

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. DeKosky⁶, Michael J. Ernanandes¹, Ivelin S. Georgiev¹, Helen J. Kim^{7,8,9}, Marie Pancera¹, Ryan P. Staube¹, Han R. Altae-Tran¹, Robert T. Bailer¹, Ema T. Crooks¹⁰, Albert Cupo¹¹, Aliaksandr Druz¹, Nigel J. Garrett⁵, Kam H. Hoi¹², Rui Kong¹, Mark K. Louder¹, Nancy S. Longo¹, Krisha McKee¹, Molati Nonyane³, Sijy O'Dell¹, Ryan S. Roark¹, Rebecca S. Rudicell¹, Stephen D. Schmidt¹, Daniel J. Sheward¹³, Cinque Soto¹, Constantinos Kurt Wibmer^{3,4}, Yongping Yang¹, Zhenhai Zhang², NISC Comparative Sequencing Program†, James C. Mullikin^{14,15}, James M. Binley¹⁰, Rogier W. Sanders¹⁶, Ian A. Wilson^{7,8,9,17}, John P. Moore¹¹, 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. Kwong¹, Lawrence Shapiro^{1,2} & John R. Mascola¹

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 donor CAP256 (from the Centre for the AIDS Programme of Research in South Africa (CAPRISA)); each antibody contained the protruding tyrosine-sulphated, anionic antigen-binding loop (complementarity-determining region (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 and potency 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. These data provide important insights relevant to HIV-1 vaccine development.

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}. Almost 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 neutralizing antibodies, 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 insights can be gained with longitudinal sampling from early after the time of HIV-1 infection¹⁵. 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.

Neutralizing antibodies 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 sulphated^{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 use antibody isolation, B-cell next-generation sequencing, structural characterization, and viral single-genome amplification (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.

¹Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. ²Department of Biochemistry, Columbia University, New York, New York 10032, USA. ³Center for HIV and STIs, National Institute for Communicable Diseases of the National Health Laboratory Service (NHLS), Johannesburg, 2131, South Africa. ⁴Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, 2050, South Africa. ⁵Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Congella, 4013, South Africa. ⁶Department of Chemical Engineering, University of Texas at Austin, Austin, Texas 78712, USA. ⁷Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California 92037, USA. ⁸Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery, The Scripps Research Institute, La Jolla, California 92037, USA. ⁹IAVI Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, California 92037, USA. ¹⁰Torrey Pines Institute, San Diego, California 92037, USA. ¹¹Weill Medical College of Cornell University, New York, New York 10065, USA. ¹²Department of Biomedical Engineering, University of Texas at Austin, Austin, Texas, USA. ¹³Institute of Infectious Diseases and Molecular Medicine, Division of Medical Virology, University of Cape Town and NHLS, Cape Town 7701, South Africa. ¹⁴NISC Comparative Sequencing program, National Institutes of Health, Bethesda, Maryland 20892, USA. ¹⁵NIH Intramural Sequencing Center, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. ¹⁶Department of Medical Microbiology, Academic Medical Center, Amsterdam 1105 AZ, Netherlands. ¹⁷Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, USA. ¹⁸Department of Molecular Biosciences, University of Texas at Austin, Austin, Texas 78712, USA. ¹⁹Department of Epidemiology, Columbia University, New York, New York 10032, USA.

*These authors contributed equally to this work.

†A list of authors and their affiliations appears in the Supplementary Information.

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 PCR with reverse transcription (RT-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 (Kabat²⁹ numbering) (Fig. 1b and Extended Data Fig. 1a). The heavy and light chains exhibited somatic mutation of 4–15% from their germline-encoded *V*-genes, VH3-30 and V λ 1-51, 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). 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 used neutralization fingerprints¹⁸; binding assays for HIV-1 Envelope (Env) in soluble, cell surface³⁰, and viral particle³¹ contexts; and negative stain electron microscopy (EM) of Fab CAP256-VRC26.09 bound to a soluble cleaved version of the HIV-1 trimer^{24,32,33} (Fig. 2a–c, Supplementary Fig. 3 and Extended Data Figs 3 and 4). Recognition of Env by CAP256-VRC26 antibodies was similar to PG9-class neutralizing antibodies that recognize the trimeric V1V2 cap²⁴, 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 neutralizing antibodies^{22,23,34}, although 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 analysed 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 used identity-divergence plots¹² 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 V_H:V_L paired sequencing³⁵ 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), probably the result of VDJ recombination with a single *D*-gene, IgHD3-3*01 and non-templated (N)-nucleotide insertions of 34 and 31 nucleotides at each

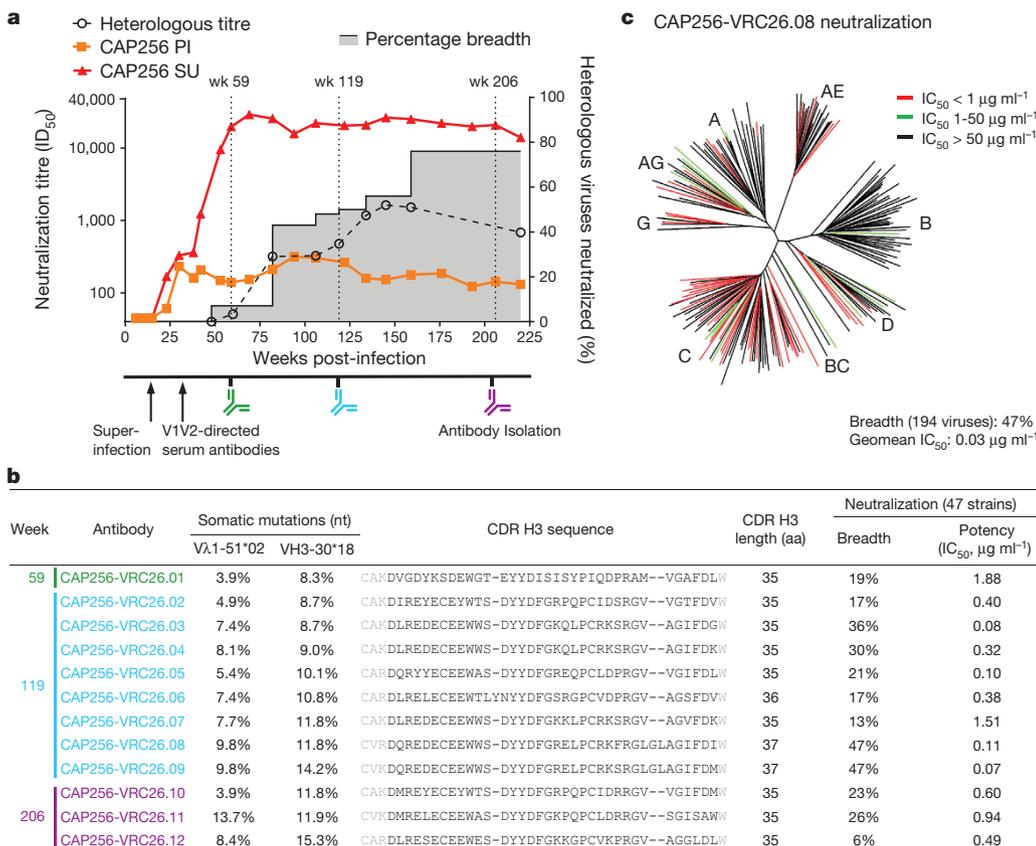


Figure 1 | Development of broad neutralization by donor CAP256 and isolation of neutralizing antibodies. **a**, Timing of antibody isolation in relation to plasma neutralization titres against the primary infecting virus (PI), the superinfecting virus (SU), and a panel of 40 heterologous viruses (geometric mean titre shown). Percentage breadth (grey area), percentage of viruses neutralized with plasma median inhibitory dilution (ID₅₀) > 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 light grey. 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.

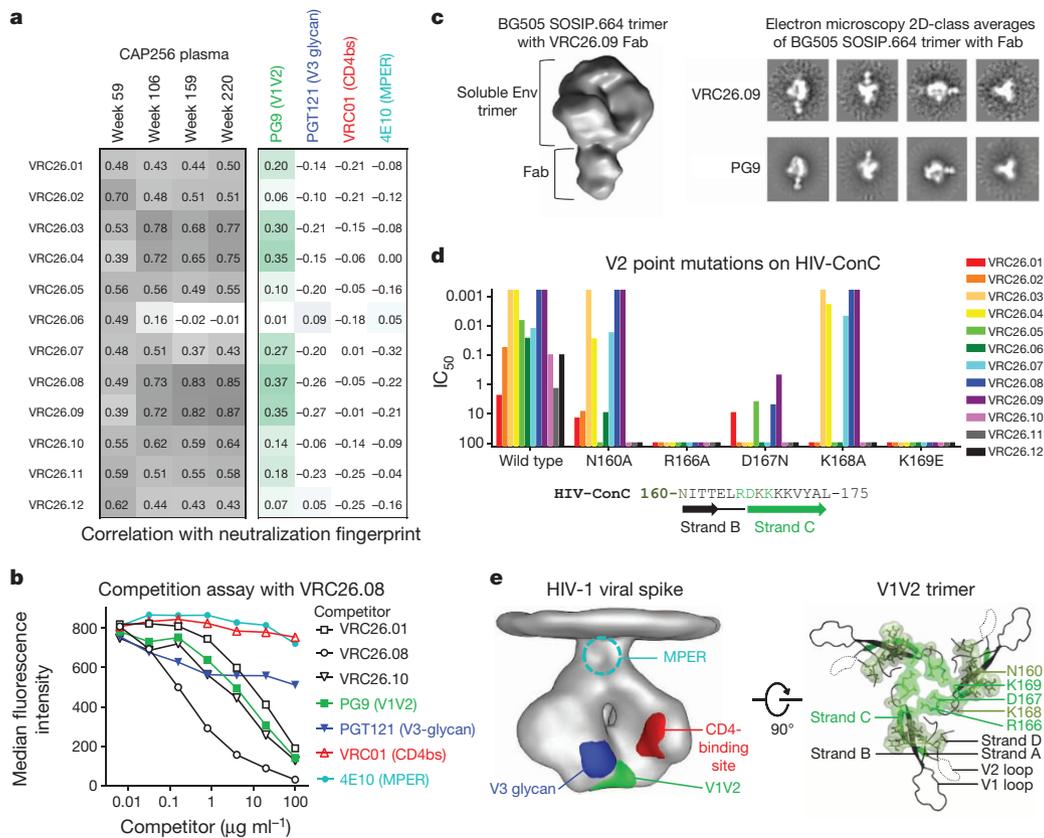


Figure 2 | Mapping of CAP256-VRC26 epitope on the HIV-1 Env spike. **a**, Correlations between neutralization fingerprints (see Methods) of CAP256-VRC26 antibodies and CAP256 plasma (left). Darker grey indicates stronger correlation. Correlations between neutralization fingerprints of CAP256-VRC26 antibodies and representative antibodies targeting the major HIV-1 neutralization epitopes (right). Correlations are colour-coded by antibody; darker shades indicate stronger correlations. **b**, Competition assay. Binding to ZM53-Env-expressing 293T cells by labelled CAP256-VRC26.08 and unlabelled competitor antibodies measured by flow cytometry. Assay shown is representative of three experiments. **c**, Negative stain electron microscopy (EM) 3D reconstruction of CAP256-VRC26.09 Fab in complex with soluble

cleaved BG505 SOSIP.664 trimer (left); 2D-class averages of VRC26.09 and PG9 in complex with BG505 SOSIP.664 trimer (right). **d**, Neutralization of Env-pseudoviruses with HIV-ConC and V2 point mutants. Sequence shows amino acids 160–175. **e**, Location of HIV-1 epitopes. EM density of viral spike⁵⁰, with viral membrane at top and major sites of vulnerability shown as determined by structural mapping of antibody interactions²⁴ (left). The gp41 membrane proximal external region (MPER) is shown schematically. Model of V1V2 based on EM structure of BG505 SOSIP.664 trimer^{24,32}, viewed looking towards the viral membrane along the trimer axis (right). Green ribbon, strand C. V2 mutations from panel **d** are shown with surface representation; brighter green indicates more potent effects on neutralization.

