinson). Binding was measured as the median fluorescence intensity (MFI) for each sample minus the MFI of cells stained with secondary antibody only.

Polyreactivity analysis of antibodies

Antibody binding to cardiolipin was determined as in [42]. Briefly, using the QUANTA Lite ACA IgG III ELISA kit (Zeus Scientific) per manufacturer's protocol, each antibody was diluted to 100 μg/ml in the kit sample diluent and tested in 3-fold serial dilutions. Results shown are representative of at least two independent ELISAs. Positive and negative controls were included on each plate, and values three times above background were considered positive. Antibody reactivity to a human epithelial cell line (HEp-2) was determined with the ANA/HEp-2 Cell Culture IFA Test System (Zeus Scientific) per manufacturer's protocol, as described in[42]. Antibodies were diluted to 50 μg/ml and 25 μg/ml in ZOBRA-NS diluent. Positive and negative controls were included on each slide. Antibodies were scored negative, indeterminate, or positive (1+ to 4+) at each dilution. Results are representative of at least two independent experiments.

Electron Microscopy (EM) and Image Processing

VRC26.09 Fabs in complex with BG505 SOSIP.664 gp140 produced in HEK 293S cells were analyzed by negative stain EM. A 3 uL aliquot of ~8 ug/ml of the complex was applied for 15s onto a glow discharged, carbon coated 400 Cu mesh grid and stained with 2% uranyl formate for 20s. Grids were imaged using a FEI Tecnai T12 electron microscope operating at 120 kV using a 52,000 × magnification and electron dose of 25 e-/Å2, which resulted in a pixel size of 2.05 Å at the specimen plane. Images were acquired with a Tietz 4k × 4k CCD camera in 5° tilt increments from 0° to 50° at a defocus of 1000 nm using LEGINON[55].

Particles were picked automatically by using DoG Picker and put into a particle stack using the Appion software package [56,57]. Initial reference free 2D class averages were calculated using particles binned by 2 via the Xmipp Clustering 2D Alignment and sorted into 128 classes[58]. Particles corresponding to the complexes were selected into a substack and another round of reference free alignment was carried out with unbinned particles using Xmipp Clustering 2D alignment and IMAGIC softwares [59]. To generate an ab initio 3D starting model, a template stack of 44 images of 2D class averages was used without imposing symmetry. The resulting starting model was refined against 2D class averages for 9 cycles and subsequently with 6,763 raw particles for 9 cycles using EMAN[60]. The resolution of the final reconstruction was calculated to be 28Å using an FSC cut-off value of 0.5.

High-throughput sequencing

Amplicon for 454 next-generation sequencing was prepared as described [12,14] with slight modifications as indicated. Briefly, mRNA was prepared from 10–15 million PBMC using an Oligotex kit (Qiagen). cDNA was synthesized using Superscript II reverse transcriptase (Invitrogen) and oligo-dT(12–18) primers. Individual PCR reactions were performed with Phusion polymerase for 30 cycles. Primers (Supplementary Table 4) consisted of pools of 5–7 oligonucleotides specific for all lambda gene families or VH3 family genes, and had adapters for 454 next generation sequencing. For week 176 only, heavy-chain PCR was performed with primers for all VH families, and mixed lambda and kappa primers were used for light chain (Supplementary Table 4). PCR products were gel-purified (Qiagen). Pyrosequencing of the PCR products was performed on a GSFLX sequencing instrument (Roche-454 Life Sciences, Bradford, CT, USA) on a half chip per reaction (full chips for week 176). On average, ~250,000 raw reads were produced.

High-throughput linkage of VH and VL transcripts was performed in single cell emulsions generated using a flow focusing apparatus[35] (B.J.D., manuscript in preparation). CD27+ B cells were isolated from CAP256 PBMCs collected at 34, 48, 59, 69, and 119 weeks post-infection by magnetic bead sorting (Miltenyi Biotec, Auburn, CA). Cells from weeks 34 and 119 were divided in two groups and half of the cells were analyzed with FR1 primers[35] while the other half were analyzed with leader peptide primers[41] (Supplementary Table 5). All other time points were analyzed in a single group using only FR1 primers (Supplementary Table 1). Overlap extension RT-PCR was performed as previously reported[35], with extension time increased to 125 seconds. Nested PCR was performed as described previously with a 23-second extension time and PCR products were sequenced

using the Illumina 2×250 bp MiSeq platform. Raw reads were quality-filtered for an Illumina Q-score of 20 in 50% of bases. VRC26-class VH and paired VL sequences were identified via BLAST against CDR-H3 nucleotide sequences of the 12 culture-isolated antibodies.

Antibodyomics pipeline

Raw 454 data was processed using a pipeline implemented in Python, similar to one we reported previously [14]. Briefly, reads were filtered for length, keeping only those between 300 and 600 nucleotides. Germline V genes were then assigned to each read using BLAST with empirically optimized parameters. Reads for which no V gene match was found with an e-value 10-10 were discarded. For reads assigned to any VH3-30 or Vλ1-51 allele, (the CAP256-VRC26 germline genes), ClustalW2 [61] was used to calculate the sequence identity to the germline and each isolated antibody. These data were plotted as density heat maps using ggplot2 in R to produce identity-divergence plots (Fig. 3b and Supplementary Fig. 4).

Finding clonally related sequences

Reads that were assigned to the same V genes as CAP256-VRC26, VH3-30 and Vλ1-51, were submitted to IMGT High-Vquest [62] (http://www.imgt.org/IMGTindex/IMGTHighV-QUEST.html), and the results, including automated sequence corrections, were used to further sieve for lineage-related sequences. Reads assigned to J genes matching CAP256-VRC26 (JH3 or Jλ1), and having similar divergence (+/− 15%) in the V and J genes, similar (+/− 10%) nucleotide and amino acid divergences in the V gene, and containing a continuous open reading frame throughout the entire variable region, were selected for further processing. Next, reads from all time points were pooled and clustered at 97.25% sequence identity (twice the standard deviation of expected 454 sequencing error) [14] using CDHit [63]. For each cluster, a representative sequence was chosen from the earliest possible time point. The choice of cluster representatives from the earliest time points at which they appeared was critical to maintaining information on the chronology of lineage development in subsequent analyses. This procedure yielded 8,485 unique heavy chain and 6,410 unique light chain sequences.

To identify CAP256-VRC26 lineage-member heavy chains, we performed intra-donor phylogenetic analysis [14] on the unique 454 sequence set using the heavy chain sequences of the 12 isolated CAP256-VRC26 antibodies. 707 sequences were identified as likely lineage members, of which 27 were discarded after manual inspection, resulting in a total of 680 unique CAP256-VRC26 lineage heavy chain sequences.

To identify light chain lineage members, a sieve requiring at least 92% sequence identity in CDR L3 to one of the isolated antibodies resulted in 495 sequences. Joinsolver [64] was used to examine the V-J junctions of these sequences in detail, to ensure that the recombination points matched those known for the isolated antibodies (Supplementary Fig. 5). This gave a total of 472 unique CAP256-VRC26 lineage light chain sequences.

Paired reads that were identified as members of the CAP256-VRC26 lineage were clustered using CDHit [63] at 95% sequence identity and consensus VH and VL sequences were generated for each cluster containing two or more pairs. Blast was then used to align the

resulting sequences to all clonally related sequences identified from the 454 sequencing as described above. Gapless alignments covering at least 190 nucleotides at 97% or greater sequence identity were considered to be matches. Two of the 157 paired sequences determined to be members of the CAP256-VRC26 lineage matched known CAP256-VRC26 lineage sequences in both VH and VL 454 data sets. An additional 4 VH sequences and 1 VL sequence were found in the 454 data, but their light or heavy chain partners were not present.

Computation of phylogenetic trees

Phylogenetic trees were constructed from 454 data and the sequences of antibodies isolated from B cell culture. Raw data is shown in Nexus format in Supplementary Figures 12 and 13. MEGA5 [65] was used to select the general time-reversible model with a gamma-distributed rate parameter (GTR+G[66]) as the best mathematical model for building a maximum-likelihood tree from the CAP256-VRC26 lineage sequences. FASTML[67] was then used to estimate the gamma parameter and build separate maximum likelihood trees for heavy and light chain sequences (including the isolated antibodies) and these were rooted on the germline V gene sequences. Two branches of the light chain tree were manually moved to match their positioning in the heavy chain tree based on the evidence from trees constructed solely with the 12 isolated antibodies. Analysis with DNAML from PHYLIP (Phylogeny Inference Package) version 3.6 (Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle) (http://cmgm.stanford.edu/phylip/dnaml.html) showed that these rearrangements did not significantly alter the log-likelihood score of the tree.

To create a condensed version of the heavy chain phylogenetic tree (Fig. 4c), CDR H3 sequences were clustered using a 95% sequence identity threshold and requiring that all CDR H3s in a cluster have the same length. Isolated antibodies and monophyletic clusters with at least five members were represented by a single leaf, while all other sequences were removed from the tree. In cases were an internal node was deleted, branch lengths above and below that node were summed, so that the tree depths of all remaining sequences were maintained.

UCA and inferred intermediates

The phylogenetic trees of all heavy and all light chain lineage members calculated above (Fig. 3c and Extended Data Fig. 1) were input into the DNAML maximum likelihood software package to infer ancestral sequences. These are a direct consequence of the input sequences and the mathematical model used to build the trees; the gamma distribution found by FASTML above was used and the topology of the tree was held fixed, so no further information was added. The calculated heavy chain UCA was identical to the germline VH3-30*18 allele. Although the VH3-30*03 allele is only one nucleotide different from *18, germline sequencing of this donor showed that she carries the *18 allele and not the *03 allele (Cathrine Mitchell, personal communication). The inferred UCA is very similar to low-divergence sequences found in the week 38 data set (Extended Data Fig 6).

To test intermediates in the development of CAP256-VRC26.01, two internal nodes were chosen from the phylogenetic trees to be approximately equally spaced in terms of evolutionary distance and the inferred sequences were retrieved using DNAML. Successful complementation of inferred heavy and light chains for each intermediate suggests that the lineage is well sampled by the 454 data and that the calculated phylogenetic trees successfully capture the coupled evolutionary dynamics of heavy and light chains.

Logograms for CDRH3s were made with Weblogo [68].

Protein Crystallization

VRC26.UCA Fab was prepared by digesting purified IgG with Lys-C at 37°C for 2 hours. The reaction was then quenched by the addition of cOmplete protease inhibitors (Roche). For VRC26.01, VRC26.03, VRC26.04, VRC26.06, VRC26.07 and VRC26.10 Fab preparation, an HRV3C recognition site (GLEVLFQGP) was inserted after Lys_235 and purified IgG was incubated with HRV3C protease overnight at 4 °C. For all, the digested antibodies were passed over Protein A agarose to remove the Fc fragment. The Fab was further purified over a Superdex 200 gel filtration column and concentrated aliquots were stored at −80°C. All Fabs were screened against 576 crystallization conditions using a Cartesian Honeybee crystallization robot. Initial crystals were grown by the vapor diffusion method in sitting drops at 20°C by mixing 0.2 μl of protein complex with 0.2 μl of reservoir solution. Crystals were manually reproduced in hanging drops by mixing 1.0 μl protein complex with 1.0 μl reservoir solution. VRC26-UCA was crystallized with a reservoir solution of 27% PEG 8000 and 0.1M Hepes pH 7.5 and was flash frozen in liquid nitrogen with 20% PEG 400 as a cryoprotectant. VRC26.01 was crystallized with a reservoir solution of 32% PEG 400, 4% PEG 3350 and 0.1M Na Acetate pH 5.5 and was flash frozen in liquid nitrogen with 20% ethylene glycol as a cryoprotectant. VRC26.03 was crystallized with a reservoir solution of 22% PEG 8000, 5% MPD and 0.1M imidazole pH 6.5 and was flash frozen in liquid nitrogen with 20% xylitol as a cryoprotectant. VRC26.04 was crystallized with a reservoir solution of 14% PEG 3350, 25% ispropanol and 0.1M Tris pH 8.5 and was flash frozen in liquid nitrogen with 20% ethylene glycol as a cryoprotectant. VRC26.06 was crystallized with a reservoir solution of 3M Na formate and 0.1M Tris pH 7.5 and was flash frozen in liquid nitrogen with 20% xylitol as a cryoprotectant. VRC26.07 was crystallized with a reservoir solution of 4% PEG 8000, 0.1M Zn acetate and 0.1M MES pH 6 and was flash frozen in liquid nitrogen with 20% glycerol as a cryoprotectant. VRC26.10 was crystallized with a reservoir solution of 22% PEG 4000, 0.4M Na Acetate and 0.1 M Tris pH 7.5 and was flash frozen in liquid nitrogen with no cryoprotectant.