junction (Supplementary Fig. 5). This inferred UCA was further supported by very-low-divergence sequences among the lineage members identified from the week 38 heavy chain data. Five unique sequences were found, all of which had CDR H3s matching the inferred UCA in at least 30 of 35 amino acids while containing three or fewer nucleotide changes in VH and JH combined (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 $\sim 20\text{\AA}$ above the antigen-combining surface of the heavy chain and contained a 2-stranded β -sheet, *O*-sulphated tyrosines, and an intra-CDR H3 disulphide bond (Fig. 4a, b). The CDR H3s of the UCA and CAP256-VRC26.01 lacked a CDR H3 disulphide bond, exhibited greater disorder and were positioned more proximal to the light chain (Fig. 4c); the appearance of the disulphide 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 cysteine 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 neutralizing antibodies. Development over four years involves the introduction of almost 20 light chain and over 30 heavy chain mutations, including a disulphide bond. The CDR H3 changes its overall orientation while losing negative charge and maintaining tyrosine sulphation (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 primary infecting virus and superinfecting virus Env sequences included V2 residues 165 and 169, and an N160 glycan in the superinfecting virus that was not present in the primary infecting virus (Fig. 5b and Extended Data Fig. 8a, b). Notably, compared to the primary

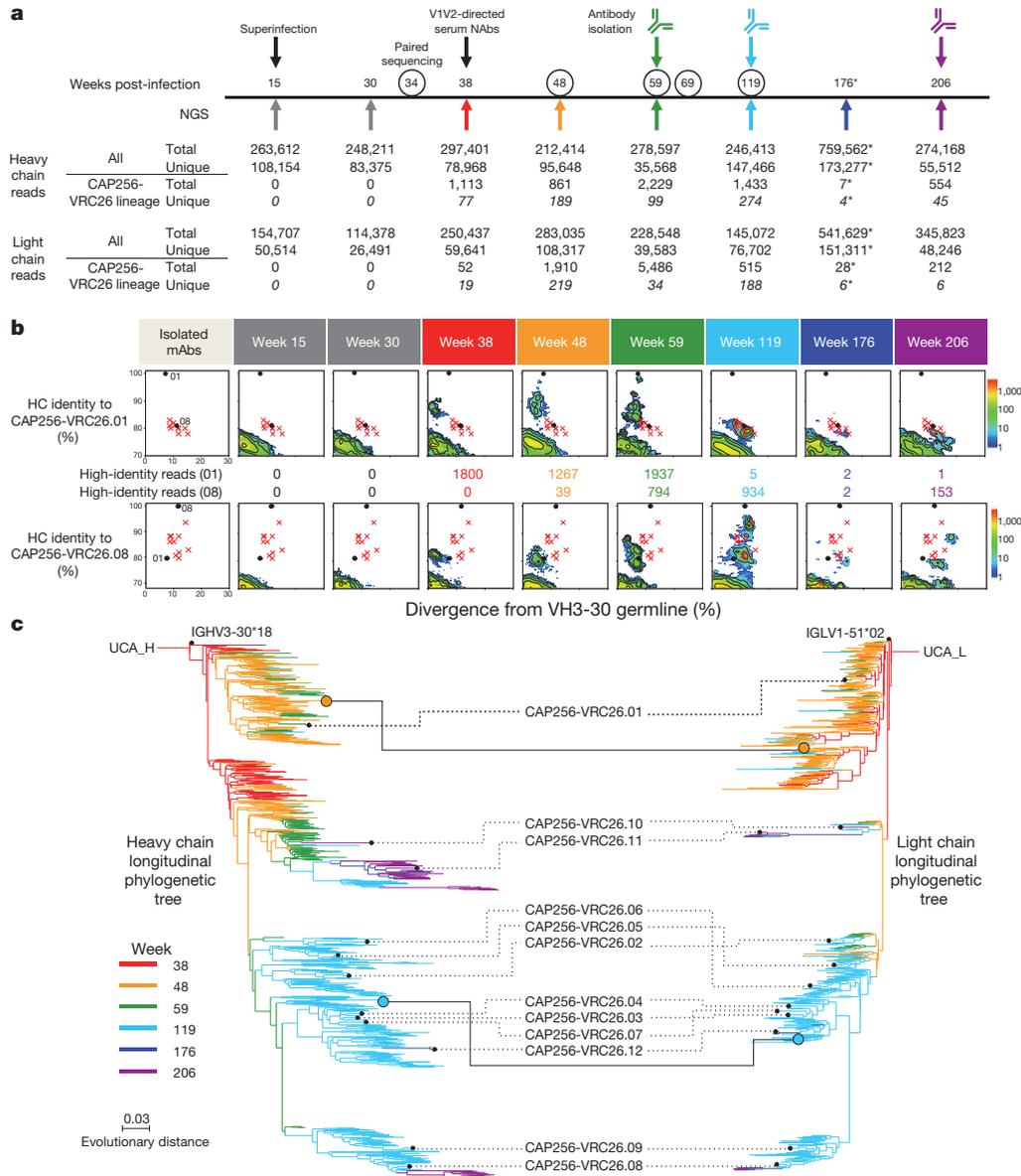


Figure 3 | Maturation of the CAP256-VRC26 lineage revealed by NGS and VH:VL paired sequencing of B cell transcripts. **a**, Timeline of longitudinal peripheral blood samples with quantification of all NGS sequence reads (total and unique), and CAP256-VRC26 lineage-related reads (total and unique). Arrows below the line indicate time points of 454 pyrosequencing for heavy and light chain sequences. Circles indicate time points for pyrosequencing of sorted B cells (see 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) 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' marks 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 (VRC26.01 (top), VRC26.08 (bottom)). **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, labelled with antibody name). Branches are coloured 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.

infecting virus, the superinfecting virus contained V2 residues that are more commonly found among circulating viruses (Extended Data Fig. 8a). All 12 antibodies neutralized the superinfecting virus, and, with the exception of CAP256-VRC26.06, failed to neutralize the primary infecting virus, suggesting the superinfecting virus 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 primary infecting virus (Fig. 5a–c

and Supplementary Figs 8 and 9) and were therefore largely neutralization resistant (Fig. 5d and Supplementary Fig. 10). Among superinfecting-virus-like sequences, a rare K169I mutation arose under strong directional selection (Supplementary Table 3) as the CAP256-VRC26 lineage emerged, which rendered the superinfecting virus resistant to only the earliest antibody (Extended Data Fig. 8d, e), indicating 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 substantial shift (Fig. 5a and Supplementary Figs 8 and 9), with the

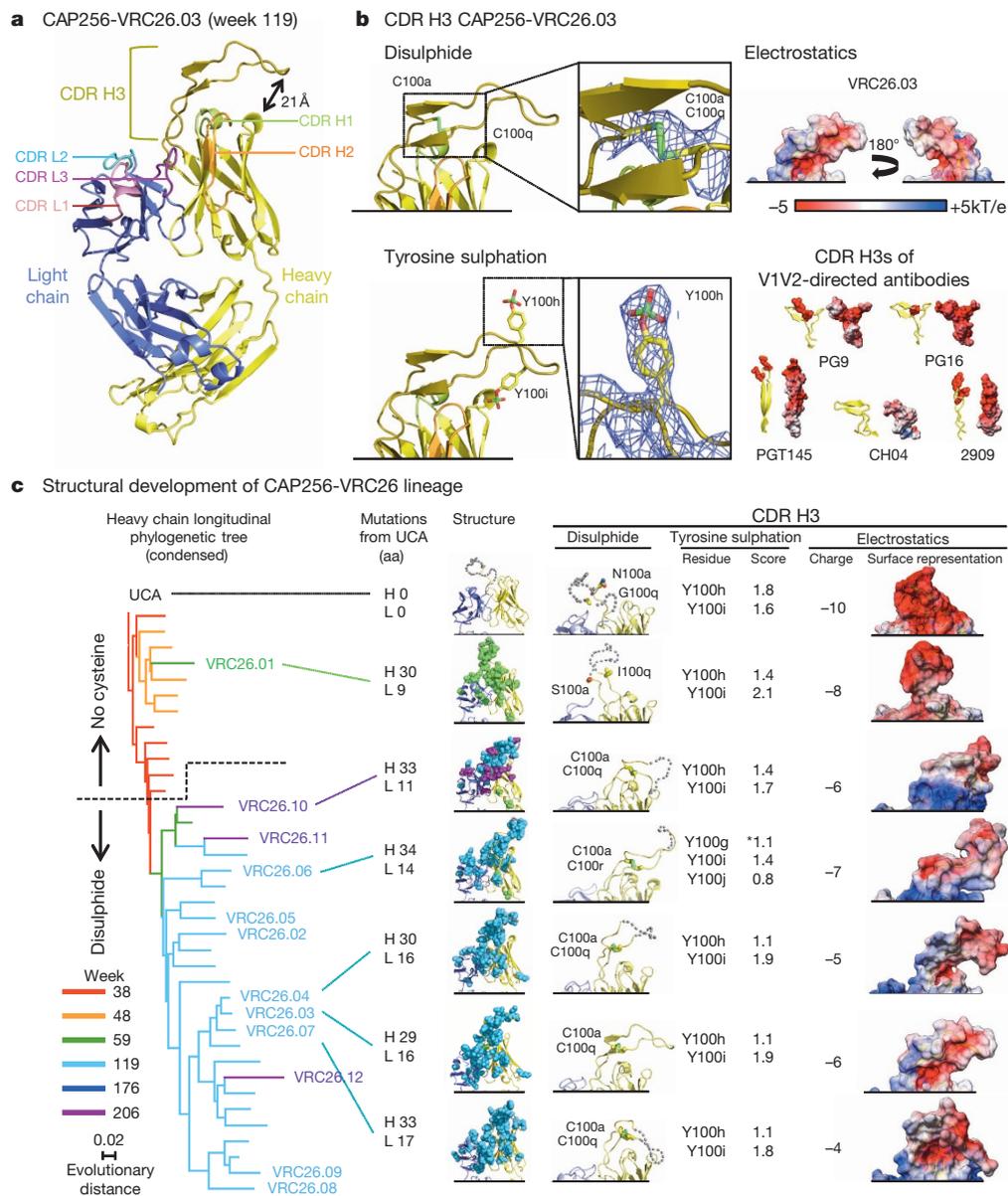


Figure 4 | Structural characteristics of the developing CAP256-VRC26 lineage. **a**, Crystal structure of the antigen-binding fragment (Fab) of CAP256-VRC26.03 shown in ribbon diagram representation. **b**, The intra-loop disulphide bond and tyrosine sulphation are shown in stick representation, and enlarged to show electron density (blue mesh, $2F_o - F_c$ at 1σ) (left). Molecular surface, with electrostatic potentials coloured red for acidic and blue for basic (right). CDR H3 regions of broadly neutralizing V1V2-directed antibodies are shown for comparison, with the left image in ribbon representation (tyrosine sulphates highlighted in red) and the right image in electrostatic representation. **c**, A condensed heavy chain phylogenetic tree highlights the isolated antibodies (left). Scale, rate of nucleotide change between nodes. The number of amino acid (aa) mutations to the heavy chain (H) and light chain (L) relative to the UCA are shown. Structures of the variable regions (middle). Mutations from the UCA are represented as spheres coloured according to the week of antibody isolation at which the mutations first appear. CDR H3 details (right). Residues that are (or evolve to become) cysteines are labelled (grey dotted lines indicate modelled disordered regions). The position of tyrosines predicted to be sulphated (scores >1) are noted and were included in the formal charges shown for each CDR H3 and the electrostatic representations (far right). Asterisk denotes Tyr insertion in VRC26.06.

superinfecting-virus-like V1V2 dominating just before the development of neutralization breadth. Neutralization of Env clones by later antibodies (CAP256-VRC26.02-12) tracked with the presence of superinfecting-virus-like V1V2 sequences (black bar, Fig. 5c) until escape occurred through mutations at positions 166 or 169 (Fig. 5c, d and Extended Data Fig. 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 superinfecting-virus-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 superinfecting virus weakly, but did not bind or 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 primary infecting virus 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 a superinfecting-virus-like V1V2. Subsequent affinity maturation, focused within 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. On the basis of 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

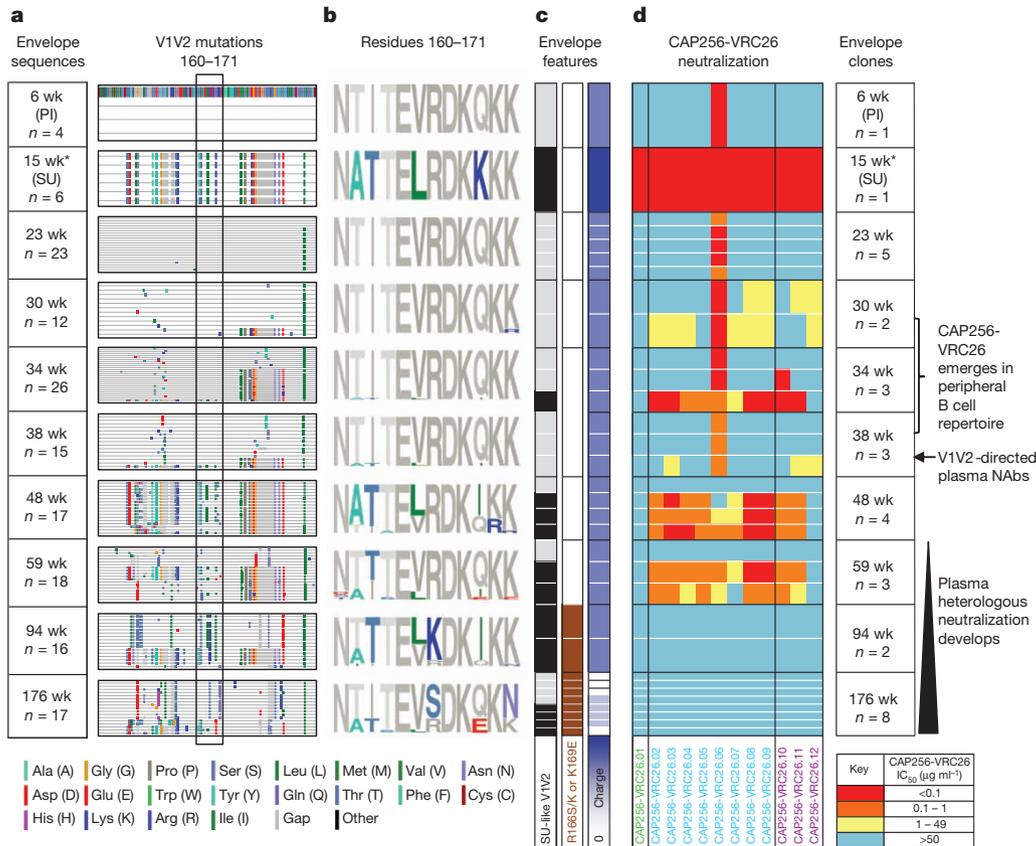


Figure 5 | HIV-1 Env evolution and the development of the CAP256-VRC26 lineage. **a**, V1V2 sequences are shown in highlighter format with the primary infecting virus (PI) designated as master and V2 residues 160 to 171 boxed. Asterisk at week 15 denotes sequences amplified with strain-specific primers matching the superinfecting virus (SU) virus. **b**, Logogram of the V2 epitope for all CAP256 sequences, with mutations away from the PI (master sequence) in colour. **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

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 neutralizing antibodies^{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 infection³⁶ and viral sieve analysis showed immune pressure in the same region³⁷. Although 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 neutralizing antibodies^{1–3,38}. Previously described V1V2 neutralizing 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 donors³⁹, 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 (≥ 24 amino acids, international immunogenetics database (IMGT)⁴⁰

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 monoclonal antibodies of representative longitudinal Env clones isolated between 6 and 176 weeks post-infection (weeks shown at far right). The CAP256 monoclonal antibodies are coloured 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.

definition) or very long (≥ 28 amino acids) CDR H3s constitute $\sim 3.5\%$ and 0.4% , respectively, of naive B cells³⁹. 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 autoreactivity⁴². The UCA and mature CAP256-VRC26 antibodies demonstrated little or no reactivity with 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 ≥ 28 amino acids (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 (ref. 15), the autologous virus in CAP256 showed extensive diversification before 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 neutralizing antibodies revealed by the CAP256-VRC26 lineage indicates that neutralization potency and breadth can be achieved without extraordinary levels of somatic hypermutation. Although some neutralizing antibodies 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

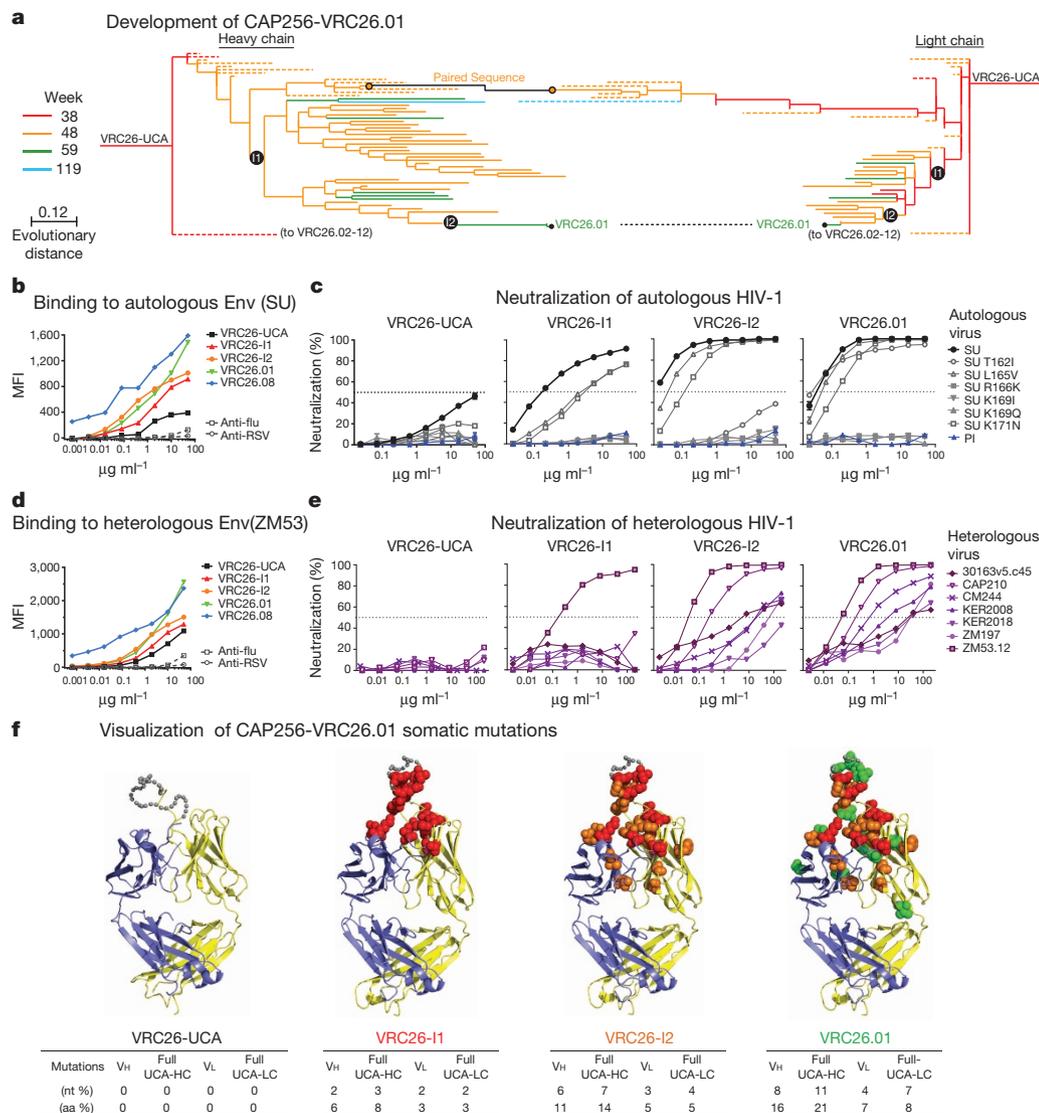


Figure 6 | Development from UCA to CAP256-VRC26.01. **a**, Expanded view of the phylogenetic trees from Fig. 3c, highlighting the maturation pathway of CAP256-VRC26.01. Off-pathway branches were collapsed and are shown as dashed lines. Inferred intermediates VRC26-I1 and VRC26-I2 were expressed for functional 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 PI, SU and point mutants (**c**) and seven heterologous viruses (**e**). Bars, standard error of the mean (triplicates). **f**, Structural models of VRC26.01 lineage antibodies. Affinity matured residues are shown as spheres coloured 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 (V_H or V_L), or from the full UCA (UCA-HC or UCA-LC).

with a protruding, tyrosine-sulphated, 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 neutralizing antibodies as an indicator of the ability to engage an appropriate naive 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.

METHODS SUMMARY

Serial blood samples were collected from HIV-1-infected subject CAP256 from 6 to 225 weeks after infection. Monoclonal antibodies CAP256-VRC26.01-12 were generated by single-B cell culture, microneutralization screening, RT-PCR, subcloning, and expression as described in^{27,28,45}. CDR lengths used Kabat notation²⁹ except as indicated. Binding of CAP256-VRC26 antibodies to virus-like particles was assessed by ELISA³¹ and binding to cell-surface expressed Env was measured by flow cytometry³⁰. HIV-1 neutralizing activity of patient plasma and monoclonal antibodies was determined with Env-pseudoviruses using the TZM-bl cell line^{46,47}. Neutralization fingerprints are the rank-order of neutralization potencies for an antibody against a set of diverse viral strains, calculated as in ref. 18. A 28Å reconstruction of the BG505 SOSIP.664 gp140 trimer with a single VRC26.09 Fab was obtained by negative stain EM using Appion, Xmipp, IMAGIC, and EMAN software. 454 pyrosequencing was performed as previously described^{12,14} on samples

from 8 time points after HIV-1 infection. High-throughput V_H:V_L pairing of peripheral blood CD27⁺ B cells was performed in single cell emulsions generated using a flow focusing apparatus³⁵. Phylogenetic analysis, inference of UCA, and identification, synthesis, and expression of clone members were performed as described in the Methods. Epitope mapping onto the spike trimer was performed with the software package UCSF Chimera, using experimental data as described in Methods. Crystallographic analysis of Fab fragments was performed as described in the Methods. Structure modelling of disordered residues in Fab crystal structures was performed using Loopy software. Single-genome amplification and expression of *env* genes was performed as described in Methods and in refs 48, 49.

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

Received 13 September 2013; accepted 16 January 2014.

Published online 2 March 2014.

- Haynes, B. F., Kelsae, G., Harrison, S. C. & Kepler, T. B. B-cell-lineage immunogen design in vaccine development with HIV-1 as a case study. *Nature Biotechnol.* **30**, 423–433 (2012).
- Kong, L. & Sattentau, Q. J. Antigenicity and immunogenicity in HIV-1 antibody-based vaccine design. *J. AIDS Clin. Res. (Suppl. 8)* 003 (2012).
- Mascola, J. R. & Haynes, B. F. HIV-1 neutralizing antibodies: understanding nature's pathways. *Immunol. Rev.* **254**, 225–244 (2013).
- Richman, D. D., Wrinn, T., Little, S. J. & Petropoulos, C. J. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc. Natl Acad. Sci. USA* **100**, 4144–4149 (2003).
- Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).