Data for all crystals were collected at a wavelength of 1.00 Å at SER-CAT beamlines ID-22 and BM-22 (Advanced Photon Source, Argonne National Laboratory). All diffraction data were processed with the HKL2000 suite [69] and model building and refinement were performed in COOT [70] and PHENIX [71], respectively. For VRC26.03 Fab data, a molecular replacement solution consisting of one Fab molecule per asymmetric unit was obtained using PHASER with a search model from PDB ID 3F12. VRC26.03 then served as a search model for all remaining VRC26 Fabs. Throughout the refinement processes, a cross validation (Rfree) test set consisting of 5% of the data was used and hydrogen atoms were

included in the refinement model. Structure validations were performed periodically during the model building/refinement process with MolProbity [72]. All graphical representations with protein crystal structures were made with Pymol [73].

Structure modeling on trimers

Defined locations of the V1V2, V3-glycan and CD4-binding sites were mapped directly onto EM density of the unliganded HIV-1 BAL spike (EMD-5019)[50] using the software package UCSF Chimera [74]. The CD4-binding site was defined by aligning density of the VRC01-bound BAL spike (EMD-5457)[75] with the unliganded map and fitting a crystal structure of VRC01-bound gp120 (PDB id 3NGB)[76] to the density. EM density in close proximity to the Fab structure was colored to highlight the region of contact. The same procedure was used to define the V3-glycan region using a PGT128-bound trimer (EMD-1970) and crystal structure (PDB id 3TYG)[77] and the V1V2 region using the PG9-bound BG505 SOSIP trimer (EMD-2241) [24] and a crystal structure of V1V2-bound PG9 (PDB id 3U4E) [22]. The fit of the PG9-V1V2 crystal structure to the SOSIP trimer was used to model the trimeric orientation of V1V2 using the 3-fold symmetry of the HIV-1 spike. The BG505.664 SOSIP crystal structure[33] was presented to highlight the quaternary location of V1V2 point mutations. Side chains of residues 166 and 167, not seen in the crystal structure were modeled. The Man5 glycan at N160, also not seen in the crystal structure, is represented as in the crystal structure of the PG9-V1V2 complex (PDB id 3U4E).

Loop modeling

Two intermediates were calculated at approximately equal maturation distance along the VRC26-UCA to VRC26.01 pathway. Mutations associated with the intermediates were mapped directly onto the structure of VRC26.01. 14 of the 35 residues in the VRC26.01 structure are disordered and were modeled with Loopy [78] (http://wiki.c2b2.columbia.edu/honiglab_public/index.php/Software:Loopy) and represented as grey dots. Mutations of the intermediates were colored according to approximate time of occurrence based on the longitudinal phylogenetic tree highlighting the timeline of the structural development. These, and the other antibodies with modeled loops (Fig. 4), were modeled in a single loop prediction involving four steps. In the first step, Loopy was used to predict 10 loop conformations. The number of initial loop conformations to be sampled was set to 50,000 (and the not the default value of 2,000). In the second step all 10 loop conformations were refined using the Protein Preparation Wizard in Maestro (http://www.schrodinger.com/). In the third step, sulfate groups were added to tyrosine at position 100 of the heavy chain and the entire structure was then subjected to all-atom energy minimization in Maestro. A fourth and final step was needed to ensure a reasonable sampling of the rotameric states for the sulfated tyrosines. The Rapid Torsion Scan module in Maestro was used to sample the chi angle involving the sulfate moiety in steps of 20 degrees. The model with the lowest energy after application of the Rapid Torsion Scan module was considered as the best prediction.

Tyrosine sulfation predictions were carried out in GPS-TPS (Z. Pan et al, http://tsp.biocuckoo.org).

Single Genome Amplification, Sequencing and Cloning

HIV-1 RNA was isolated from plasma using the Qiagen QIAamp Viral RNA kit, and reverse transcribed to cDNA using SuperScript III Reverse Transcriptase (Invitrogen, CA). The envelope genes were amplified from single genome templates [49] and amplicons were directly sequenced using the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA) and resolved on an ABI 3100 automated genetic analyzer. The full-length env sequences were assembled and edited using Sequencher v.4.5 software (Genecodes, Ann Arbor, MI). Multiple sequence alignments were performed using Clustal X (ver. 1.83) and edited with BioEdit (ver. 7.0.9) Sequence alignments were visualized using Highlighter for Amino Acid Sequences v1.1.0 (beta).

For analysis of selection pressure, and to account for recombination between the SU and PI, sequences were partitioned into two alignments (an SU-related, and a PI-related alignment) based on the inferred recombination breakpoints using an in-house script. Breakpoints were identified by a shift in identity from one reference towards the other, and required at least two sequential polymorphisms in common with a corresponding PI/SU-related virus in order to be considered. Phylogenies for both alignments were then reconstructed using FastTree[79] with a GTR+CAT model, and rooted on the PI/SU. Signals of selective pressure were detected with MEME (episodic diversifying selection)[80] and DEPS (directional selection)[81] using the FastTree-generated trees, implemented in Hyphy [82].

The frequencies of specific amino acids at a site and the distribution of net charges in the V2 epitope were calculated from the 2012 filtered web alignment (N=3990) from the Los Alamos HIV database (http://www.hiv.lanl.gov/).

Selected envelope amplicons were cloned into the expression vector pcDNA 3.1 (directional) (Invitrogen) by re-amplification of SGA first-round products using Pfu Ultra II enzyme (Stratagene) with the EnvM primer, 5′-TAG CCC TTC CAG TCC CCC CTT TTC TTT TA-3′ [83] and directional primer, EnvAstop, 5′-CAC CGG CTT AGG CAT CTC CTA TGG CAG GAA GAA-3′ [48]. Cloned env genes were sequenced to confirm that they exactly matched the sequenced amplicon. Autologous clones were mutated at key residues within the C-strand using the Stratagene QuickChange II kit (Stratagene) as described by the manufacturer. Mutations were confirmed by sequencing. Envelope clones were used to generate single round of replication Env-pseudoviruses as described above.

Extended Data

a

b

Extended Data Figure 1. Sequences of CAP256-VRC26 heavy and light chains
Sequences of the 12 B-cell culture derived antibodies, inferred germline V and J genes, and
inferred intermediates are compared to the predicted UCA. **a**, heavy chain. **b**, lambda light
chain.

a

Autologous	IC50	CAP256-VRC26.01	CAP256-VRC26.02	CAP256-VRC26.03	CAP256-VRC26.04	CAP256-VRC26.05	CAP256-VRC26.06	CAP256-VRC26.07	CAP256-VRC26.08	CAP256-VRC26.09	CAP256-VRC26.10	CAP256-VRC26.11	CAP256-VRC26.12
	CAP256.PI	>50	>50	>50	>50	>50	0.030	>50	>50	>50	>50	>50	>50
	CAP256.SU	0.14	0.015	0.004	0.003	0.007	0.055	0.008	0.003	0.002	0.031	0.055	0.032

Heterologous by clade	IC50	CAP256-VRC26.01	CAP256-VRC26.02	CAP256-VRC26.03	CAP256-VRC26.04	CAP256-VRC26.05	CAP256-VRC26.06	CAP256-VRC26.07	CAP256-VRC26.08	CAP256-VRC26.09	CAP256-VRC26.10	CAP256-VRC26.11	CAP256-VRC26.12
C	96ZM651.02	>50	>50	>50	>50	>50	>50	>50	2.630	0.991	>50	>50	>50
	CAP210.E8	0.166	0.062	0.008	0.005	0.013	1.200	0.021	0.003	0.002	0.100	0.073	0.072
	CAP244.D3	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	CAP45.G3	31.8	>50	16.0	>50	>50	>50	>50	4.00	16.0	>50	>50	>50
	DU156.12	>50	>50	0.090	10.4	>50	>50	>50	0.006	0.016	>50	>50	>50
	DU172.17	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	DU422.01	>50	>50	0.224	4.17	>50	>50	>50	0.008	0.023	8.20	11.5	>50
	ZM109.4	>50	>50	>50	>50	26.5	>50	>50	>50	>50	>50	>50	>50
	ZM135.10a	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	ZM197.7	22.1	1.97	0.302	0.254	0.113	>50	>50	0.022	0.011	0.346	0.055	>50
	ZM214.15	>50	>50	0.100	>50	>50	>50	>50	3.23	0.333	>50	>50	>50
	ZM233.6	0.701	0.701	0.032	0.060	0.003	22.600	>50	0.001	0.001	0.228	0.322	>50
	ZM249.1	>50	>50	0.004	11.4	>50	>50	44.1	0.076	0.010	>50	34.5	>50
	ZM53.12	0.036	0.043	0.008	0.004	0.012	0.082	0.014	0.004	0.002	0.072	0.081	0.211
A	0260.v5.c36	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	BG505.w6m	>50	>50	4.760	8.430	>50	>50	>50	0.143	0.054	>50	4.200	>50
	KER2008.12	0.308	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	KER2018.11	5.95	0.18	0.014	0.008	0.012	>50	3.100	42.7	0.002	0.335	1.76	>50
	MB201	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	MB539.2B7	>50	>50	>50	>50	>50	>50	>50	31.9	40.7	>50	>50	>50
	Q168.a2	>50	>50	>50	>50	>50	>50	>50	0.148	0.181	>50	>50	>50
	Q23.17	>50	>50	0.867	23.6	>50	>50	>50	0.911	9.50	2.59	>50	>50
	Q259.d2.17	>50	>50	0.002	0.025	0.023	>50	42.5	0.001	0.008	24.6	>50	>50
	Q461.e2	>50	>50	>50	>50	>50	>50	>50	0.502	0.705	>50	>50	>50
	Q769.d22	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	Q842.d12	>50	>50	>50	>50	>50	>50	>50	>50	10.1	>50	>50	>50
	RW020.2	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	UG037.8	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
B	6535.3	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	AC10.29	>50	>50	>50	>50	>50	0.030	>50	>50	>50	>50	>50	>50
	CAAN.A2	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	PVO.04	>50	7.98	>50	>50	44.400	0.947	>50	11.800	>50	2.48	10.8	>50
	QH0692.42	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	REJO.67	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	RHPA.7	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	SC422.8	>50	>50	>50	>50	>50	0.439	>50	>50	>50	>50	>50	>50
	THRO.18	>50	>50	>50	>50	>50	3.30	>50	>50	>50	>50	>50	>50
	TRJO.58	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	TRO.11	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	WITO.33	>50	>50	>50	>50	>50	0.005	>50	>50	>50	>50	>50	>50
D	191821.E6.1	>50	>50	1.50	>50	>50	>50	>50	0.003	0.054	>50	>50	>50
	3016.V5.c36	16.4	0.057	0.450	0.283	0.011	>50	>50	0.005	0.004	0.017	0.300	7.93
	6405.v4.c34	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
AE	C1080.c3	>50	>50	0.008	0.212	>50	>50	>50	0.023	0.019	>50	0.018	>50
	CM244.ec1	3.32	2.030	0.007	0.361	0.517	>50	6.90	1.00	0.105	0.914	26.7	>50
	CNE3	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	TH976.17	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	% breadth geomean	19%	17%	36%	30%	21%	17%	13%	47%	47%	23%	26%	6%
	IC50	1.88	0.40	0.08	0.32	0.10	0.38	1.51	0.11	0.07	0.60	0.94	0.49

b

Extended Data Figure 2. Neutralization breadth and potency of CAP256-VRC26 antibodies **a**, Neutralization of autologous (CAP256 PI and SU) and 47 heterologous viruses by CAP256-VRC26 antibodies. Neutralization was measured using a TZM-bl assay with Env-pseudoviruses. Geometric mean was calculated for values <50 mg/ml. **b**, Breadth-potency curves. Neutralization of a 194-virus panel was measured for VRC26.08, PG9, PGT145, and CH01. The curves show the percent of viruses neutralized at any given IC_{50}.
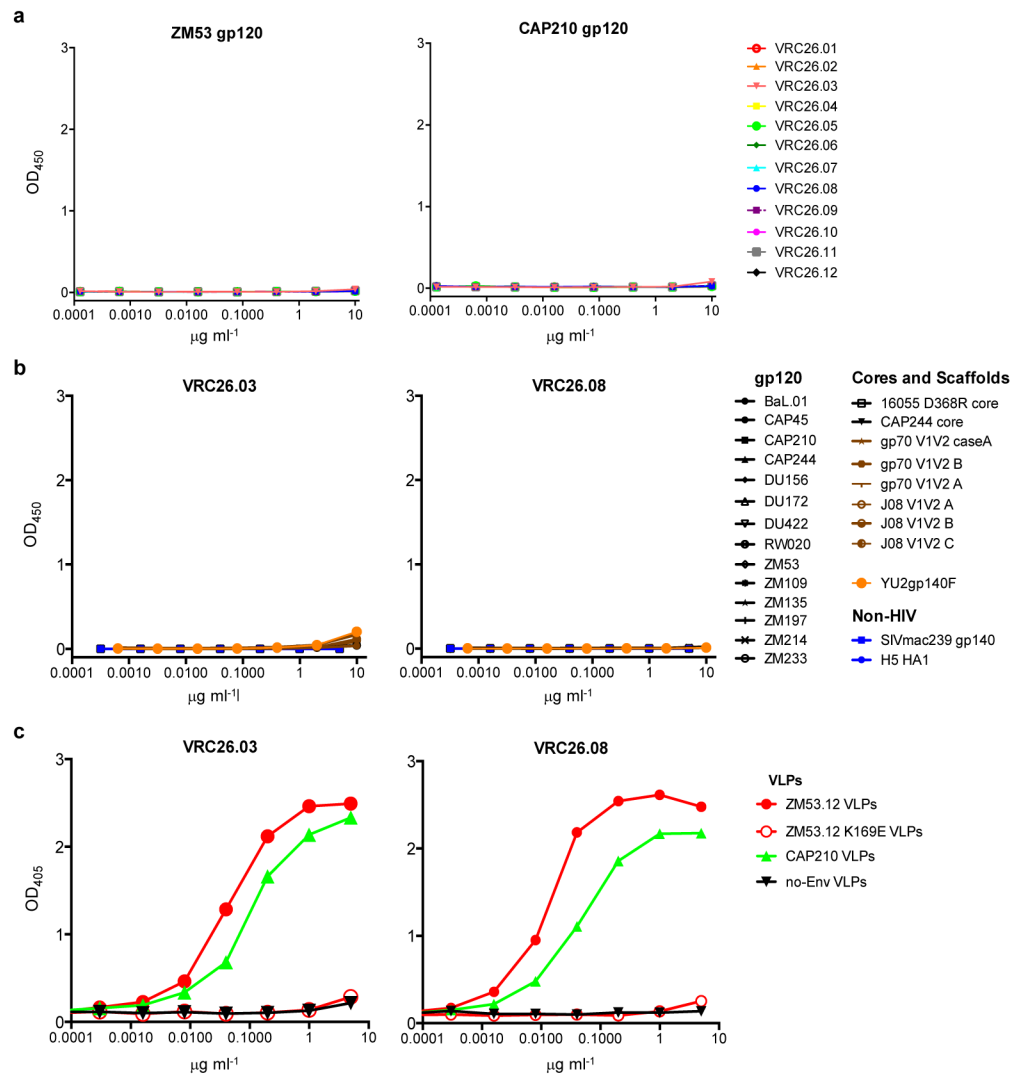
Extended Data Figure 3. CAP256-VRC26 antibodies recognize a quaternary epitope

a, All 12 CAP256-VRC26 mAbs were tested by ELISA against gp120 from ZM53 and CAP210. Positive control antibody PG9 bound to both gp120s (not shown). **b**, 23 proteins and scaffolded V1V2 constructs were tested by ELISA for binding of CAP256-VRC26.03 and CAP256-VRC26.08. PG9 bound to several of these (not shown). Similar data were observed for CAP256-VRC26.06, .07 and .09. **c**, Binding of CAP256-VRC26.03 and CAP256-VRC26.08 to virus-like particles (VLP). VLP expressing ZM53, ZM53.K169E, CAP210, or no Env were concentrated by pelleting and used to coat ELISA plates; assays were performed without detergent to preserve the trimer spikes. Similar data were observed for CAP256-VRC26.06, .07 and .09.
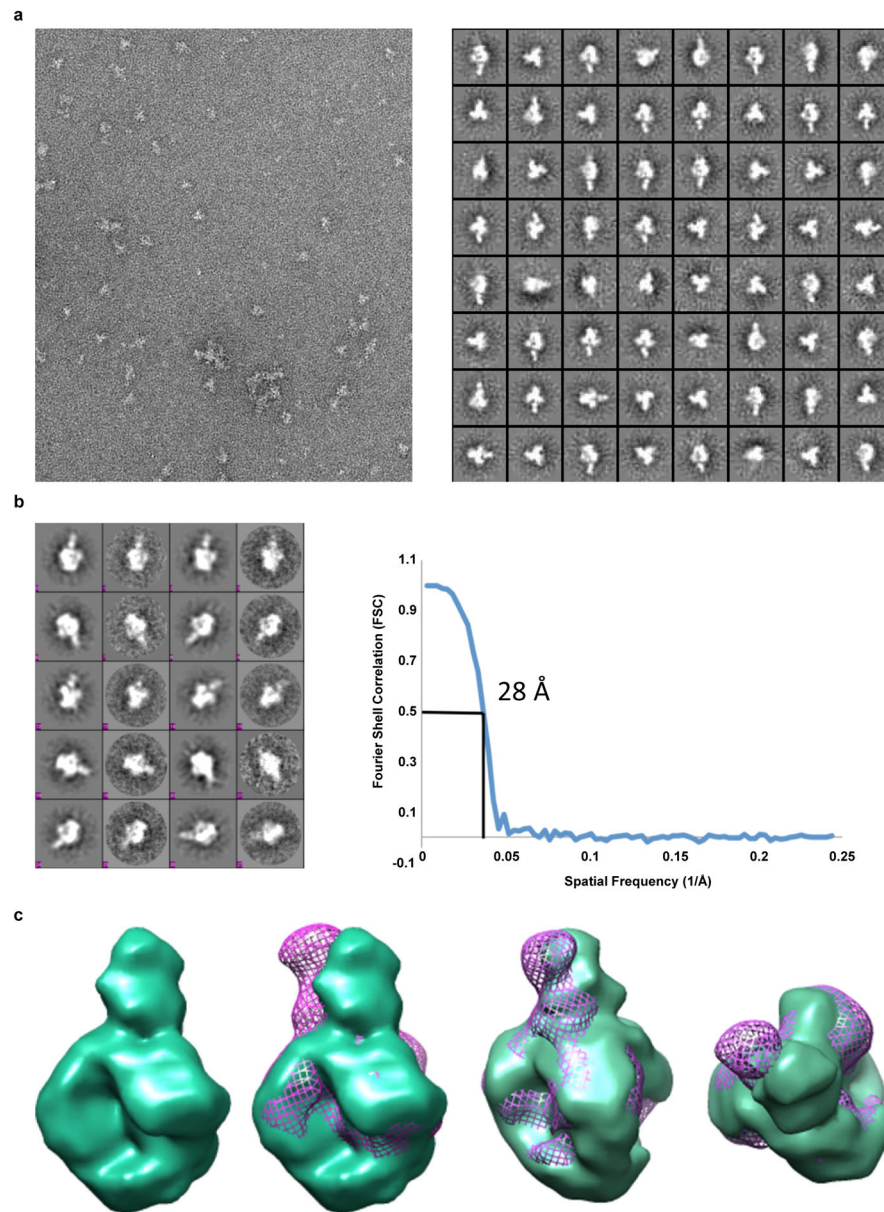
Extended Data Figure 4. Visualization of CAP256-VRC26.09 bound to Env trimers by negative-stain electron microscopy

a, Raw micrograph and corresponding reference free 2D class averages of VRC26.09 in complex with cleaved soluble BG505 SOSIP.664 gp140 trimers. **b**, Projection matching of 3D model refinement and FSC curve used to calculate resolution. Resolution, 28 Å at FSC=0.5. **c**, 3D reconstruction of VRC26.09:BG505 SOSIP.664 complex (green surface) alone and overlayed with PG9:SOSIP (purple mesh). The reconstructions are nearly identical in the trimer portion while displaying small differences in the Fab angles.
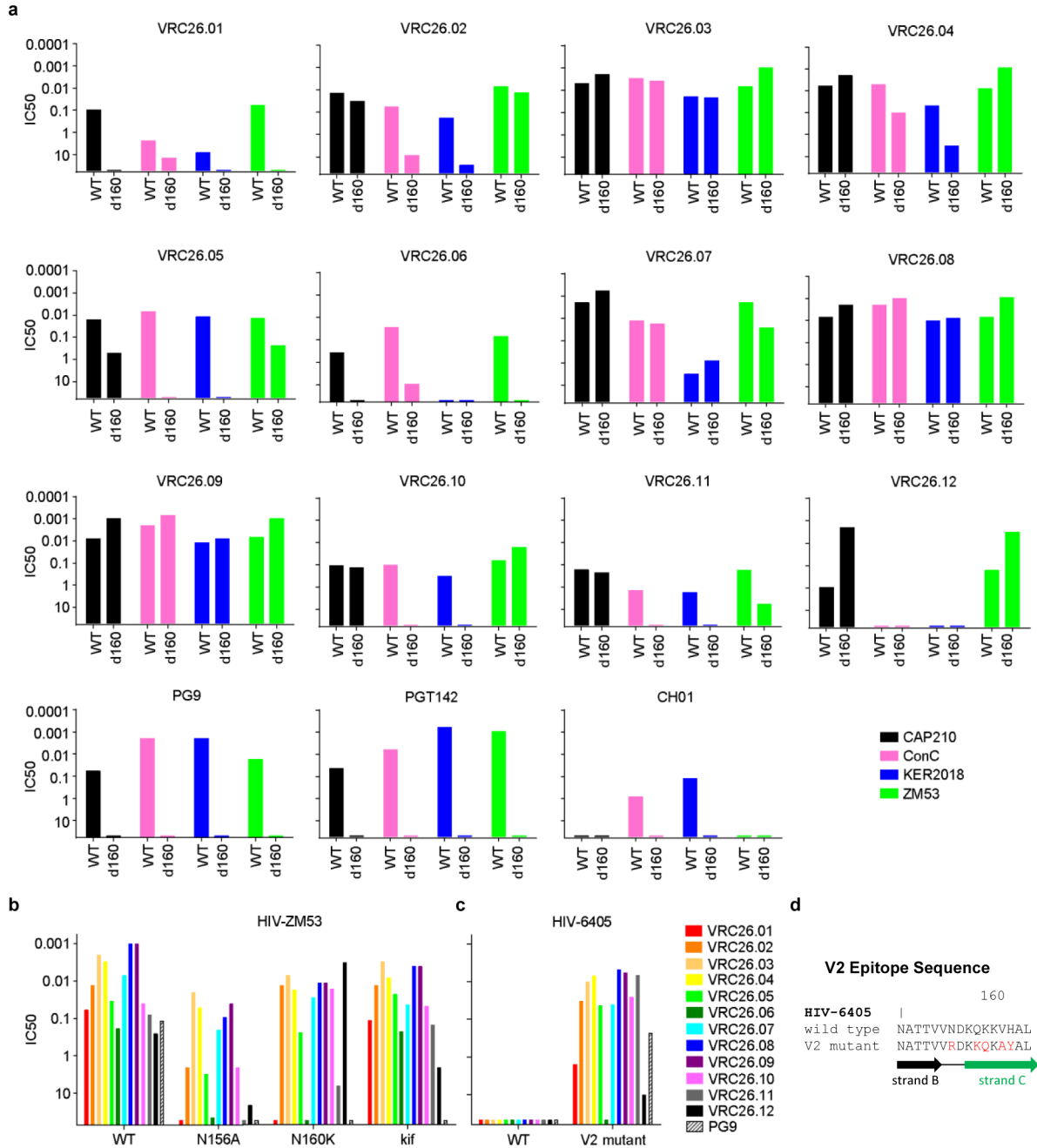
Extended Data Figure 5. Effects of V2 mutations on neutralization activity of CAP256-VRC26 antibodies

a, Each panel shows neutralization of wild-type and N160 glycan mutant CAP210.E8, ConC, KER2018.11, and ZM53.12 viruses. CAP256-VRC26 mAbs are partially and variably affected by loss of N160 glycan, in a virus-strain specific manner. In contrast, PG9-class antibodies PG9, PGT142, and CH01 are uniformly knocked out by N160 mutation. **b,** CAP256-VRC26 mAbs are partially and variably affected by changes in V2 glycans. Neutralization by each antibody was measured against wild-type ZM32.12, mutants N156A and N160K, and ZM53.12 grown in the presence of kifunensine, an inhibitor of glycan

processing. In contrast to CAP256-VRC26 antibodies, PG9 activity is knocked out by the mutations and by kifunensine. **c**, HIV-6405 wild type is resistant to PG9 and CAP256-VRC26 antibodies, while a PG9-sensitive mutant[34] is also sensitive to CAP256-VRC26 antibodies. **d**, Sequences of wild type and mutant HIV-6405.