6. Gray, E. S. *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. *J. Virol.* **85**, 4828–4840 (2011).
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. *J. Virol.* **83**, 10269–10274 (2009).
8. Sather, D. N. *et al.* Factors associated with the development of cross-reactive neutralizing antibodies during human immunodeficiency virus type 1 infection. *J. Virol.* **83**, 757–769 (2009).
9. Doria-Rose, N. A. *et al.* Breadth of human immunodeficiency virus-specific neutralizing activity in sera: clustering analysis and association with clinical variables. *J. Virol.* **84**, 1631–1636 (2010).
10. Glanville, J. *et al.* Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proc. Natl Acad. Sci. USA* **106**, 20216–20221 (2009).
11. Briney, B. S., Willis, J. R., McKinney, B. A. & Crowe, J. E. Jr. High-throughput antibody sequencing reveals genetic evidence of global regulation of the naive and memory repertoires that extends across individuals. *Genes Immun.* **13**, 469–473 (2012).
12. Wu, X. *et al.* Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* **333**, 1593–1602 (2011).
13. Zhu, J. *et al.* Mining the antibodyome for HIV-1-neutralizing antibodies with next-generation sequencing and phylogenetic pairing of heavy/light chains. *Proc. Natl Acad. Sci. USA* **110**, 6470–6475 (2013).
14. Zhu, J. *et al.* Somatic populations of PGT135–137 HIV-1-neutralizing antibodies identified by 454 pyrosequencing and bioinformatics. *Front. Microbiol.* **3**, 315 (2012).
15. Liao, H. X. *et al.* Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* **496**, 469–476 (2013).
16. Walker, L. M. *et al.* A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals. *PLoS Pathog.* **6**, e1001028 (2010).
17. Lynch, R. M. *et al.* The B cell response is redundant and highly focused on V1V2 during early subtype C infection in a Zambian seroconverter. *J. Virol.* **85**, 905–915 (2011).
18. Georgiev, I. S. *et al.* Delineating antibody recognition in polyclonal sera from patterns of HIV-1 isolate neutralization. *Science* **340**, 751–756 (2013).
19. Walker, L. M. *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326**, 285–289 (2009).
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. *J. Virol.* **85**, 9998–10009 (2011).
21. Walker, L. M. *et al.* Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477**, 466–470 (2011).
22. McLellan, J. S. *et al.* Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480**, 336–343 (2011).
23. Pancera, M. *et al.* Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1–V2-directed antibody PG16. *Nature Struct. Mol. Biol.* **20**, 804–813 (2013).
24. Julien, J. P. *et al.* Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. *Proc. Natl Acad. Sci. USA* **110**, 4351–4356 (2013).
25. Moore, P. L. *et al.* Potent and broad neutralization of HIV-1 subtype C by plasma antibodies targeting a quaternary epitope including residues in the V2 loop. *J. Virol.* **85**, 3128–3141 (2011).
26. Moore, P. L. *et al.* Multiple pathways of escape from HIV broadly cross-neutralizing V2-dependent antibodies. *J. Virol.* **87**, 4882–4894 (2013).
27. Huang, J. *et al.* Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* **491**, 406–412 (2012).
28. Tiller, T. *et al.* Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J. Immunol. Methods* **329**, 112–124 (2008).
29. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. *Sequences of Proteins of Immunological Interest* (U.S. Department of Health and Human Services, National Institutes of Health, 1991).
30. Pancera, M. & Wyatt, R. Selective recognition of oligomeric HIV-1 primary isolate envelope glycoproteins by potentially neutralizing ligands requires efficient precursor cleavage. *Virology* **332**, 145–156 (2005).
31. Tong, T., Crooks, E. T., Osawa, K. & Binley, J. M. HIV-1 virus-like particles bearing pure env trimers expose neutralizing epitopes but occlude nonneutralizing epitopes. *J. Virol.* **86**, 3574–3587 (2012).
32. Lyumkis, D. *et al.* Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1484–1490 (2013).
33. Julien, J. P. *et al.* Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1477–1483 (2013).
34. Doria-Rose, N. A. *et al.* A short segment of the HIV-1 gp120 V1/V2 region is a major determinant of resistance to V1/V2 neutralizing antibodies. *J. Virol.* **86**, 8319–8323 (2012).
35. DeKosky, B. J. *et al.* High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nature Biotechnol.* **31**, 166–169 (2013).
36. Haynes, B. F. *et al.* Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N. Engl. J. Med.* **366**, 1275–1286 (2012).
37. Rolland, M. *et al.* Increased HIV-1 vaccine efficacy against viruses with genetic signatures in Env V2. *Nature* **490**, 417–420 (2012).
38. Overbaugh, J. & Morris, L. The antibody response against HIV-1. *Cold Spring Harb. Perspect. Med.* **2**, a007039 (2012).
39. Briney, B. S., Willis, J. R. & Crowe, J. E. Jr. Human peripheral blood antibodies with long HCDR3s are established primarily at original recombination using a limited subset of germline genes. *PLoS ONE* **7**, e36750 (2012).
40. Lefranc, M. P. *et al.* IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev. Comp. Immunol.* **27**, 55–77 (2003).
41. Wardemann, H. *et al.* Predominant autoantibody production by early human B cell precursors. *Science* **301**, 1374–1377 (2003).
42. Haynes, B. F. *et al.* Cardioliipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science* **308**, 1906–1908 (2005).
43. Kwong, P. D. & Mascola, J. R. Human antibodies that neutralize HIV-1: identification, structures, and B cell ontogenies. *Immunity* **37**, 412–425 (2012).
44. Burton, D. R. *et al.* A blueprint for HIV vaccine discovery. *Cell Host Microbe* **12**, 396–407 (2012).
45. Huang, J. *et al.* Isolation of human monoclonal antibodies from peripheral blood B cells. *Nature Protocols* **8**, 1907–1915 (2013).
46. Shu, Y. *et al.* Efficient protein boosting after plasmid DNA or recombinant adenovirus immunization with HIV-1 vaccine constructs. *Vaccine* **25**, 1398–1408 (2007).
47. Montefiori, D. C. Measuring HIV neutralization in a luciferase reporter gene assay. *Methods Mol. Biol.* **485**, 395–405 (2009).
48. Kraus, M. H. *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* **397**, 346–357 (2010).
49. Salazar-Gonzalez, J. F. *et al.* Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J. Virol.* **82**, 3952–3970 (2008).
50. Liu, J., Bartesaghi, A., Borgnia, M. J., Sapiro, G. & Subramaniam, S. Molecular architecture of native HIV-1 gp120 trimers. *Nature* **455**, 109–113 (2008).

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

Acknowledgements We thank the participants in the CAPRISA 002 study for their commitment. For technical assistance and advice, we thank: K. Mlisana, S. Sibeko, N. Naicker, the CAPRISA 002 clinical team, N. Samsunder, S. Heeralal, B. Lambson, M. Madzivhandila, T. Khoza, C. Mitchell Scheepers, E. Turk, C.-L. Lin, M. Roederer, J. Stuckey, B. Hartman, G. Loots, J. H. Lee, G. Ippolito, B. Briney, S. Hunnicke-Smith and J. Wheeler, and members of the WCMC HIVRAD Core and the NIH Vaccine Research Center HIMS, HMC, SBS and SBIS sections. We thank J. Baalwa, D. Ellenberger, F. Gao, B. Hahn, K. Hong, J. Kim, F. McCutchan, D. Montefiori, J. Overbaugh, E. Sanders-Buell, G. Shaw, R. Swanstrom, M. Thomson, S. Tovanabutra and L. Zhang for contributing the HIV-1 Envelope plasmids used in our neutralization panel. Funding was provided by the intramural research programs of the Vaccine Research Center and NIAID, the Fogarty International Center, NHGRI, and NIGMS of the National Institutes of Health, USA; the International AIDS Vaccine Initiative; the National Science Foundation; Scripps CHAV-ID; the South African Department of Science and Technology; and fellowships from the Wellcome Trust, Hertz Foundation, Donald D. Harrington Foundation, Poliomyelitis Research Foundation and the National Research Foundation of South Africa. Use of sector 22 (Southeast Region Collaborative Access team) at the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under contract number W-31-109-Eng-38.

Author Contributions N.A.D.-R., C.A.S., J.G. and P.L.M. contributed equally to this work. N.A.D.-R., C.A.S., J.G., P.L.M. and J.N.B., designed and performed experiments, analysed data and wrote the manuscript. L.M., P.D.K., L.S. and J.R.M. conceived and designed the experiments, analysed data, and wrote the manuscript. B.J.D., M.J.E., I.S.G., H.J.K., M.P. and R.P.S. conducted experiments and analysed data. H.R.A.-T., B.T.B., E.T.C., A.C., K.H.H., R.K., M.K.L., K.M., M.N., S.O., Ry.S.R., Re.S.R., S.D.S., C.K.W., Y.Y., J.C.M. and NISC conducted experiments. C.W. and A.D. contributed analysis tools and data analysis. S.S.A.K. and N.J.G. conceived and managed the CAPRISA cohorts. J.M.B., R.W.S., I.A.W., J.P.M., A.B.W., G.G., N.S.L., D.J.S., C.S. and Z.Z. analysed data.

Author Information Coordinates and structure factors for CAP256-VRC26 lineage Fabs have been deposited with the Protein Data Bank under accession codes 4ODH, 4OCR, 4OD1, 4ORC, 4OCW, 4OD3 and 4OCS. The EM reconstruction density for the CAP256-VRC26.09 complex with BG505 SOSIP.664 trimer has been deposited with the Electron Microscopy Data Bank under accession code EMD-5856. We have also deposited deep sequencing data used in this study to National Center for Biotechnology Information Short Reads Archives (SRA) under accession numbers SRP034555 and SRP017087. Information deposited with GenBank includes: the heavy- and light-chain variable region sequences of cloned antibodies CAP256-VRC26.01-12, UCA, I1 and I2 (accession numbers KJ134860–KJ134889); bioinformatically identified VRC26-related sequences from B cell transcripts: 680 heavy chains and 472 light chains (accession numbers KJ133708 – KJ134387, KJ134388 – KJ134859); and CAP256 Env sequences (accession numbers KF996576 – KF996716). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed for CAPRISA and viral evolution to L.M. (lynm@nicd.ac.za), for crystallography to P.D.K. (pdkwong@nih.gov), for NGS to L.S. (iss8@columbia.edu), and for isolated antibodies to J.R.M. (jrmascola@nih.gov).

METHODS

Study subject. CAPRISA participant CAP256 was enrolled into the CAPRISA acute infection study⁵¹ 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 breadth⁶. 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 2005–09.

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 ref. 18. Cells were plated at two B cells per 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 refs 27 and 45. Culture supernatants were screened by microneutralization as described in ref. 52 against HIV-1 ZM53.12 and either CAP45.G3 or CAP210.E8 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 ref. 18. Heavy chains were reconstituted as IgG1. The efficiency of cloning was as follows. For week 59, a total of 15,000 B cells (7,500 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, a total of 45,000 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, a total of 42,000 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 time point 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, titred and used to infect TZM-bl target cells as described previously^{64,47}. 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 half-maximum inhibitory concentration (IC₅₀) in micrograms per micro-litre for monoclonal antibodies, or reciprocal dilution (ID₅₀) for plasma samples.

Neutralization fingerprints. Owing 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 strains¹⁸.

The correlations between the neutralization fingerprints of the CAP256-VRC26 antibodies and the neutralization patterns of four longitudinal serum time points (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, CI080.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 time points 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 were averaged and consolidated. All correlations are based on the Spearman's rank correlation coefficient.