a

```
                                              CDR H1                        CDR H2
                                        _____                  _____
UCA              QVQLVESGGG VVQPGRSLRL SCAASGFTFS SYGMHWVRQA PGKGLEWVAV ISYDGSNKYY ADSVKGRFTI
038-190676 (1x)  .......... .......... .......... .......... .......... .......... ..........
038-195559 (1x)  .......... .....G.... .......S.. .......... .......... .......... ..........
038-006719 (35x) .......... .......... .......S.. .......... .......... .......... ..........
038-223402 (1x)  .......... .......... .......S.. .......... .......... .......... ..........
038-226205 (1x)  .......... .I........ .......S.. .......... .......... .......... ..........
```

```
                                                      CDR H3
                                        _____
UCA              SRDNSKNTLY LQMNSLRAED TAVYYCAKDL GESENEEWAT DYYDFSIGYP GQDPRGVVGA FDIWGQGTMV TVSS
038-190676       .......... .......... .......... .........S .....R.... .......... .......... ....
038-195559       .......... .......... ..L....... .........S .....GRD.. .........T .......... ....
038-223402       .......... .......... ..L....... .........S .....GRD.. .........T .......... ....
038-223402       .......... .......... ..L-...... .........S .....GRD.. .........T .......... ....
038-226205       .......... .......... ..L....... .........S ....GRD.. .........T .......... ....
```

b

Anti-cardiolipin activity

- CAP256-VRC26-UCA
- CAP256-VRC26.01
- CAP256-VRC26.02
- CAP256-VRC26.03
- CAP256-VRC26.04
- CAP256-VRC26.05
- CAP256-VRC26.06
- CAP256-VRC26.07
- CAP256-VRC26.08
- CAP256-VRC26.09
- CAP256-VRC26.10
- CAP256-VRC26.11
- CAP256-VRC26.12
- VRC01
- 4E10

c

Antibody	Hep2 staining
CAP256-VRC26-UCA	-
CAP256-VRC26.01	
CAP256-VRC26.02	-
CAP256-VRC26.03	
CAP256-VRC26.04	
CAP256-VRC26.05	-
CAP256-VRC26.06	
CAP256-VRC26.07	-
CAP256-VRC26.08	
CAP256-VRC26.09	-
CAP256-VRC26.10	
CAP256-VRC26.11	-
CAP256-VRC26.12	
VRC01	-
4E10	+

d

All VH reads   VH3 reads

- □ CAP256, wk 30
- ▲ CAP256, wk 176
- ● HIV Uninfected Donors

CDRH3 length

Extended Data Figure 6. Origins of long CDRH3s in donor CAP256

a, Week 38 sequences from 454 that support the calculation of the UCA. Unique amino acid sequences with 2–5 residue changes in the CDR H3 are compared to the calculated UCA sequence. Each contained fewer than 3 combined nucleotide mutations in VH and JH. Parentheses, number of corresponding reads in the raw 454 data. **b**, **c**, Lack of autoreactivity. **b**, ELISA for binding to cardiolipin. 4E10 was strongly positive, CAP256-VRC26.03 was weakly positive, and the other 11 CAP256-VRC26 mAbs and the UCA were negative along with control antibody VRC01. **c**, Staining on Hep2 cells was assessed at 50 and 25 mg/ml. Only the positive control, mAb 4E10, showed positive staining. **d**, Distribution of CDRH3

lengths among 454 sequencing reads of B cell transcripts. The percentage of high-quality NGS reads that have CDR H3 ≤24 or ≥28 are shown for three HIV-1 uninfected donors (solid circles on both right and left plots) and for donor CAP256 (week 176) amplified with all-VH primers donor, and CAP256 (week 30) amplified with VH3 primers. High-quality reads are defined as successful V and J assignments and a continuous open reading frame. CDRH3 lengths use the IMGT definitions.

a

c

	VRC26.03	
IC50	wild type	- disulphide
CAP256-PI	>50	>50
CAP256-SU	0.004	3.1
3016.v5.v45	0.118	>50
CAP210.E8	0.011	>50
CM244.ec1	0.000	>50
KER2008.12	>50	>50
KER2018.11	0.003	>50
Q259.17	0.014	>50
ZM197.7	0.160	>50
ZM233.6	0.032	>50
ZM53.12	0.008	>50

b

```
VRC26.01         CAKDVGDYKSDEWGTEYYDISISYPIQDPRAMVGAFDLW
VRC26.03         CAKDLREDECEEWWSDYYDFGKQLPCRKSRGVAGIFDGW
VRC26.03_mutant  CAKDLREDESEEWWSDYYDFGKQLPIRKSRGVAGIFDGW
```

Extended Data Figure 7. Loss of flexibility at the base of the CDR H3

a, Top, logograms of CDR H3 sequences extracted from the heavy chain phylogenetic tree from weeks 59 and 119. The height of each letter is proportional its frequency in the population. Sequences that lack a disulfide bond contain a highly conserved glycine at the

3rd position of the CDR H3 (residue 97, Kabat definition). The appearance of the two cysteines that form the disulfide bond coincides with a glycine to arginine mutation at this site. (Bottom) Overlay and close-up of crystal structures from Supplementary Figure 6A. Loss of the glycine limits flexibility at the base of the CDR H3 and is shown in the crystal structures to be the initial site of divergence in the CDR H3 loops between the antibodies without the disulfide bond (UCA and CAP256-VRC26.01) and those with it (CAP256-VRC26.03, .04, .06, .07, .10). This mutation may contribute to the conserved trajectory of the CDR H3 protrusion towards the heavy chain that is seen in the more mature antibody structures. **b**, CDRH3 and flanking sequences for VRC26.01, VRC26.03, and a mutant VRC26.03 in which the conserved cysteines are changed to the corresponding amino acids found in VRC26.01. **c**, neutralization activity of VRC26.03 and the mutant shown in panel b. The mutant shows reduced activity against CAP256 SU and complete loss of heterologous activity.

a, Primary infecting virus (PI) — N160 glycan, 165, 169; Prevalence within all HIV (n = 3,990): 9%, 11%, 8%.

Superinfecting virus (SU) — N160 glycan, 165, 169; Prevalence within all HIV (n = 3,990): 91%, 34%, 41%.

b, % of viruses (n = 3,990) vs Net charge of V2 epitope (residues 160 – 171). PI, SU, and 176 wk indicated.

c, Infecting viruses and effect of polymorphisms in V2 epitope

Env clone	Residues 160 - 171	Neutralization IC$_{50}$ titers (µg ml^{-1})											
PI	NTITEVRDKQKK	>50	>50	>50	>50	>50	0.03	>50	>50	>50	>50	>50	>50
SU	NATTELRDKKKK	0.04	0.008	0.003	0.005	0.009	0.06	0.008	0.004	0.003	0.01	0.03	0.01
SU T162I	NAITELRDKKKK	0.02	0.05	0.002	0.003	1.7	0.04	0.01	0.005	0.004	0.4	0.1	0.1
SU L165V	NATTEVRDKKKK	0.05	0.01	0.003	0.004	0.007	0.05	0.04	0.006	0.008	0.02	0.04	0.05
SU K169Q	NATTELRDKQKK	>50	0.05	0.03	0.06	0.1	0.8	1.1	0.04	0.02	0.10	0.09	>50
		01	02	03	04	05	06	07	08	09	10	11	12

d, Validation of escape mutations

Env clone	Residues 160 - 171	Neutralization IC$_{50}$ titers (µg ml^{-1})												Mutations first detected
SU K169I	NATTELRDKIKK	>50	0.1	0.02	0.05	0.20	0.05	1.4	0.04	0.02	0.08	0.1	>50	34 weeks
SU K171N	NATTELRDKKKN	0.1	0.03	0.004	0.01	0.03	0.16	0.02	0.005	0.004	0.03	0.02	>50	94 weeks
SU R166K	NATTELKDKKKK	>50	>50	4.8	16	>50	>50	>50	0.1	0.2	>50	>50	>50	
SU R166S	NATTELSDKKKK	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	176 weeks
SU K169E	NATTELRDKEKK	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	
		01	02	03	04	05	06	07	08	09	10	11	12	

e, Reversion of I169 escape mutation in a 34 week clone

34 wk clones	Residues 160 - 171	Neutralization IC$_{50}$ titers (µg ml^{-1})											
34 wk wt	NATTELRDKIKK	>50	0.05	0.02	0.7	0.1	0.2	1	0.05	0.03	0.08	0.04	>50
34 wk I169K	NATTELRDKKKK	0.05	<0.02	<0.02	<0.02	<0.02	<0.02	<0.02	<0.02	<0.02	<0.02	<0.02	<0.02

Column labels: CAP256-VRC26.01, CAP256-VRC26.02, CAP256-VRC26.03, CAP256-VRC26.04, CAP256-VRC26.05, CAP256-VRC26.06, CAP256-VRC26.07, CAP256-VRC26.08, CAP256-VRC26.09, CAP256-VRC26.10, CAP256-VRC26.11, CAP256-VRC26.12
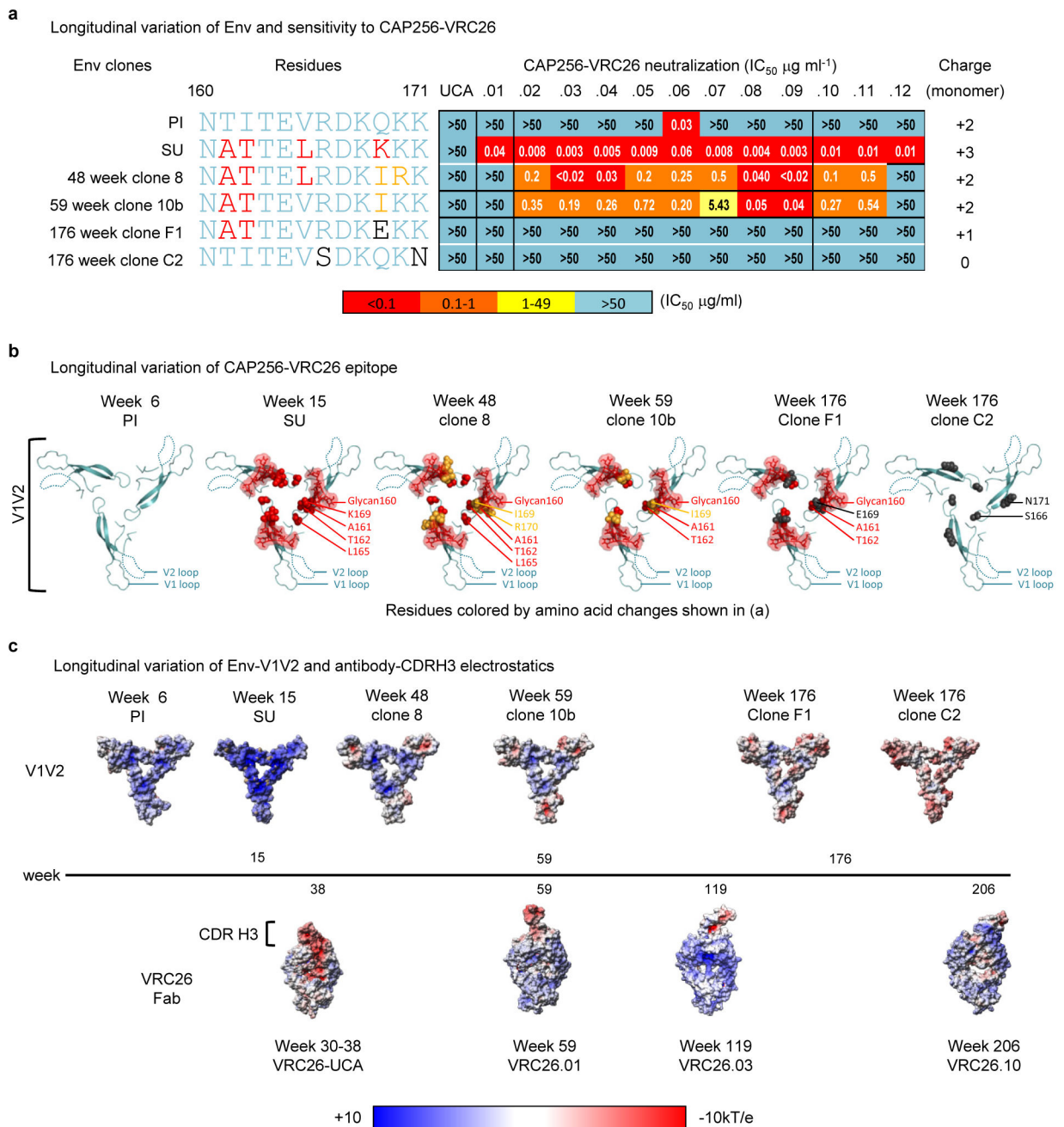
Extended Data Figure 8. Viral polymorphisms and escape mutations
a, Frequency of CAP256 PI and SU polymorphisms at positions 160-162 (glycosylation sequon), 165, and 169. Colored slices on pie charts and percentages indicate prevalence of these polymorphisms within global circulating viruses in the Los Alamos Sequence Database (n=3,990). **b,** Distribution of net charge of the V2 epitope, defined as residues 160 – 171, within global circulating viruses (n=3,990). The charge of the PI, SU and 176 week clones are indicated. **c,** CAP256-VRC26 mAb neutralization of the SU and PI viruses, and of the SU virus mutated to contain PI polymorphisms 162I, 165V or 169Q. **d,** CAP256-VRC26 mAb neutralization of the SU virus mutated to contain known CAP256 escape mutations in the V2 epitope. **e.** CAP256-VRC26 mAb neutralization of 34 week clone (designated wildtype, wt) with an SU-like V1V2, compared to the I169K back mutant. **(c–e)**