Virus-like particle ELISA. VLP ELISAs were performed as described previously³¹. 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 20× the

concentration in transfection supernatants. Monoclonal antibody 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 tail³⁰. For binding experiments: after 2 days, the cells were stained with ViVid viability dye (Invitrogen) followed by serial dilutions of antibodies, two washes with PBS/5% FBS, then R-PE-conjugated F(ab) goat anti-human IgG specific for the Fc fragment (Jackson ImmunoResearch) at a 1:200 dilution⁵⁴. For competition assays, the cells were stained with ViVid viability dye followed by biotinylated CAP256-VRC26.08 (0.8 µg ml⁻¹) premixed with serially diluted unlabelled competitor antibodies. After incubation and 2 washes, cells were stained with streptavidin-PE (Invitrogen) at 1:200 dilution. Cells were analysed 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 ref. 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⁻¹ in the kit sample diluent and tested in threefold 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 ref. 42. Antibodies were diluted to 50 µg ml⁻¹ and 25 µg ml⁻¹ 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 trimer produced in HEK 293S cells were analysed by negative stain EM. A 3 µl aliquot of ~8 µg ml⁻¹ of the complex was applied for 15 s onto a glow discharged, carbon-coated 400 Cu mesh grid and stained with 2% uranyl formate for 20 s. 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 1,000 nm using LEGION⁵⁵.

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 classes⁵⁸. 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 softwares⁵⁹. 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 EMAN⁶⁰. The resolution of the final reconstruction was calculated to be 28 Å 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 apparatus³⁵ (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 analysed with FR1 primers³⁵, while the other half were analysed with leader peptide primers⁴¹ (Supplementary Table 5). All other time points were analysed in a single group using only FR1 primers (Supplementary Table 1). Overlap extension RT-PCR was performed as previously reported³⁵, with extension time increased to 125 s. Nested PCR was performed as described previously with a 23-s

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 previously¹⁴. 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 $\leq 10^{-10}$ were discarded. For reads assigned to any VH3-30 or V λ 1-51 allele, (the CAP256-VRC26 germline genes), ClustalW2 (ref. 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 V λ 1-51, were submitted to IMGT High-Vquest⁶² (<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 J λ 1), and having similar divergence ($\pm 15\%$) in the V and J genes, similar ($\pm 10\%$) 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 CDHit⁶³. 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 intradonor phylogenetic analysis¹⁴ 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. Joinsolve⁶⁴ 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 CDHit⁶³ 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 are shown in Nexus format in Supplementary Figs 12 and 13. MEGA5 (ref. 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. FASTML⁶⁷ 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 where 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 Scheepers, 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 CDR H3s were made with Weblogo⁶⁸.

X-ray crystallography. VRC26.UCA Fab was prepared by digesting purified IgG with Lys-C at 37°C for 2 h. 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 (GLEVLFGQP) 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 vapour diffusion method in sitting drops at 20°C by mixing 0.2 μ l of protein complex with 0.2 μ l of reservoir solution. Crystals were manually reproduced in hanging drops by mixing 1.0 μ l protein complex with 1.0 μ l reservoir solution. VRC26-UCA was crystallized with a reservoir solution of 27% PEG 8000 and 0.1 M 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 0.1 M sodium 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 0.1 M 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% isopropanol and 0.1 M 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 3 M sodium formate and 0.1 M 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, 0.1 M zinc acetate and 0.1 M 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, 0.4 M sodium 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 suite⁶⁹ and model building and refinement were performed in COOT⁷⁰ and PHENIX⁷¹, 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 MolProbity⁷². Ribbon diagram representations of protein crystal structures were made with PyMOL⁷³ and electrostatics were calculated and rendered with UCSF Chimera⁷⁴.

Structure modelling 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 Chimera⁷⁴. 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 accession number 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 accession number 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 threefold symmetry of the HIV-1 spike. The BG505.664 SOSIP crystal structure³³, PDB 4NCO, 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 modelled. 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 accession code 3U4E).

Loop modelling. 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 modelled with Loopy⁷⁸ (http://wiki.c2b2.columbia.edu/honiglab_public/index.php/Software:Loopy) and represented as grey dots. Mutations of the intermediates were coloured 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 modelled loops (Fig. 4), were modelled 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, sulphate 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 sulphated tyrosines. The Rapid Torsion Scan module in Maestro was used to sample the chi angle involving the sulphate 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 sulphation predictions were carried out in GPS-TPS (Z. Pan *et al.*, <http://tsp.biocuckoo.org>).

Single genome amplification (SGA), 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 templates⁴⁹ 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 analyser. 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 FastTree⁷⁹ 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 HyPhy⁸².

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 ($n = 3,990$) 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'-TAGCCCTCCAGTCCCCCTTTCTTTA-3' (ref. 83) and directional primer, EnvAstop, 5'-CACCGGCTTAGGCATCTCCTATGGCAGGAAGAA-3' (ref. 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.

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* **3**, e1954 (2008).

52. Doria-Rose, N. *et al.* High throughput HIV-1 microneutralization assay. *Protocol Exchange* <http://dx.doi.org/10.1038/protex.2013.069> (2013).
53. Wu, X. *et al.* Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329**, 856–861 (2010).
54. Julien, J. P. *et al.* Broadly neutralizing antibody PGT121 allosterically modulates CD4 binding via recognition of the HIV-1 gp120 V3 base and multiple surrounding glycans. *PLoS Pathog.* **9**, e1003342 (2013).
55. Suloway, C. *et al.* Automated molecular microscopy: the new Legimon system. *J. Struct. Biol.* **151**, 41–60 (2005).
56. Lander, G. C. *et al.* Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* **166**, 95–102 (2009).
57. Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S. & Carragher, B. DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol.* **166**, 205–213 (2009).
58. Sorzano, C. O. *et al.* A clustering approach to multireference alignment of single-particle projections in electron microscopy. *J. Struct. Biol.* **171**, 197–206 (2010).
59. van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R. & Schatz, M. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116**, 17–24 (1996).
60. Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97 (1999).
61. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
62. Alamyar, E., Giudicelli, V., Li, S., Duroux, P. & Lefranc, M. P. IMGT/HighV-QUEST: the IMGT web portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high throughput and deep sequencing. *Immunome Res.* **8** (2012).
63. Li, W., Jaroszewski, L. & Godzik, A. Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics* **17**, 282–283 (2001).
64. Souto-Carneiro, M. M., Longo, N. S., Russ, D. E., Sun, H. W. & Lipsky, P. E. Characterization of the human Ig heavy chain antigen binding complementarity determining region 3 using a newly developed software algorithm, JOINSOLVER. *J. Immunol.* **172**, 6790–6802 (2004).
65. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
66. Waddell, P. J. & Steel, M. A. General time-reversible distances with unequal rates across sites: mixing gamma and inverse Gaussian distributions with invariant sites. *Mol. Phylogenet. Evol.* **8**, 398–414 (1997).
67. Ashkenazy, H. *et al.* FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Res.* **40**, W580–W584 (2012).
68. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
69. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
70. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
71. Adams, P. D. *et al.* Recent developments in the PHENIX software for automated crystallographic structure determination. *J. Synchrotron Radiat.* **11**, 53–55 (2004).
72. Davis, I. W., Murray, L. W., Richardson, J. S. & Richardson, D. C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615–W619 (2004).
73. DeLano, W. L. The PyMOL Molecular Graphics System. <http://www.pymol.org> (DeLano Scientific, San Carlos, California, 2002).
74. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
75. Tran, E. E. *et al.* Structural mechanism of trimeric HIV-1 envelope glycoprotein activation. *PLoS Pathog.* **8**, e1002797 (2012).
76. Zhou, T. *et al.* Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329**, 811–817 (2010).
77. Pejchal, R. *et al.* A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. *Science* **334**, 1097–1103 (2011).
78. Soto, C. S., Fasnacht, M., Zhu, J., Forrest, L. & Honig, B. Loop modeling: Sampling, filtering, and scoring. *Proteins* **70**, 834–843 (2008).
79. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
80. Murrell, B. *et al.* Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* **8**, e1002764 (2012).
81. Kosakovsky Pond, S. L., Poon, A. F., Leigh Brown, A. J. & Frost, S. D. A maximum likelihood method for detecting directional evolution in protein sequences and its application to influenza A virus. *Mol. Biol. Evol.* **25**, 1809–1824 (2008).
82. Pond, S. L., Frost, S. D. & Muse, S. V. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679 (2005).
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. *J. Virol.* **70**, 7013–7029 (1996).

a

| | <u>CDRH1</u> | <u>CDRH2</u> |
|-------------|---|--------------|
| IGHV3-30*18 | | |
| IGHJ3-01 | | |
| VRC26-UCA_H | QVQLVESGGGVVQPGRSRLRSLCAASGFTFSSYGMHWVRQAPGKGLEWVAVISYDGSNKYYADS VKGRFTISRDN SKNTLY | |
| VRC26-I1 |N.A.D.....S.....N..... | |
| VRC26-I2 |NFA.G.....F..S.....N.G.....F | |
| VRC26.01 | E..V.....T.....NFA.G.....F..S.....N.G.....VF | |
| VRC26.02 | E.....Q.S.GN.....A..FA.TKTN.....V. | |
| VRC26.03 | E.....K.....R.S.NR.....A.....TD.H.K.W..... | |
| VRC26.04 | E.....K.....Q.S.NR.....A.....TD.H.K.W..... | |
| VRC26.05 | E.....H.SL.....TG..FA.TKT..G..R...V..I...F | |
| VRC26.06 | E..I.....G...S.NN.....G..FA.IK..GT.....NF. | |
| VRC26.07 | E.....VG.Q.S.NR.....G..F..TDR.H..N.W..... | |
| VRC26.08 | E.....T.....Q...N.....SV.N..TK..HG...W.....F | |
| VRC26.09 | E.....T.....Q.N.AN.....S..N..TK..HEE..W...K...S..F | |
| VRC26.10 | .AI.....Q...GH.L.....S..FA.TKMD.....A..... | |
| VRC26.11 | E.....P.K.T..V..R...AF.....S..FA.IK.....S.....~.F | |
| VRC26.12 | E...Q.....PY.E.GR..F.....S..FA.RDY.HSP..W..... | |

CDRH3

| | |
|-------------|--|
| IGHV3-30*18 | |
| IGHJ3-01 | |
| VRC26-UCA_H | LQMNSLRAEDTAVVYCAKDLGESENEEWAT~DYYDFSIGYPGQDPR~~GVVGFADIWGQGTMTVTVSS |
| VRC26-I1 |NK.D...~L.A.V...~A..... |
| VRC26-I2 |L.....V..YKSD...~L.A.I...~AM..... |
| VRC26.01 |V...L.....V.DYKSD..G~E...I..S..I...~AM...L.....P |
| VRC26.02 |L.F...IR.Y.C.Y.TS~...GRPQ.CI.S...~T..V..... |
| VRC26.03 |L.....R.D.C...WS~...GKQL.CRKS...~A.I..G..... |
| VRC26.04 |L.....R.D.C...WS~...GKQL.CRKS...~A.I..K.....I.. |
| VRC26.05 |PD...M...R.QRYC...S~...GREQ.CL...~I..L..... |
| VRC26.06 |V...L...R..R.L.C...TLYN...GSRG.CV...~A.S..V..... |
| VRC26.07 |S.....L.....R.D.C...WS~...GKKL.CRKS...~A.V..K..... |
| VRC26.08 |S.....F.VR.QR.D.C...WS~...GREL.CRKF.GL.LA.I...H..... |
| VRC26.09 |I...L.F.V..QR.D.C...WS~...GREL.CRKS.GL.LA.I..M..H..... |
| VRC26.10 |V...L.....MR.Y.C.Y.TS~...GRPQ.CI.R...~I..M...T |
| VRC26.11 |G..G...GL.H.V..MR.L.C...S~...GKPQ.CL.R...~S.ISAW..P..... |
| VRC26.12 |I...L.F..R..R...C...ES~...GKKG.CVK...~A.GL.L...I... |

b

| | <u>CDRL1</u> | <u>CDRL2</u> |
|-------------|--|--------------|
| IGLV1-51*02 | | |
| IGLJ1-02 | | |
| VRC26-UCA L | QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNYVSWYQLPGTAPKLLIYENNKRPSPGIPDRFSGSKSGTSATLGITGLQ | |
| VRC26-I1-L |K.E..... | |
| VRC26-I2-L |T.....L.....K.D..... | |
| VRC26.01 |T.....RL.....K.DN.....S... | |
| VRC26.02 |N...KD...I.A.....S.....A... | |
| VRC26.03 |F...R...I.....ET.....A... | |
| VRC26.04 | .P.....N...G...F...R...S...T.....A...A... | |
| VRC26.05 | .A.....N...V.....V.....V..... | |
| VRC26.06 |T.K.Q.S.....DR...T.....A... | |
| VRC26.07 | .A.....N...G...F...R...S.....A... | |
| VRC26.08 | .A.....N...F...V...L.TYR.....A... | |
| VRC26.09 | .A.....N...F...R.....R.....A.S... | |
| VRC26.10 |D...H...TR.....A... | |
| VRC26.11 |N...PP.H.EK.D...R..M..MV..SKR..L..V..A.R..S..V..... | |
| VRC26.12 |T...T..V...GT..T..F...H.....L.V...A...A..... | |