The V2 epitope sequence, with mutated residues in red is shown on the left, IC_{50} values in the middle, and the time point when mutations were first detected in Env sequences on the right (weeks post-infection).

a, Longitudinal variation of Env and sensitivity to CAP256-VRC26

Env clones	Residues (160–171)	UCA	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10	.11	.12	Charge (monomer)
PI	NTITEVRDKQKK	>50	>50	>50	>50	>50	>50	0.03	>50	>50	>50	>50	>50	>50	+2
SU	NATTELRDKKKK	>50	0.04	0.008	0.003	0.005	0.009	0.06	0.008	0.004	0.003	0.01	0.01	0.01	+3
48 week clone 8	NATTELRDKIRK	>50	>50	0.2	<0.02	0.03	0.2	0.25	0.5	0.040	<0.02	0.1	0.5	>50	+2
59 week clone 10b	NATTEVRDKIKK	>50	>50	0.35	0.19	0.26	0.72	0.20	5.43	0.05	0.04	0.27	0.54	>50	+2
176 week clone F1	NATTEVRDKEKK	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	+1
176 week clone C2	NTITEVSDKQKN	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	0

CAP256-VRC26 neutralization (IC_{50} μg ml^{-1})

<0.1 | 0.1-1 | 1-49 | >50 (IC_{50} μg/ml)

b, Longitudinal variation of CAP256-VRC26 epitope

Week 6 PI — Week 15 SU — Week 48 clone 8 — Week 59 clone 10b — Week 176 Clone F1 — Week 176 clone C2

Residues colored by amino acid changes shown in (a)

c, Longitudinal variation of Env-V1V2 and antibody-CDRH3 electrostatics

Week 6 PI — Week 15 SU — Week 48 clone 8 — Week 59 clone 10b — Week 176 Clone F1 — Week 176 clone C2

Week 30-38 VRC26-UCA — Week 59 VRC26.01 — Week 119 VRC26.03 — Week 206 VRC26.10

+10 to -10kT/e

Extended Data Figure 9. Longitudinal changes in CAP256 V1V2
a, Variation in the V1V2 sequence of six Env clones. Amino acid mutations from residues 160-171 are highlighted and corresponding changes in neutralization for the six Env clones by CAP256-VRC26.01-.12 and the UCA are shown. The charge of the displayed sequences that make up the central region of the trimer are shown on the right. **b,** Residue changes

highlighted in a were mapped onto the V1V2 domain in the crystal structure of the HIV-1 BG505.664 SOSIP Env trimer. The structure is viewed looking towards the viral membrane along the trimer axis. Mutations are colored as in panel a and represented as spheres (amino acids) or stick and surface (glycan). **c**, Electrostatic surface representations of (top row) the full V1V2 region for each Env clone, (bottom row) Fabs. Timeline of infection is shown in the middle. V1V2 sequences were modeled with SWISS-MODEL using the BG505.664 SOSIP as a template. Escape mutations R166S, K171N, and K169E resulted in a net charge change in the V2 epitope from +3 (SU) to a rare 0. Antibody CDR H3s became less negatively charged over time, suggesting co-evolution of the viral epitope and the antibody paratope.

Extended Data Table 1

Genetic characteristics of CAP256-VRC26 antibodies and V1V2-directed broadly neutralizing antibodies from other donors.

Data are from the present study and from references (Walker 2009, Walker 2011, Bonsignori2011JVI). CAP256-VRC26.01-12 are derived from B cell culture, while CAP256.VRC26-I1 and –I2 (in italics) are inferred intermediates. CDRH3 lengths use Kabat notation. **a**, nucleotides. **b**, amino acids.

a

Donor	Antibody	VH gene	VL gene	% Divergence, Nucleotide			
				from VH	from VL	from UCA-H	from UCA-H CDRH3
CAP256	*VRC26-I1*	IGHV3-30*18	IGLV1-51*02	2%	4%	3%	8%
	VRC26-I2			6%	7%	7%	13%
	VRC26.01			8%	4%	11%	18%
	VRC26.02			9%	5%	13%	27%
	VRC26.03			9%	7%	13%	28%
	VRC26.04			9%	8%	13%	28%
	VRC26.05			10%	5%	13%	24%
	VRC26.06			11%	7%	16%	30%
	VRC26.07			12%	8%	14%	26%
	VRC26.08			12%	10%	17%	33%
	VRC26.09			14%	10%	19%	34%
	VRC26.10			12%	4%	15%	26%
	VRC26.11			12%	14%	18%	32%
	VRC26.12			15%	8%	18%	27%
IAVI24	PG9	IGHV3-33*05	IGLV2-14*01	12%	8%		
	PG16			15%	12%		
CH0219	CH01	IGHV3-20*01	IGKV3-20*01	16%	11%		
	CH02			15%	14%		
	CH03			14%	11%		
	CH04			14%	11%		
IAVI84	PGT141	IGHV1-8*01	IGKV2-28*01	16%	13%		

a

Donor	Antibody	VH gene	VL gene	% Divergence, Nucleotide			
				from VH	from VL	from UCA-H	from UCA-H CDRH3
	PGT142			16%	13%		
	PGT143			16%	13%		
	PGT144			17%	12%		
	PGT145			17%	17%		

b

Donor	Antibody	VH gene	VL gene	CDRH3 Length	% Divergence, Amino Acid			
					from VH	from VL	from UCA-H	from UCA-H CDRH3
	VRC26-I1			35	6%	5%	8%	20%
	VRC26-I2			35	11%	7%	13%	29%
	VRC26.01			35	16%	7%	18%	40%
	VRC26.02			35	16%	7%	27%	46%
	VRC26.03			35	14%	9%	28%	46%
	VRC26.04			35	14%	8%	28%	46%
	VRC26.05			35	20%	8%	24%	40%
CAP256	VRC26.06	IGHV3-30*18	IGLV1-51*02	36	18%	9%	30%	44%
	VRC26.07			35	18%	9%	26%	46%
	VRC26.08			37	16%	9%	33%	51%
	VRC26.09			37	21%	7%	34%	54%
	VRC26.10			35	17%	5%	26%	46%
	VRC26.11			35	23%	23%	32%	49%
	VRC26.12			35	22%	14%	27%	43%
IAVI24	PG9	IGHV3-33*05	IGLV2-14*01	28	20%	15%		
	PG16			28	21%	21%		
CH0219	CH01	IGHV3-20*01	IGKV3-20*01	24	29%	17%		
	CH02			24	22%	23%		
	CH03			24	22%	19%		
	CH04			24	23%	17%		
IAVI84	PGT141	IGHV1-8*01	IGKV2-28*01	32	28%	21%		
	PGT142			32	30%	21%		
	PGT143			32	28%	22%		
	PGT144			32	31%	23%		
	PGT145			31	30%	26%		

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Haynes BF, Kelsoe G, Harrison SC, Kepler TB. B-cell-lineage immunogen design in vaccine development with HIV-1 as a case study. Nature biotechnology. 2012; 30:423–433.

2. Kong L, Sattentau QJ. Antigenicity and Immunogenicity in HIV-1 Antibody-Based Vaccine Design. J AIDS Clinic Res. 2012; S8:003.

3. Mascola JR, Haynes BF. HIV-1 neutralizing antibodies: understanding nature's pathways. Immunological reviews. 2013; 254:225–244. [PubMed: 23772623]

4. Richman DD, Wrin T, Little SJ, Petropoulos CJ. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:4144–4149. [PubMed: 12644702]

5. Wei X, et al. Antibody neutralization and escape by HIV-1. Nature. 2003; 422:307–312. [PubMed: 12646921]

6. Gray ES, 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. Journal of virology. 2011; 85:4828–4840. [PubMed: 21389135]

7. Piantadosi A, et al. Breadth of neutralizing antibody response to human immunodeficiency virus type 1 is affected by factors early in infection but does not influence disease progression. Journal of virology. 2009; 83:10269–10274. [PubMed: 19640996]

8. Sather DN, et al. Factors associated with the development of cross-reactive neutralizing antibodies during human immunodeficiency virus type 1 infection. Journal of virology. 2009; 83:757–769. [PubMed: 18987148]

9. Doria-Rose NA, et al. Breadth of human immunodeficiency virus-specific neutralizing activity in sera: clustering analysis and association with clinical variables. Journal of virology. 2010; 84:1631–1636. [PubMed: 19923174]

10. Glanville J, et al. Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:20216–20221. [PubMed: 19875695]

11. Briney BS, Willis JR, McKinney BA, Crowe JE Jr. High-throughput antibody sequencing reveals genetic evidence of global regulation of the naive and memory repertoires that extends across individuals. Genes and immunity. 2012; 13:469–473. [PubMed: 22622198]

12. Wu X, et al. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. Science. 2011; 333:1593–1602. [PubMed: 21835983]

13. Zhu J, et al. Mining the antibodyome for HIV-1-neutralizing antibodies with next-generation sequencing and phylogenetic pairing of heavy/light chains. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110:6470–6475. [PubMed: 23536288]

14. Zhu J, et al. Somatic Populations of PGT135-137 HIV-1-Neutralizing Antibodies Identified by 454 Pyrosequencing and Bioinformatics. Frontiers in microbiology. 2012; 3:315. [PubMed: 23024643]

15. Liao HX, et al. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. Nature. 2013; 496:469–476. [PubMed: 23552890]

16. Walker LM, et al. A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals. PLoS pathogens. 2010; 6

17. Lynch RM, et al. The B cell response is redundant and highly focused on V1V2 during early subtype C infection in a Zambian seroconverter. Journal of virology. 2011; 85:905–915. [PubMed: 20980495]

18. Georgiev IS, et al. Delineating antibody recognition in polyclonal sera from patterns of HIV-1 isolate neutralization. Science. 2013; 340:751–756. [PubMed: 23661761]

19. Walker LM, et al. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science. 2009; 326:285–289. [PubMed: 19729618]

20. Bonsignori M, et al. Analysis of a clonal lineage of HIV-1 envelope V2/V3 conformational epitope-specific broadly neutralizing antibodies and their inferred unmutated common ancestors. Journal of virology. 2011; 85:9998–10009. [PubMed: 21795340]

21. Walker LM, et al. Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature. 2011; 477:466–470. [PubMed: 21849977]

22. McLellan JS, et al. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature. 2011; 480:336–343. [PubMed: 22113616]

23. Pancera M, et al. Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. Nature structural & molecular biology. 2013; 20:804–813.

24. Julien JP, et al. Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110:4351–4356. [PubMed: 23426631]

25. Moore PL, et al. Potent and broad neutralization of HIV-1 subtype C by plasma antibodies targeting a quaternary epitope including residues in the V2 loop. Journal of virology. 2011; 85:3128–3141. [PubMed: 21270156]

26. Moore PL, et al. Multiple pathways of escape from HIV broadly cross-neutralizing V2-dependent antibodies. Journal of virology. 2013; 87:4882–4894. [PubMed: 23408621]

27. Huang J, et al. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. Nature. 2012; 491:406–412. [PubMed: 23151583]

28. Tiller T, et al. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. Journal of immunological methods. 2008; 329:112–124. [PubMed: 17996249]

29. Kabat, EA.; Wu, TT.; Perry, HM.; Gottesman, KS.; Foeller, C. Sequences of Proteins of Immunological Interest. U.S. Department of Health and Human Service, National Institutes of Health; Bethesda MD: 1991.

30. Pancera M, Wyatt R. Selective recognition of oligomeric HIV-1 primary isolate envelope glycoproteins by potently neutralizing ligands requires efficient precursor cleavage. Virology. 2005; 332:145–156. [PubMed: 15661147]

31. Tong T, Crooks ET, Osawa K, Binley JM. HIV-1 virus-like particles bearing pure env trimers expose neutralizing epitopes but occlude nonneutralizing epitopes. Journal of virology. 2012; 86:3574–3587. [PubMed: 22301141]

32. Lyumkis D, et al. Cryo-EM Structure of a Fully Glycosylated Soluble Cleaved HIV-1 Envelope Trimer. Science. 2013

33. Julien JP, et al. Crystal Structure of a Soluble Cleaved HIV-1 Envelope Trimer. Science. 2013

34. Doria-Rose NA, et al. A short segment of the HIV-1 gp120 V1/V2 region is a major determinant of resistance to V1/V2 neutralizing antibodies. Journal of virology. 2012; 86:8319–8323. [PubMed: 22623764]

35. DeKosky BJ, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. Nature biotechnology. 2013; 31:166–169.

36. Haynes BF, et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. The New England journal of medicine. 2012; 366:1275–1286. [PubMed: 22475592]

37. Rolland M, et al. Increased HIV-1 vaccine efficacy against viruses with genetic signatures in Env V2. Nature. 2012; 490:417–420. [PubMed: 22960785]

38. Overbaugh J, Morris L. The Antibody Response against HIV-1. Cold Spring Harbor perspectives in medicine. 2012; 2:a007039. [PubMed: 22315717]

39. Briney BS, Willis JR, Crowe JE Jr. Human peripheral blood antibodies with long HCDR3s are established primarily at original recombination using a limited subset of germline genes. PloS one. 2012; 7:e36750. [PubMed: 22590602]

40. Lefranc MP, et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. Developmental and comparative immunology. 2003; 27:55–77. [PubMed: 12477501]

41. Wardemann H, et al. Predominant autoantibody production by early human B cell precursors. Science. 2003; 301:1374–1377. [PubMed: 12920303]

42. Haynes BF, et al. Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. Science. 2005; 308:1906–1908. [PubMed: 15860590]

43. Kwong PD, Mascola JR. Human antibodies that neutralize HIV-1: identification, structures, and B cell ontogenies. Immunity. 2012; 37:412–425. [PubMed: 22999947]

44. Burton DR, et al. A Blueprint for HIV Vaccine Discovery. Cell host & microbe. 2012; 12:396–407. [PubMed: 23084910]

45. Huang J, et al. Isolation of human monoclonal antibodies from peripheral blood B cells. Nat Protoc. 2013; 8:1907–1915. [PubMed: 24030440]

46. Shu Y, et al. Efficient protein boosting after plasmid DNA or recombinant adenovirus immunization with HIV-1 vaccine constructs. Vaccine. 2007; 25:1398–1408. [PubMed: 17113201]

47. Montefiori DC. Measuring HIV neutralization in a luciferase reporter gene assay. Methods Mol Biol. 2009; 485:395–405. [PubMed: 19020839]

48. Kraus MH, et al. A rev1-vpu polymorphism unique to HIV-1 subtype A and C strains impairs envelope glycoprotein expression from rev-vpu-env cassettes and reduces virion infectivity in pseudotyping assays. Virology. 2010; 397:346–357. [PubMed: 20003995]

49. Salazar-Gonzalez JF, et al. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. Journal of virology. 2008; 82:3952–3970. [PubMed: 18256145]

50. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S. Molecular architecture of native HIV-1 gp120 trimers. Nature. 2008; 455:109–113. [PubMed: 18668044]

51. van Loggerenberg F, et al. Establishing a cohort at high risk of HIV infection in South Africa: challenges and experiences of the CAPRISA 002 acute infection study. PloS one. 2008; 3:e1954. [PubMed: 18414658]

52. Doria-Rose, N., et al. High throughput HIV-1 microneutralization assay. Protocol Exchange. 2013. http://dx.doi.org/10.1038/protex.2013.069

53. Wu X, et al. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science. 2010; 329:856–861. [PubMed: 20616233]

54. Julien JP, et al. Broadly neutralizing antibody PGT121 allosterically modulates CD4 binding via recognition of the HIV-1 gp120 V3 base and multiple surrounding glycans. PLoS pathogens. 2013; 9:e1003342. [PubMed: 23658524]

55. Suloway C, et al. Automated molecular microscopy: the new Leginon system. Journal of structural biology. 2005; 151:41–60. [PubMed: 15890530]

56. Lander GC, et al. Appion: an integrated, database-driven pipeline to facilitate EM image processing. Journal of structural biology. 2009; 166:95–102. [PubMed: 19263523]

57. Voss NR, Yoshioka CK, Radermacher M, Potter CS, Carragher B. DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. Journal of structural biology. 2009; 166:205–213. [PubMed: 19374019]

58. Sorzano CO, et al. A clustering approach to multireference alignment of single-particle projections in electron microscopy. Journal of structural biology. 2010; 171:197–206. [PubMed: 20362059]

59. van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M. A new generation of the IMAGIC image processing system. Journal of structural biology. 1996; 116:17–24. [PubMed: 8742718]

60. Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. Journal of structural biology. 1999; 128:82–97. [PubMed: 10600563]

61. Larkin MA, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948. [PubMed: 17846036]

62. Alamyar E, Giudicelli V, Li S, Duroux P, Lefranc MP. IMGT/HighV-QUEST: the IMGT(R) web portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high throughput and deep sequencing. Immunome research. 2012; 8:26.

63. Li W, Jaroszewski L, Godzik A. Clustering of highly homologous sequences to reduce the size of large protein databases. Bioinformatics. 2001; 17:282–283. [PubMed: 11294794]

64. Souto-Carneiro MM, Longo NS, Russ DE, Sun HW, Lipsky PE. Characterization of the human Ig heavy chain antigen binding complementarity determining region 3 using a newly developed software algorithm, JOINSOLVER. Journal of Immunology. 2004; 172:6790–6802.

65. Tamura K, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution. 2011; 28:2731–2739. [PubMed: 21546353]

66. Waddell PJ, Steel MA. General time-reversible distances with unequal rates across sites: mixing gamma and inverse Gaussian distributions with invariant sites. Molecular phylogenetics and evolution. 1997; 8:398–414. [PubMed: 9417897]

67. Ashkenazy H, et al. FastML: a web server for probabilistic reconstruction of ancestral sequences. Nucleic acids research. 2012; 40:W580–584. [PubMed: 22661579]

68. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome research. 2004; 14:1188–1190. [PubMed: 15173120]

69. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods in enzymology. 1997; 276:307–326.

70. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

71. Adams PD, et al. Recent developments in the PHENIX software for automated crystallographic structure determination. J Synchrotron Radiat. 2004; 11:53–55. [PubMed: 14646133]

72. Davis IW, Murray LW, Richardson JS, Richardson DC. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. Nucleic acids research. 2004; 32:W615–619. [PubMed: 15215462]

73. DeLano, WL. The PyMOL Molecular Graphics System. DeLano Scientific; San Carlos, CA: 2002.

74. Pettersen EF, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004; 25:1605–1612. [PubMed: 15264254]

75. Tran EE, et al. Structural mechanism of trimeric HIV-1 envelope glycoprotein activation. PLoS pathogens. 2012; 8:e1002797. [PubMed: 22807678]

76. Zhou T, et al. Structural Basis for Broad and Potent Neutralization of HIV-1 by Antibody VRC01. Science. 2010; 329:811–817. [PubMed: 20616231]

77. Pejchal R, et al. A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. Science. 2011; 334:1097–1103. [PubMed: 21998254]

78. Soto CS, Fasnacht M, Zhu J, Forrest L, Honig B. Loop modeling: Sampling, filtering, and scoring. Proteins. 2008; 70:834–843. [PubMed: 17729286]

79. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. PloS one. 2010; 5:e9490. [PubMed: 20224823]

80. Murrell B, et al. Detecting individual sites subject to episodic diversifying selection. PLoS genetics. 2012; 8:e1002764. [PubMed: 22807683]

81. Kosakovsky Pond SL, Poon AF, Leigh Brown AJ, Frost SD. A maximum likelihood method for detecting directional evolution in protein sequences and its application to influenza A virus. Molecular biology and evolution. 2008; 25:1809–1824. [PubMed: 18511426]

82. Pond SL, Frost SD, Muse SV. HyPhy: hypothesis testing using phylogenies. Bioinformatics. 2005; 21:676–679. [PubMed: 15509596]

83. Gao F, et al. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. Journal of virology. 1996; 70:7013–7029. [PubMed: 8794346]
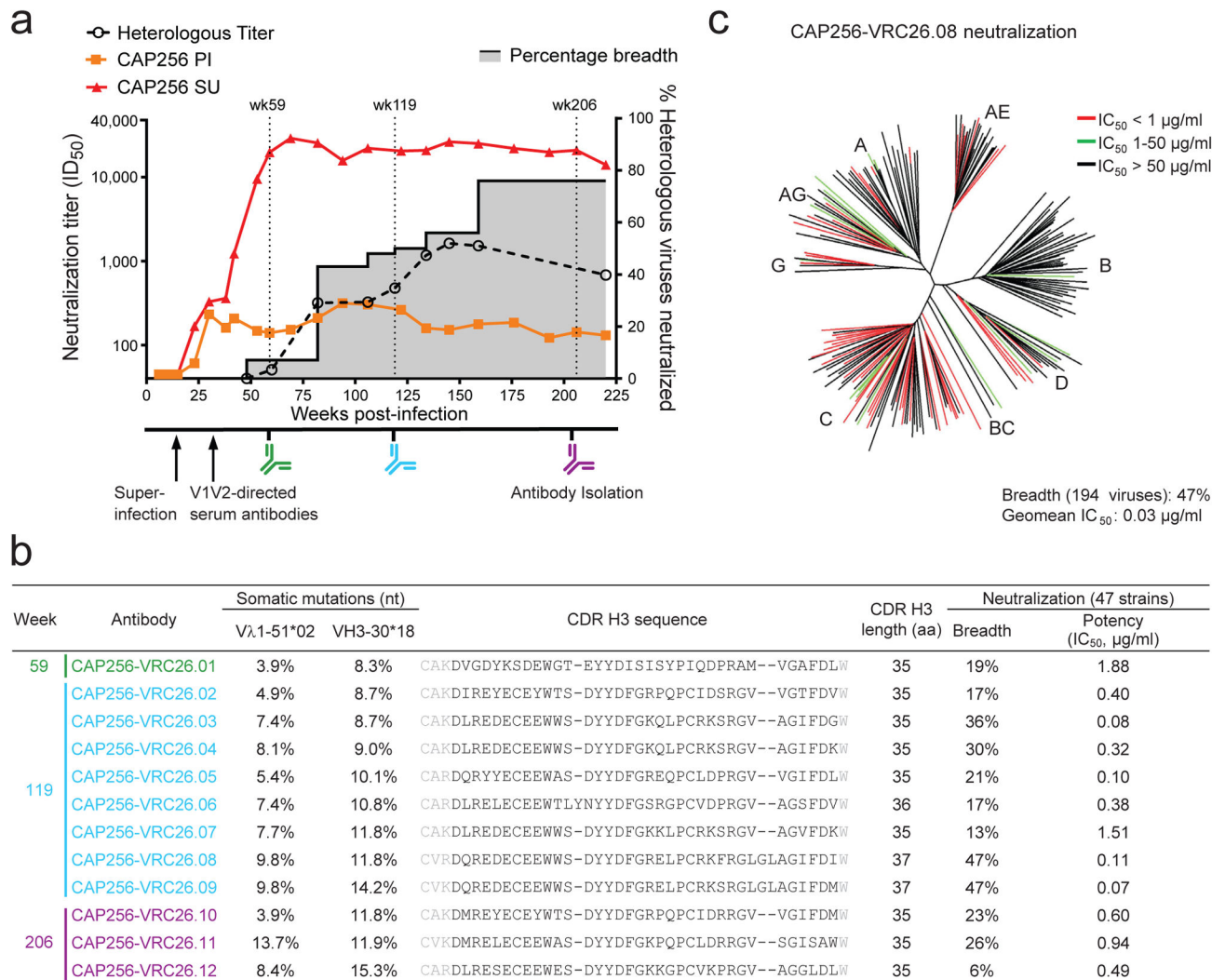
Figure 1. **Development of broad neutralization by donor CAP256 and isolation of neutralizing antibodies**