CDRL3

| | |
|-------------|-------------------------------|
| IGLV1-51*02 |T.L |
| IGLJ1-02 |Y..... |
| VRC26-UCA L | TGDEADYCGTWDSSLSAGGVFGTGKVTVL |
| VRC26-I1-L |T..... |
| VRC26-I2-L |T..... |
| VRC26.01 |T...G..... |
| VRC26.02 |A...GRV~~~S.I...N.I.. |
| VRC26.03 | .A...E...A..SA...SAR...RI.. |
| VRC26.04 |A..AA..TSAR...I.S |
| VRC26.05 |F...GG..RT..L...R... |
| VRC26.06 |E...E...G...~.. |
| VRC26.07 |Y...A..AAR.NSAR...M... |
| VRC26.08 |TV.GVRKGV.A..... |
| VRC26.09 |W.AV.GVRRG..A...A..... |
| VRC26.10 |V.RPNR.A..... |
| VRC26.11 | .A.....GRMN~~~~T.S.. |
| VRC26.12 |G...A..SG..N~~~~I...S.. |

Extended Data Figure 1 | Amino acid sequences of CAP256-VRC26 heavy and light chains. **a**, Sequences of the 12 B-cell culture derived antibodies,

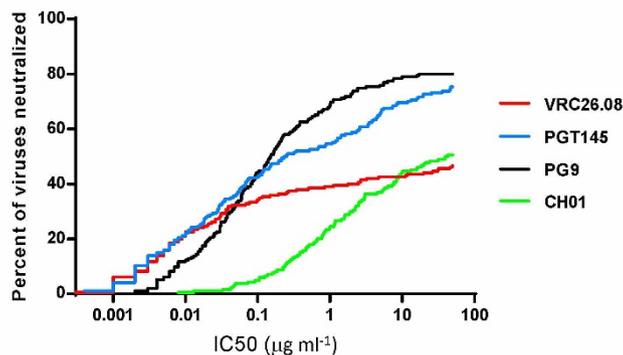
inferred germline V and J genes, and inferred intermediates are compared to the predicted UCA. **b**, Lambda light chain.

a

| 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 | |
|-----------------------|-------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----|
| Autologous | 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 | 4.780 | 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 | |
| MBS39.2B7 | | >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 | 0.006 | 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 | |

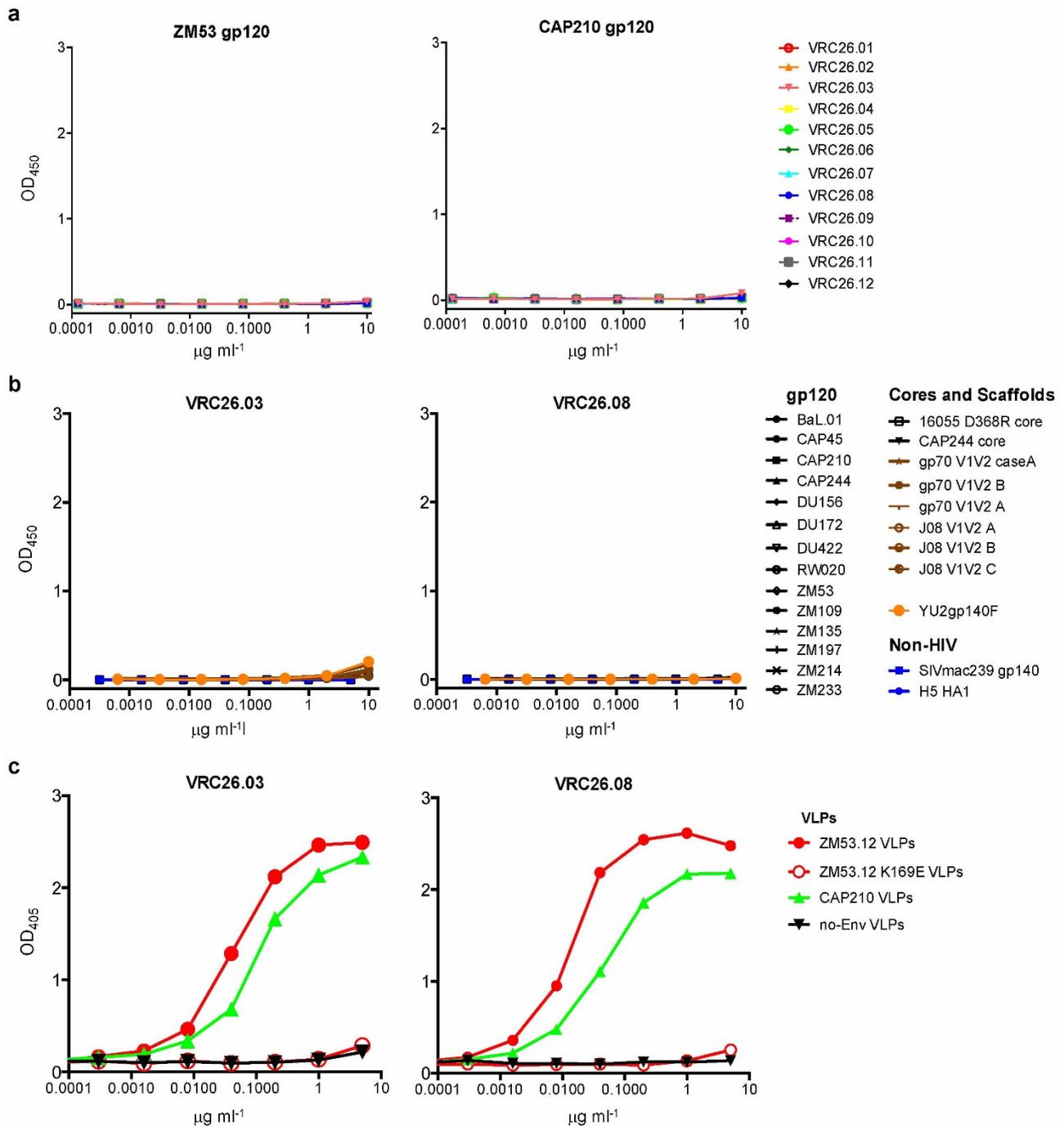
| | | | | | | | | | | | | |
|--------------|------|------|------|------|------|------|------|------|------|------|------|------|
| % breadth | 19% | 17% | 36% | 30% | 21% | 17% | 13% | 47% | 47% | 23% | 28% | 6% |
| geomean IC50 | 1.88 | 0.40 | 0.08 | 0.32 | 0.10 | 0.38 | 1.51 | 0.11 | 0.07 | 0.60 | 0.94 | 0.49 |

b



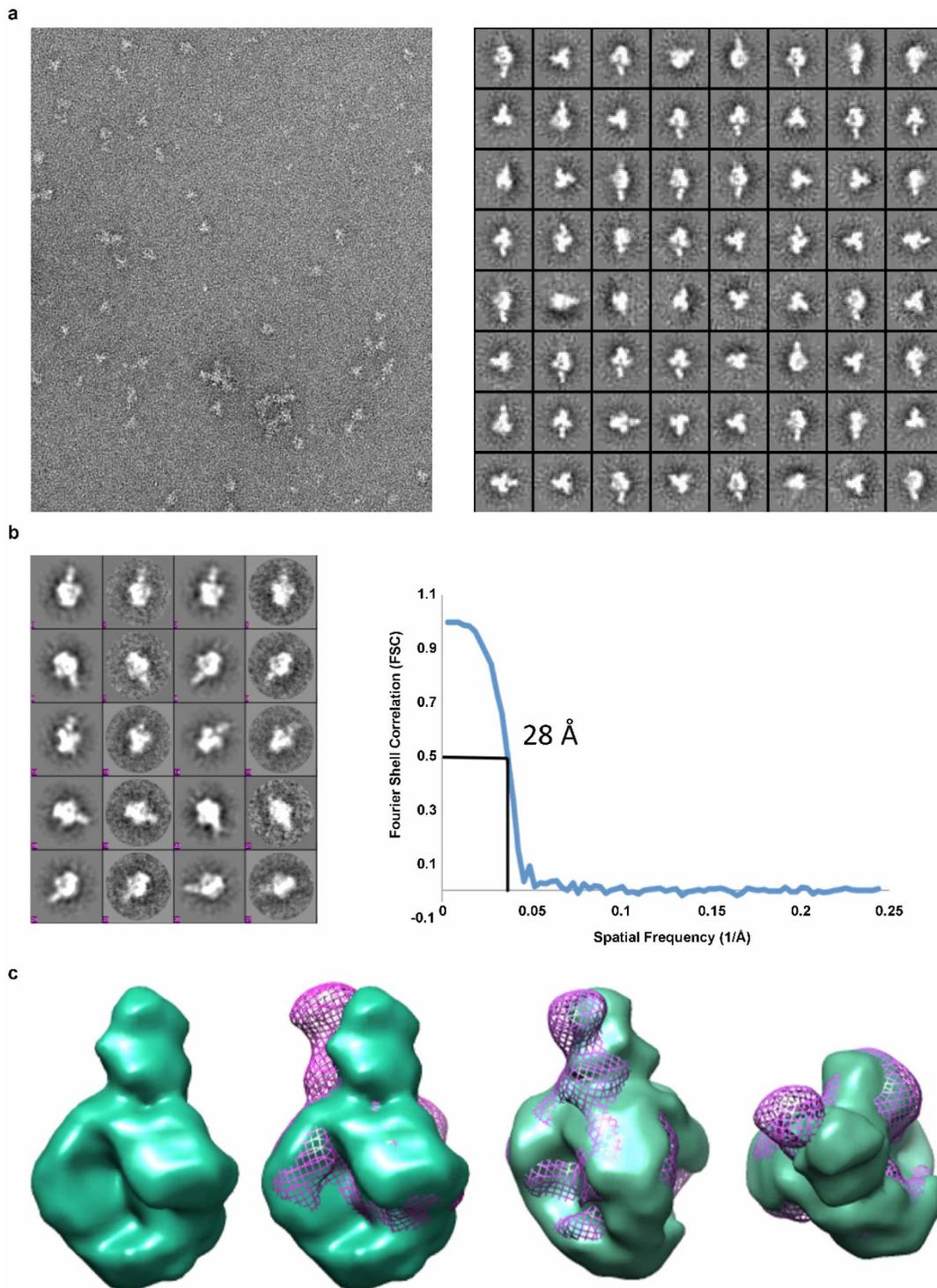
Extended Data Figure 2 | Neutralization breadth and potency of CAP256-VRC26 antibodies. a, Neutralization of autologous (CAP256 PI and SU) and 47 heterologous viruses by CAP256-VRC26 antibodies. Neutralization was measured using a TZM-bl assay with Env-pseudoviruses. Geometric mean was

calculated for values $<50 \mu\text{g ml}^{-1}$. 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} .