(a) Timing of antibody isolation in relation to plasma neutralization titers against the primary infecting virus (PI), the superinfecting virus (SU), and a panel of 40 heterologous viruses (geometric mean titer shown). Percentage breadth (gray area), % of viruses neutralized with plasma ID_{50} >45. (b) Genetic characteristics and neutralization breadth and potency of the 12 isolated antibodies. Week of antibody isolation and V-gene mutation rates are indicated. Residues flanking the Kabat-defined CDR H3 sequences are shown in gray. Neutralization was assessed against a panel of 47 heterologous viruses. (c) Breadth and potency of antibody CAP256-VRC26.08 on a panel of 194 Env-pseudoviruses. Dendrogram shows phylogenetic relatedness of Env sequences in the panel.
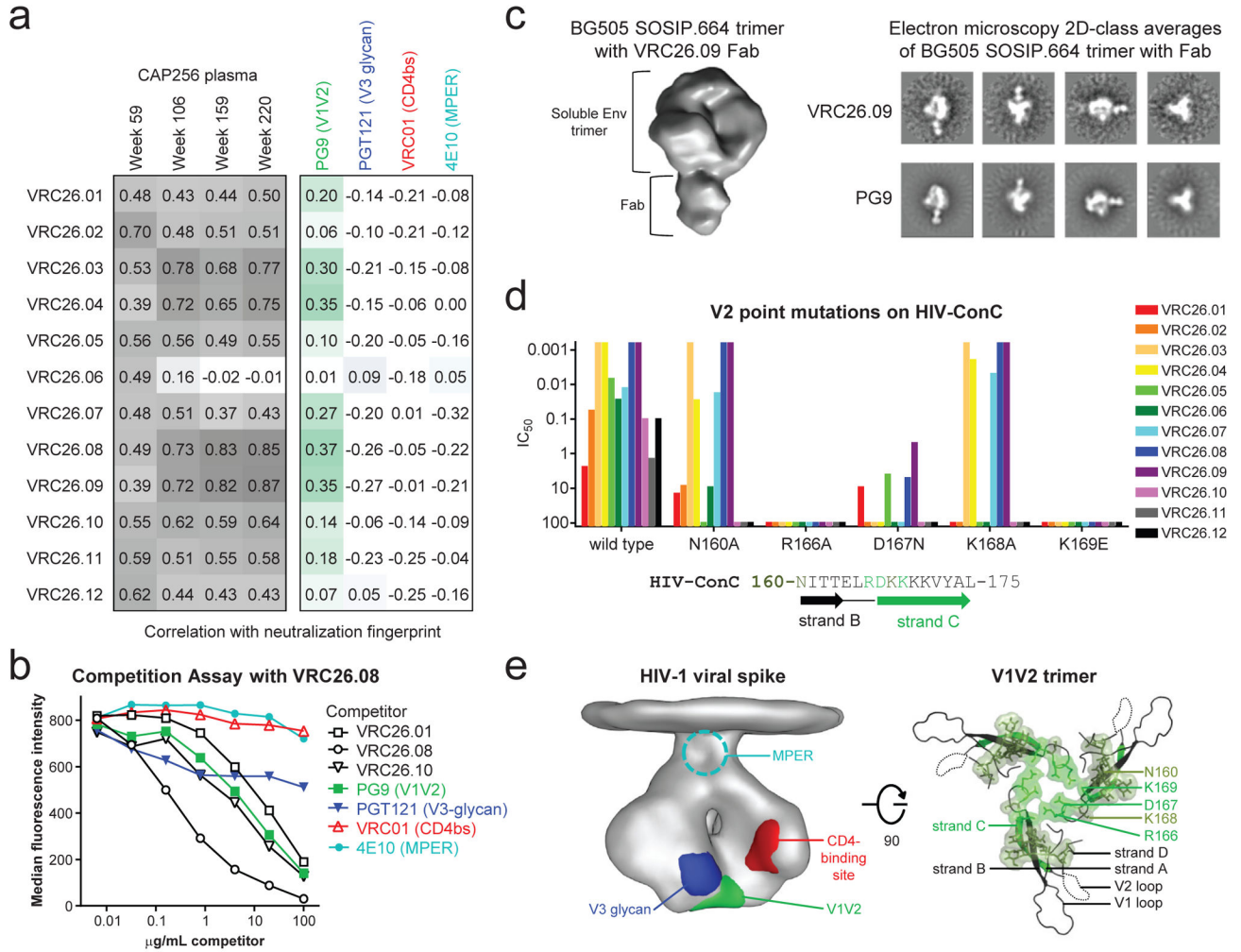
Figure 2. Mapping of CAP256-VRC26 epitope on the HIV-1 Env spike
(a) Left, correlations between neutralization fingerprints (see detailed methods) of CAP256-VRC26 antibodies and CAP256 plasma. Darker gray indicates stronger correlation. Right, correlations between neutralization fingerprints of CAP256-VRC26 antibodies and representative antibodies targeting the major HIV-1 neutralization epitopes. Correlations are color-coded by antibody; darker shades indicate stronger correlations. (b) Competition assay. Binding to ZM53-Env-expressing 293T cells by labeled CAP256-VRC26.08 and unlabeled competitor antibodies measured by flow cytometry. Assay shown is representative of three experiments. (c) Left, negative stain electron microscopy (EM) 3D reconstruction of CAP256-VRC26.09 Fab in complex with soluble cleaved BG505 SOSIP.664 trimer; right, 2D-class averages of VRC26.09 and PG9 in complex with BG505 SOSIP.664 trimer. (d) Neutralization of Env-pseudoviruses with HIV-ConC and V2 point mutants. Sequence shows amino acids 160-175. (e) Location of HIV-1 epitopes. Left, EM density of viral spike[50], with viral membrane at top and major sites of vulnerability shown as determined by structural mapping of antibody interactions[24]. The gp41 membrane proximal external region (MPER) is shown schematically. Right, model of V1V2 based on EM reconstruction of PG9 with BG505 SOSIP.664 trimer[24,32], viewed looking towards the viral membrane along the
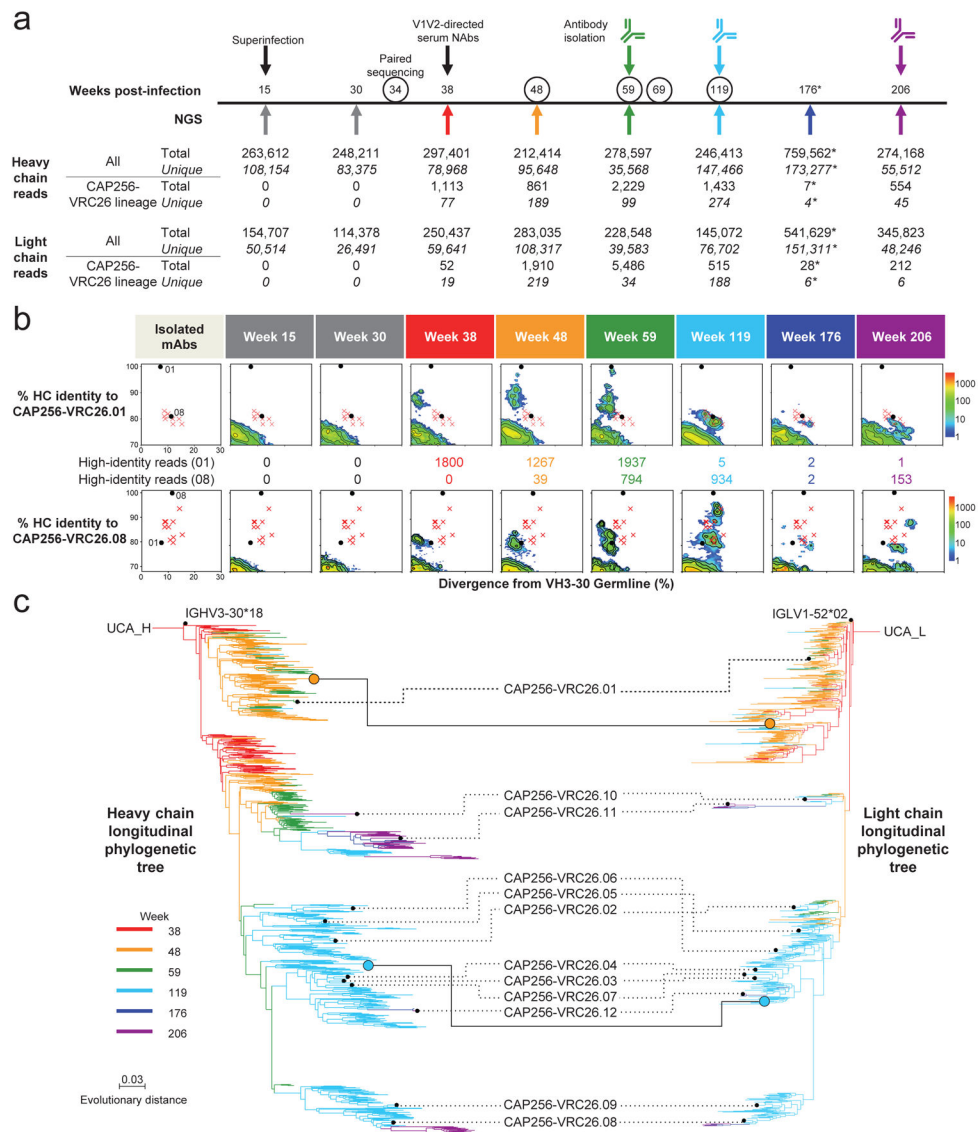
trimer axis. Green ribbon, strand C. V2 mutations from panel **d** are shown with surface representation; brighter green indicates more potent effects on neutralization.
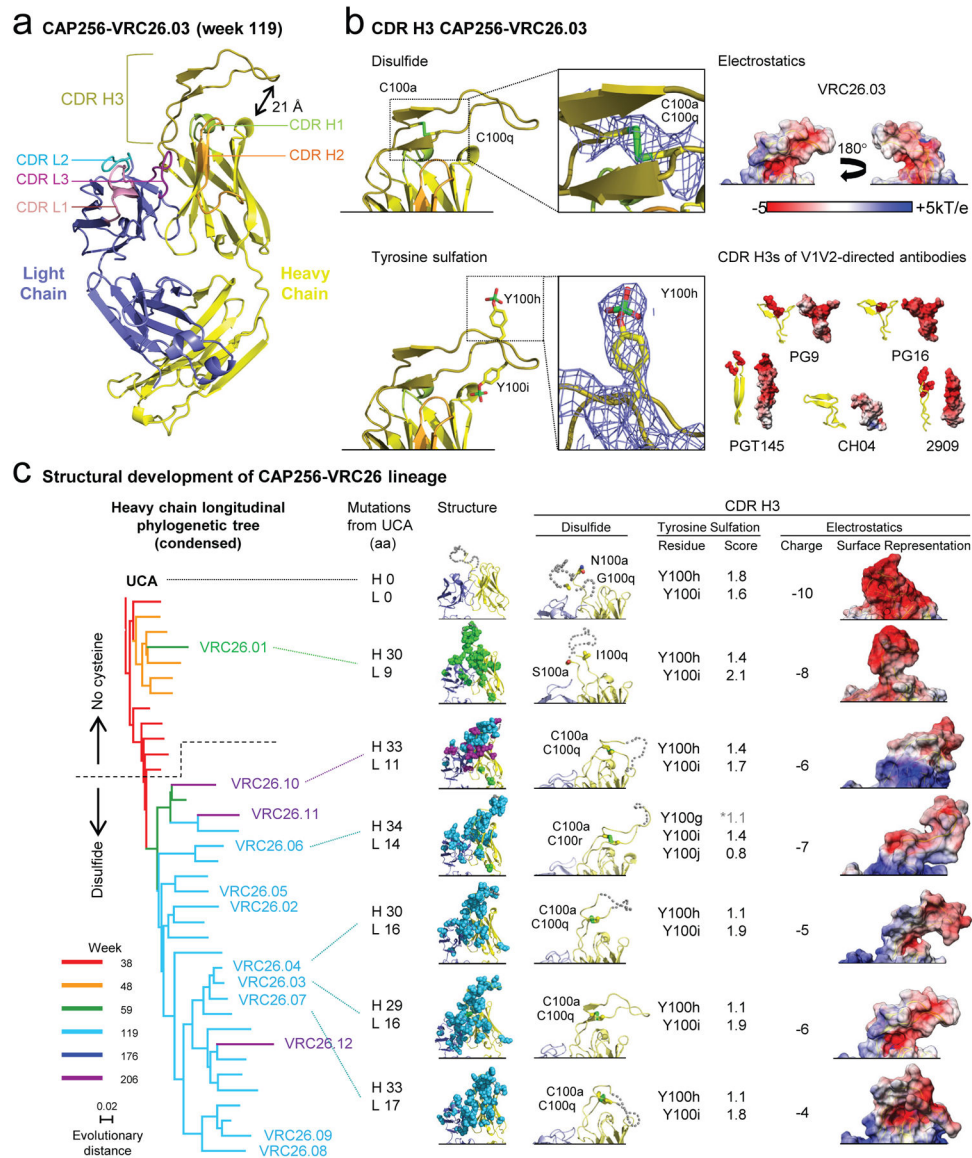
Figure 3. Maturation of the CAP256-VRC26 lineage revealed by NGS and VH:VL paired sequencing of B cell transcripts