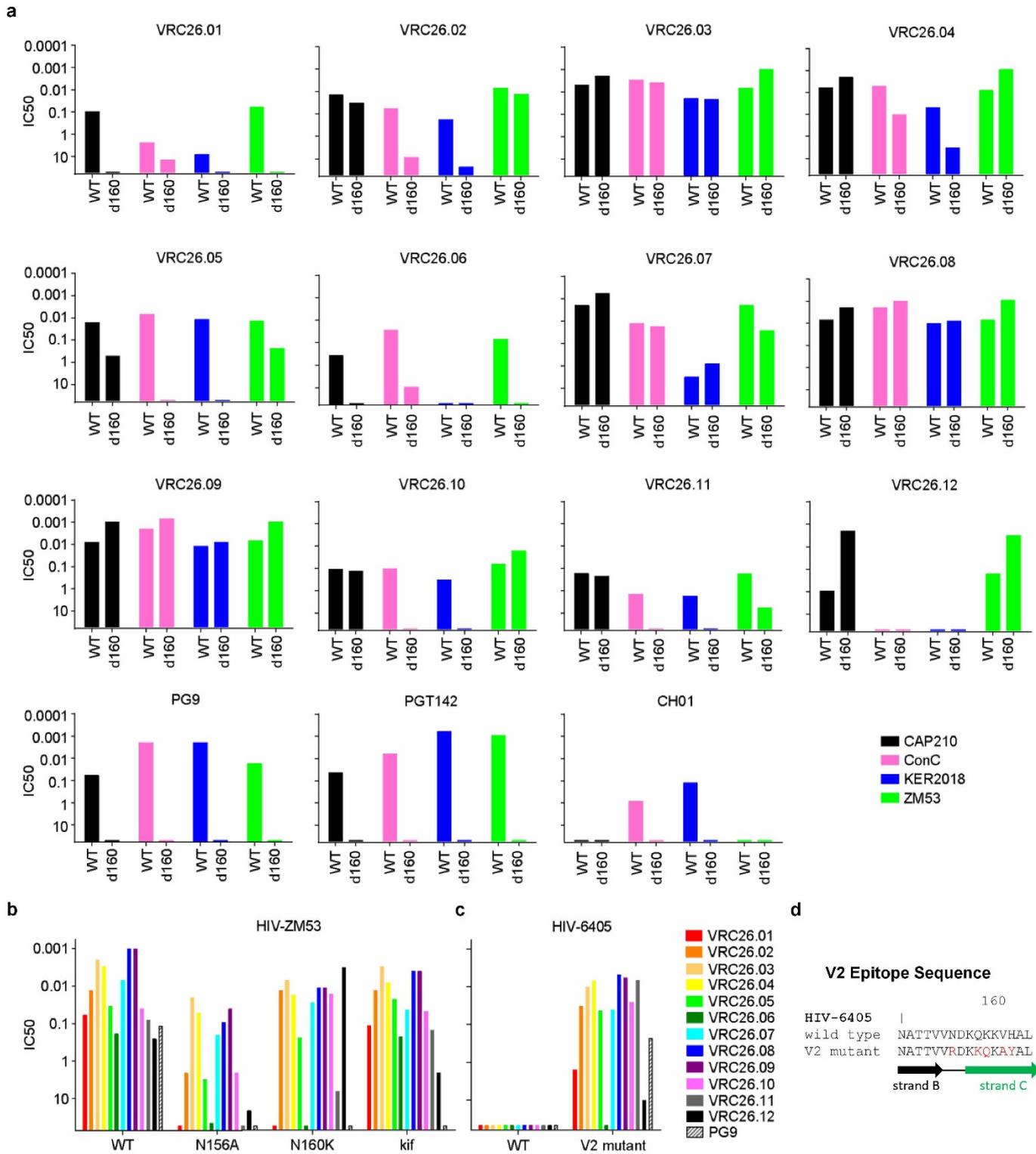
Extended Data Figure 3 | CAP256-VRC26 antibodies recognize a quaternary epitope. **a**, All 12 CAP256-VRC26 monoclonal antibodies were tested by ELISA against gp120 from ZM53 and CAP210. Positive control antibody PG9 bound to both gp120s (not shown). **b**, Twenty-three proteins and scaffolded V1V2 constructs 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 overlaid with PG9:SOSIP (purple mesh). The reconstructions are nearly identical in the trimer portion while displaying small differences in the Fab angles.



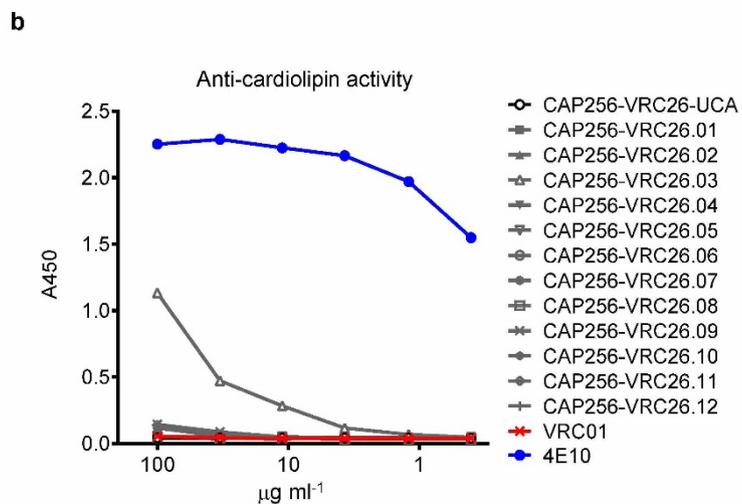
Extended Data Figure 5 | Effects of V2 mutations on neutralization activity of CAP256-VRC26 antibodies. **a**, Each panel shows neutralization of wild-type and N160 glycan mutant CAP210.E8, ConC, KER2018.11 and ZM53.12 viruses. CAP256-VRC26 monoclonal antibodies are partially and variably affected by 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 monoclonal antibodies 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, and its PG9-sensitive mutant³⁴ is also sensitive to CAP256-VRC26 antibodies. **d**, Sequences of wild type and mutant HIV-6405.

a

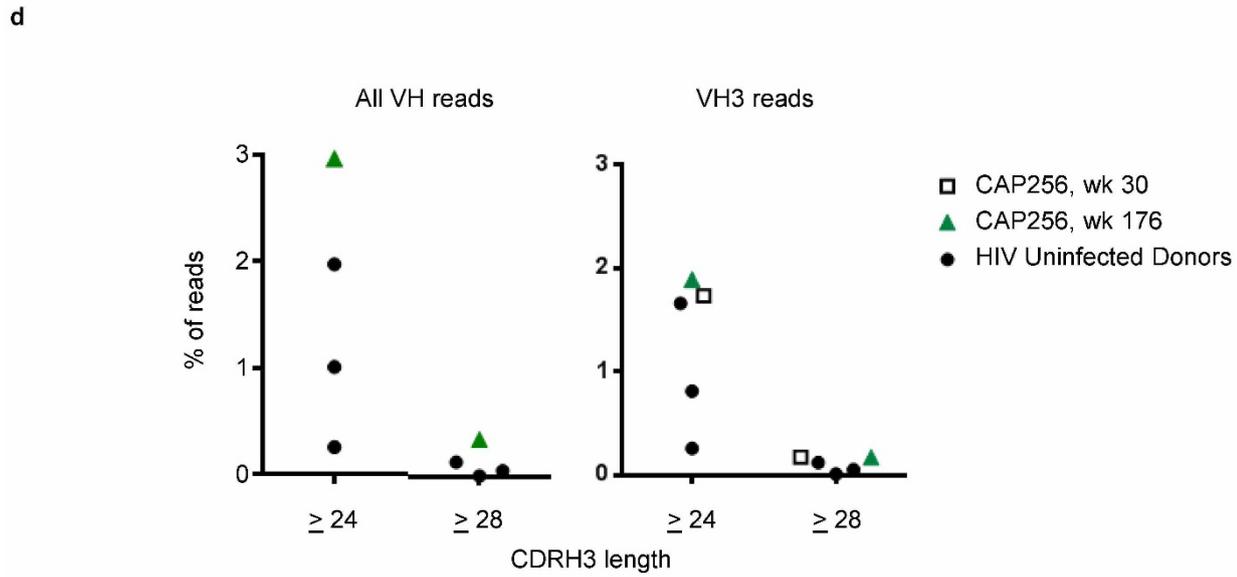
| | CDR H1 | | | CDR H2 | | |
|------------------|------------|-------------|-------------|------------|------------|-----------------------|
| UCA | QVQLVESGGG | VVQPGRSLRL | SCAASGFTFS | SYGMHWVRQA | PGKGLEWVAV | ISYDGSNKYY ADSVKGRFTI |
| 038-190676 (1x) | | | | | | |
| 038-195559 (1x) | |G..... |S..... | | | |
| 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..... |



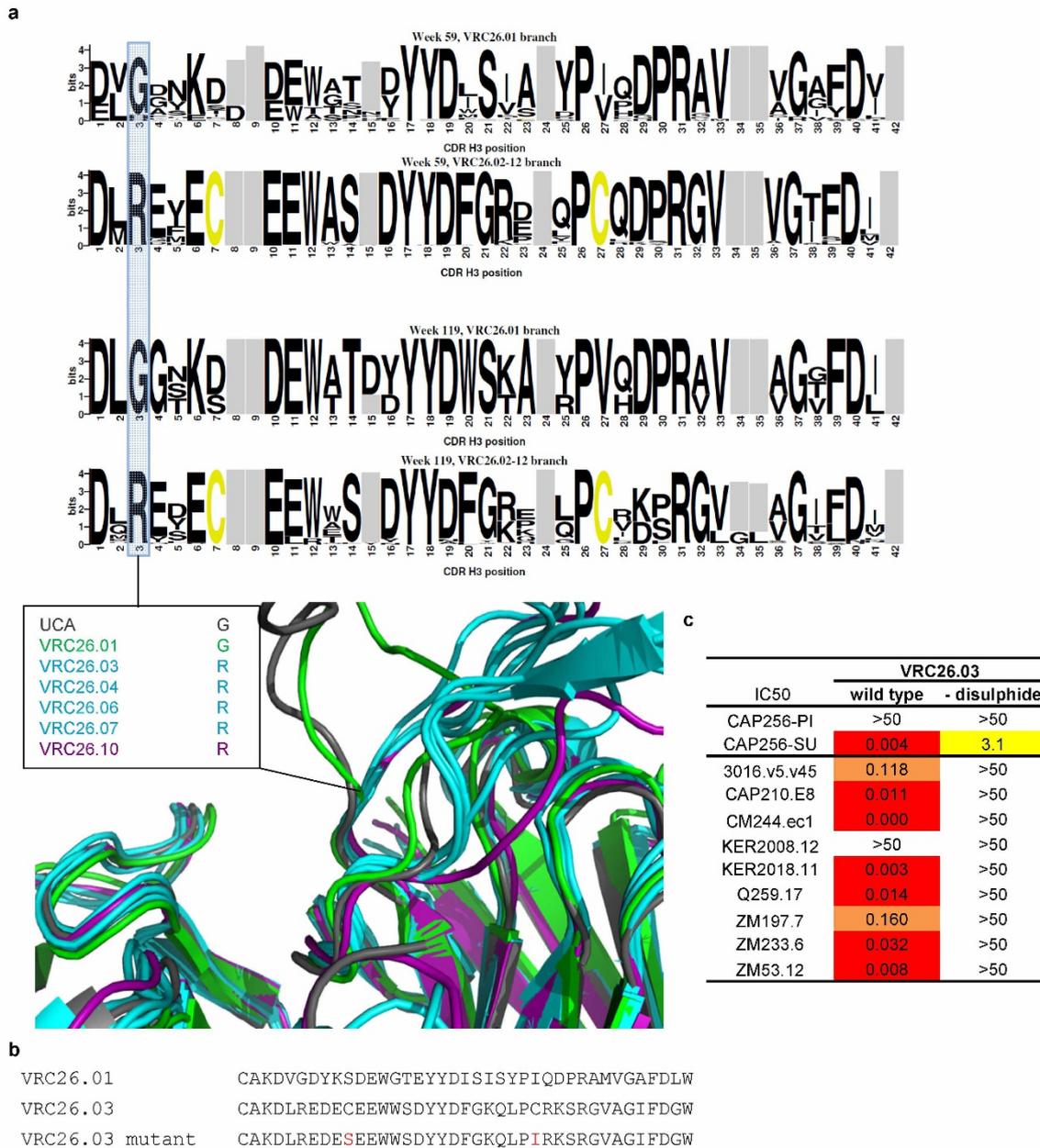
c

| 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 | + |



Extended Data Figure 6 | Origins of long CDR H3s 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 V_H and J_H. 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 monoclonal antibodies and the UCA were negative along with control antibody VRC01. **c**, Staining on

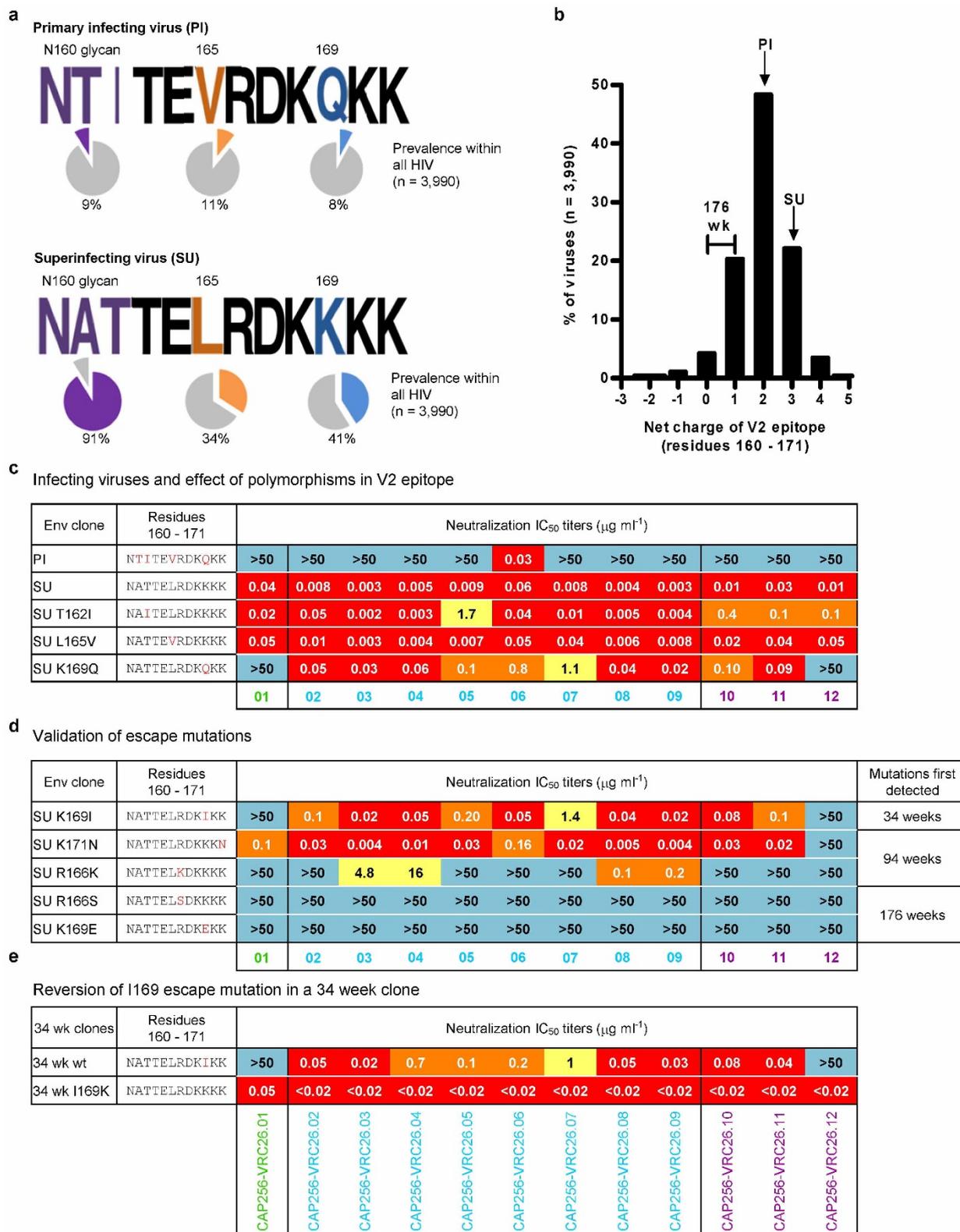
Hep2 cells was assessed at 50 and 25 µg ml⁻¹. Only the positive control, mAb 4E10, showed positive staining. **d**, Distribution of CDRH3 lengths among 454 sequencing reads of B cell transcripts. The percentage of high-quality NGS reads that have CDR H3 ≥ 24 or ≥ 28 are shown for three HIV-1 uninfected 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.



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

a, Top shows logograms of CDR H3 sequences extracted from the heavy chain phylogenetic tree from weeks 59 and 119. The height of each letter is proportional to its frequency in the population. Sequences that lack a disulphide bond contain a highly conserved glycine at the third position of the CDR H3 (residue 97, Kabat definition). The appearance of the two cysteines that form the disulphide bond coincides with a glycine to arginine mutation at this site. Bottom shows overlay and close-up of crystal structures from Supplementary Fig. 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 disulphide 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.



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. Coloured 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 monoclonal antibody neutralization of the SU and PI viruses, and of the SU virus mutated to contain

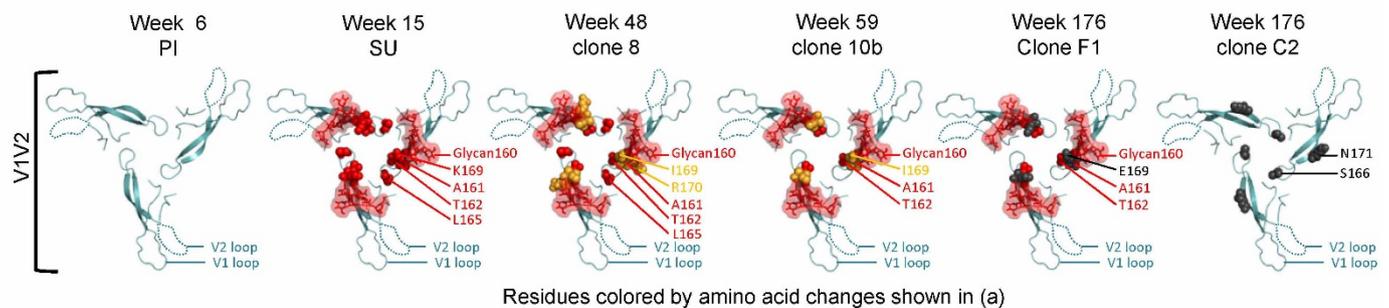
PI polymorphisms 162I, 165V or 169Q. **d**, CAP256-VRC26 monoclonal antibody neutralization of the SU virus mutated to contain known CAP256 escape mutations in the V2 epitope. **e**, CAP256-VRC26 monoclonal antibody neutralization of 34 week clone (designated wild type, wt) with an SU-like V1V2, compared to the I169K back mutant. **c–e**, The V2 epitope sequence, with mutated residues in red is shown on the left, IC₅₀ 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

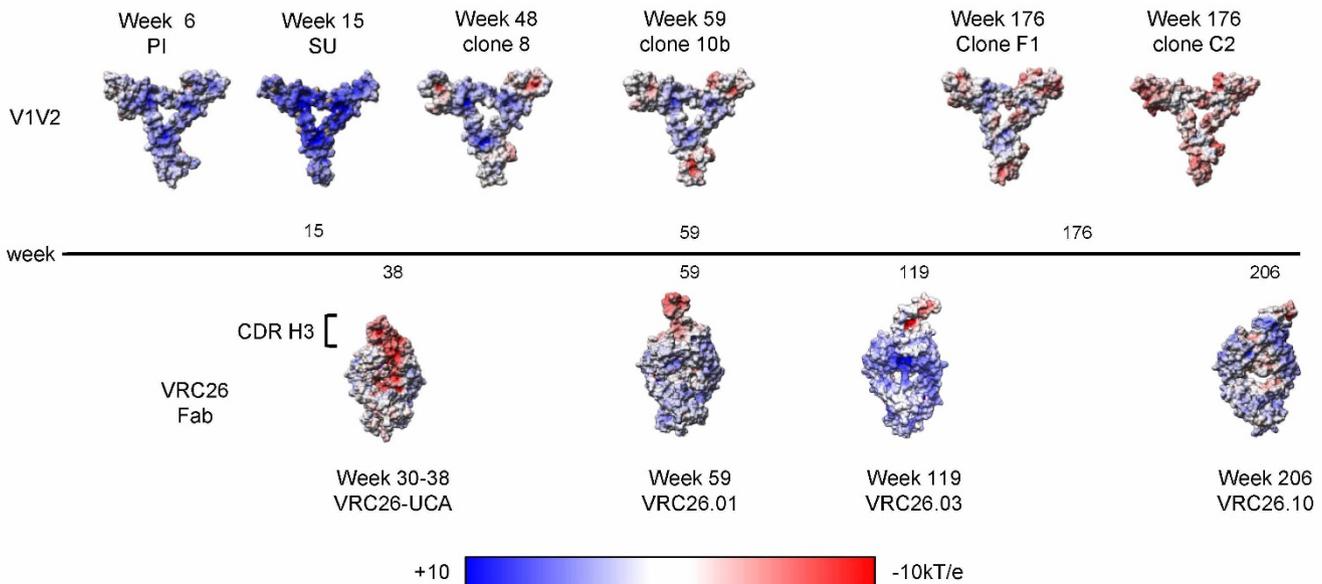
| Env clones | Residues | CAP256-VRC26 neutralization (IC_{50} $\mu g\ ml^{-1}$) | | | | | | | | | | | | Charge (monomer) | | |
|-------------------|---------------------------|--|------|-------|-------|-------|-------|------|-------|-------|-------|------|------|------------------|------|-----|
| | | 160 | 171 | UCA | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 | | .10 | .11 |
| PI | NTITEVRDKQKK | >50 | >50 | >50 | >50 | >50 | >50 | >50 | 0.03 | >50 | >50 | >50 | >50 | >50 | >50 | +2 |
| SU | NATTELRDKKK | >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 | 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 | NATTEVRDK E KK | >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 | |

<0.1
0.1-1
1-49
>50
 (IC_{50} $\mu g/ml$)

b Longitudinal variation of CAP256-VRC26 epitope



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



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 coloured as in panel **a** and

represented as spheres (amino acids) or stick and surface (glycan). **c**, Electrostatic surface representations of the full V1V2 region for each Env clone (top row), Fabs (bottom row). Timeline of infection is shown in the middle. V1V2 sequences were modelled 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

a

| Donor | Antibody | VH gene | VL gene | % Divergence, Nucleotide | | | |
|----------|-----------------|-------------|-------------|--------------------------|---------|------------|------------------|
| | | | | from VH | from VL | from UCA-H | from UCA-H CDRH3 |
| CAP256 | <i>VRC26-11</i> | IGHV3-30*18 | IGLV1-51*02 | 2% | 4% | 3% | 8% |
| | <i>VRC26-12</i> | | | 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% | | |
| | 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 |
| CAP256 | <i>VRC26-11</i> | IGHV3-30*18 | IGLV1-51*02 | 35 | 6% | 5% | 8% | 20% |
| | <i>VRC26-12</i> | | | 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% |
| | VRC26.06 | | | 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% | | |

a, b, Data are from the present study and from references 19–21. CAP256-VRC26.01–12 are derived from B cell culture, while CAP256.VRC26-11 and -12 (in italics) are inferred intermediates. CDRH3 lengths use Kabat notation. **a**, Nucleotides. **b**, Amino acids.