(a) Timeline of longitudinal peripheral blood samples with quantification of total NGS sequence reads (total), and CAP256-VRC26 lineage-related reads (total and unique). Arrows below the line indicate time points of 454 pyrosequencing for heavy and light chain sequences. Circles indicate time points of paired sequencing of sorted B cells (see detailed methods). PCR amplifications for pyrosequencing used primers specific for VH3 family sequences (heavy chain) and V lambda sequences (light chain), with the exception of the week 176 sample (asterisk), which was amplified using all-VH gene primers, resulting in fewer CAP256-VRC26 specific reads. (b) Maturation time course for CAP256-VRC26.01 (top panels) and CAP256-VRC26.08 (bottom panels). Heat map plots show sequence identity (vertical axis) versus germline divergence (horizontal axis) for NGS data. The 12 isolated antibodies are displayed as red 'x's for reference, with the exception of the CAP256-VRC26.01 and 08 antibodies which are shown as black dots. Numbers between the

top and bottom panels correspond to the number of raw reads with at least 85% identity to the indicated antibody (top: VRC26.01, bottom: VRC26.08). (c) Phylogenetic trees of the CAP256-VRC26 clonal lineage for heavy chain (left) and light chain (right) were constructed by maximum likelihood using the 454 sequences and the isolated antibodies (black dots, labeled with antibody name). Branches are colored by time point when NGS sequences were first detected. The orange and blue circles indicate linked heavy and light chain sequences from the paired sequencing data. Scale, rate of nucleotide change (per site) between nodes.

Figure 4. Structural characteristics of the developing CAP256-VRC26 lineage
(a) Crystal structure of the antigen-binding fragment (Fab) of CAP256-VRC26.03 shown in ribbon diagram representation. (b) Left, the intra-loop disulfide bond and tyrosine sulfation are shown in stick representation, and enlarged to show electron density (blue mesh, 2Fo-Fc at 1σ). Right, molecular surface, with electrostatic potentials colored red for acidic and blue for basic. CDR H3 regions of broadly neutralizing V1V2-directed antibodies are shown for comparison, with the left image in ribbon representation (tyrosine sulfates highlighted) and the right image in electrostatic representation. (c) Left, a condensed heavy chain phylogenetic tree highlights the isolated antibodies. Scale, rate of nucleotide change between nodes. The number of mutations to the heavy chain (H) and light chain (L) relative to the UCA are shown. Middle, structures of the variable regions. Mutations from the UCA are represented as spheres colored according to the week of antibody isolation at which the mutations first appear. Right, CDR H3 details. Residues that are (or evolve to become)
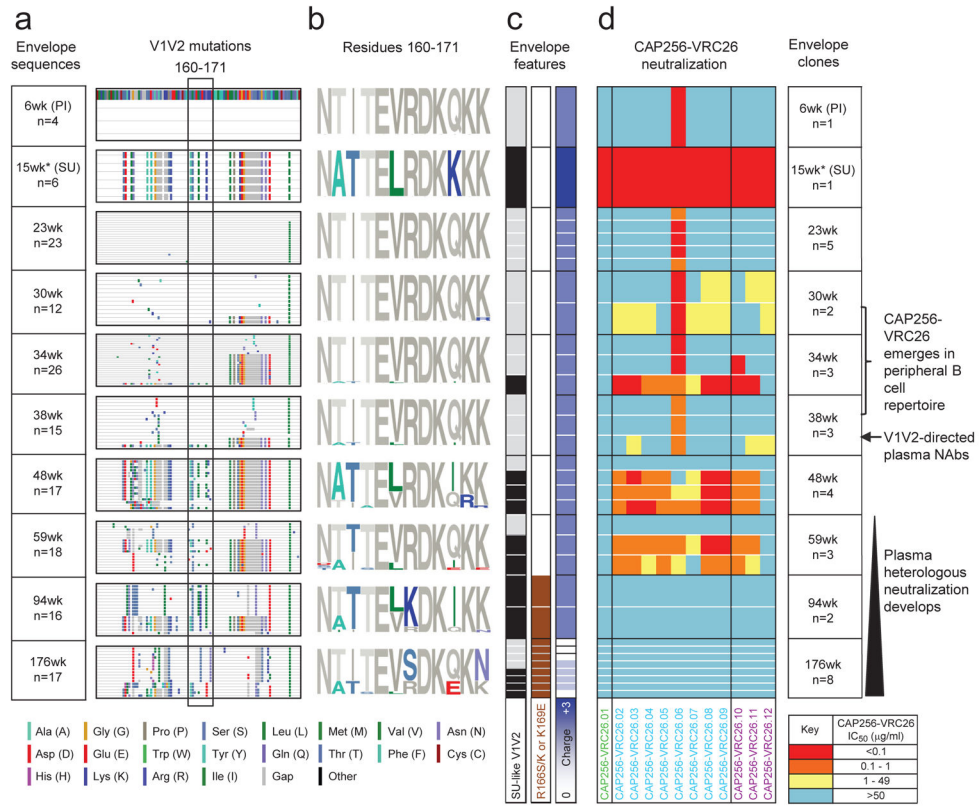
cysteines are labeled (gray dashes indicate modeled disordered regions). The position of tyrosines predicted to be sulfated (scores >1) are noted and were included in the formal charges shown for each CDR H3 and the electrostatic representations (far right).

Figure 5. HIV-1 Env evolution and the development of the CAP256-VRC26 lineage

(a) V1V2 sequences are shown in highlighter format with the primary infecting virus (PI) designated as master and V2 residues 160 to 171 boxed. Asterisk at week 15 denotes sequences amplified with strain-specific primers matching the SU virus. (b) Logogram of the V2 epitope for all CAP256 sequences, with mutations away from the PI (master sequence) in color. (c) SU-like V1V2 sequences are indicated by black (present) and grey (absent) boxes. Escape mutations (K169E or R166S/K) are indicated by brown boxes. The net charge of the V2 epitope (residues 160 to 171) is shown in purple/white, ranging from +3 to 0. White lines separate clones within a time point; black lines separate time points. (d) Neutralization by the 12 CAP256-VRC26 mAbs of representative longitudinal Env clones isolated between 6 and 176 weeks post infection (weeks shown at far right). The CAP256 mAbs are colored by time of isolation (as in Fig. 1). The development of the CAP256-VRC26 antibody lineage, V1V2-directed plasma neutralizing antibodies, and plasma heterologous neutralization, are indicated on the right.
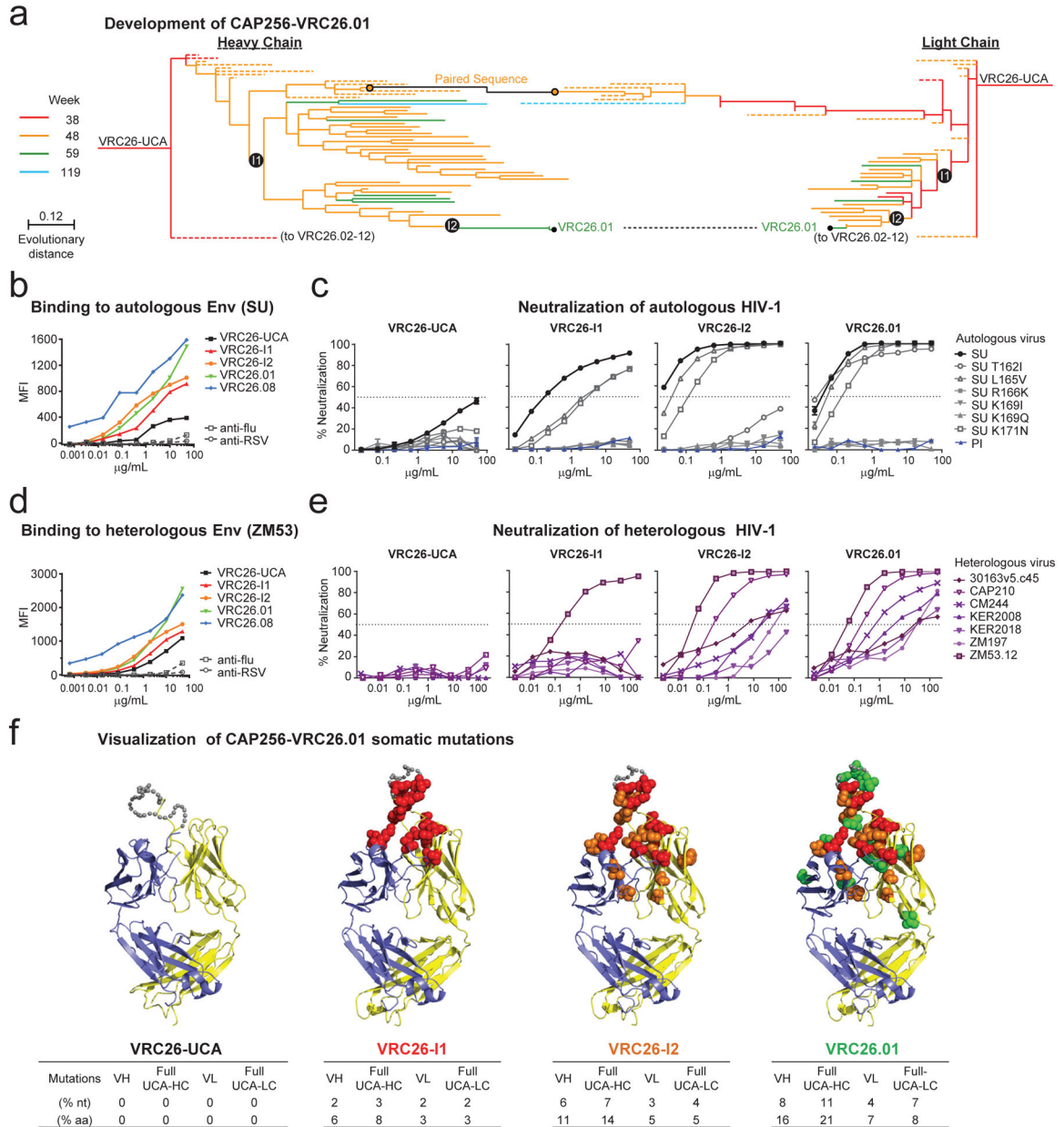
Figure 6. Development from UCA to CAP256-VRC26.01
(a) Expanded view of the phylogenetic trees from Fig. 3c, highlighting the maturation pathway of CAP256-VRC26.01. Off-pathway branches were collapsed and are shown as dashed lines. Inferred intermediates VRC26-I1 and VRC26-I2 were expressed for functional analyses. (b–e) Binding and neutralization of antibodies UCA, VRC26-I1, VRC26-I2, VRC26.01. (b, d), Binding to cell-surface expressed Env (SU and ZM53). MFI, median fluorescence intensity. (c, e) Neutralization of (c) PI, SU and point mutants, and (e) seven heterologous viruses. Bars, standard error of the mean (triplicates). (f) Structural models of VRC26.01 lineage antibodies. Affinity matured residues are shown as spheres colored according to the intermediate at which they first appear: red, VRC26-I1; orange, VRC26-I2; green, VRC26.01. Grey dots, disordered residues in the CDR H3. The number of changes

from the UCA to each intermediate are noted for V gene only (VH or VL), or from the full UCA (UCA-HC or UCA-LC).