A genetic shift in an escaped transmitted/founder virus guides combinatorial vaccine design against HIV-1

Swati Jain[1], Gherman Uritskiy[1], Marthandan Mahalingam[1], Himanshu Batra[1], Subhash Chand[1], Hung Trinh[2], Charles Beck[4], Woong-Hee Shin[5,6], Wadad AlSalmi[1], Gustavo Kijak[2], Leigh Anne Eller[2], Jerome Kim[3], Daisuke Kihara[5,7], Sodsai Tovanabutra[2], Guido Ferrari[4], Merlin Robb[2], Mangala Rao[3], and Venigalla B. Rao[1*]

[1]Bacteriophage Medical Research Center, Department of Biology, The Catholic University of America, Washington, D.C, United States of America
[2]U.S. Military HIV Research Program, Henry M. Jackson Foundation, MD, United States of America
[3]U.S. Military HIV Research Program, Walter Reed Army Institute of Research, MD, United States of America
[4]Department of Molecular Genetics and Microbiology, Duke University, Durham, NC, United States of America
[5]Department of Biological Sciences, Purdue University, IN, United States of America
[6]Department of Chemical Science Education, Sunchon National University, Suncheon 57922, Republic of Korea
[7]Department of Computer Science, Purdue University, IN, United States of America

*Corresponding author E-mail: rao@cua.edu

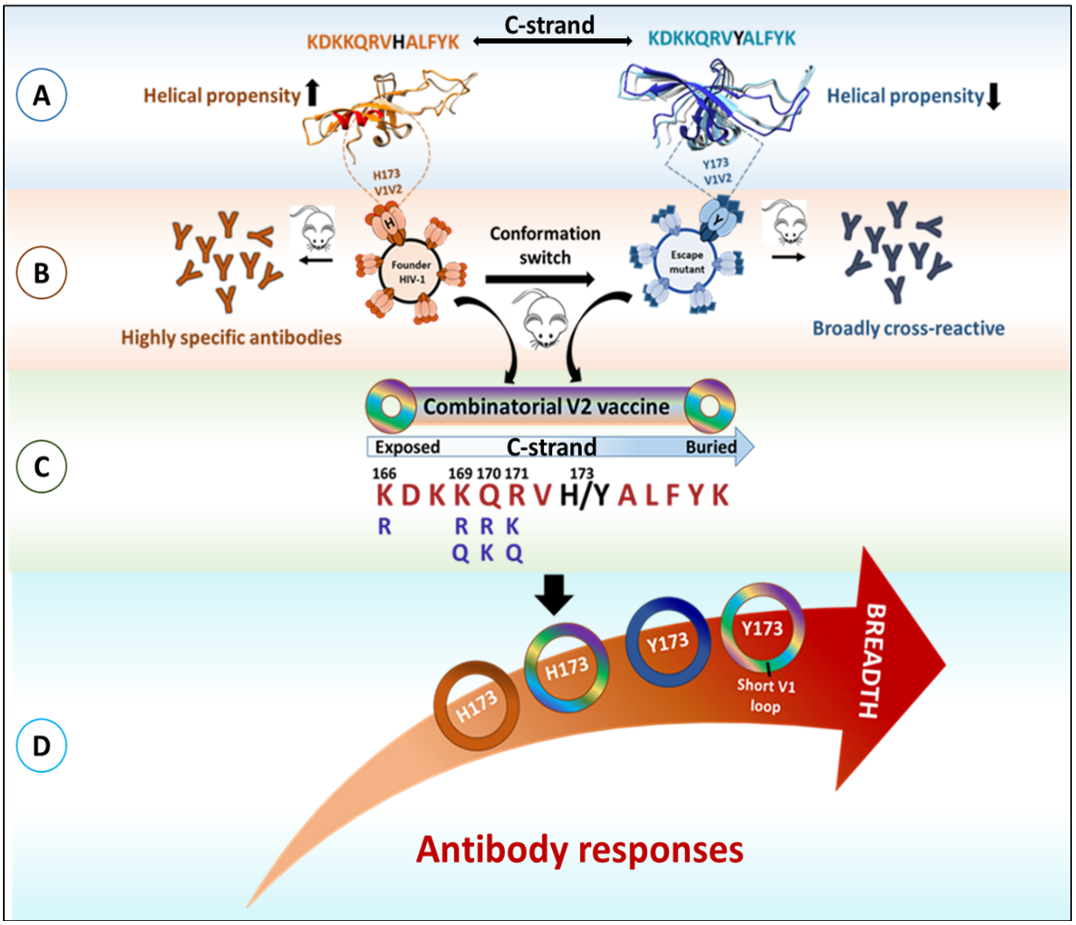
Abstract

A productive HIV-1 infection is often established by a single transmitted/founder (T/F) virus, which then evolves into a complex mixture of variants during the lifetime of infection. An effective vaccine should have sufficient breadth to block the entry of diverse T/F involved in different infections. Although the variable V1V2 domain of HIV-1 envelope protein (Env) is found to be a good target for vaccine design based on the correlates of protection in the modestly successful RV144 trial, the breadth of immune responses has to be substantially enhanced to improve vaccine efficacy and minimize the emergence of breakthrough infections. Here, we report a remarkable genetic shift in a T/F virus from a participant of an acute HIV infection cohort_RV217 study. It resulted in substitution of histidine at position 173 to tyrosine (H173Y) at week 24 (wk 24) after infection, coinciding with the disappearance of strictly H173-specific first wave antibodies. Intriguingly, a second wave antibodies emerged against the escaped Y173 variant that displayed increased breadth recognizing both the H173 and Y173 epitopes. This differential antibody responses towards variant epitopes were recapitulated in a mouse model. Structural analyses suggest distinct conformations for H173 and Y173 variants which might have led to antibody responses with different reactivity and breadth. Given the occurrence of conformational dynamism in the V2 region, combinatorial V2 vaccine candidates consisting of numerous conformational variants in the natural HIV-1 diversity were designed and tested as an immunogen. These libraries, especially the Y173 variant libraries that also contained a V1 loop deletion showed increased breadth and cross-reactivity to diverse HIV-1 Env proteins. This combinatorial design might be a powerful strategy in the future design of highly efficacious HIV-1 vaccine candidates.

Author summary

After numerous HIV-1 vaccine failures, only one human clinical study, RV144 has demonstrated protection with modest efficacy (31.2%) by virtue of antibody responses directed to V2-region of the Env that were correlated with reduced HIV acquisition. In natural infection, there occurs an evolutionary race between virus and the host immune system, as a result of which mutants of virus that escape the immune pressure are positively selected. Similarly, in vaccinated individuals, breakthrough infections occur when the infecting virus is able to escape from vaccine generated immune response compromising vaccine efficacy. Systematic learning of how these escape signatures are positively selected during the course of infection can have implications in designing a vaccine that can effectively counteract breakthrough infections. In the current study, we took an unprecedented approach of studying evolution of T/F viruses in acutely infected individuals to identify early V2-specific escape mutations and further study these mutations through biochemical, immunological and structural aspects. The knowledge obtained from these analyses was used to rationally design combinatorial vaccine libraries encompassing V2 variants mirroring natural HIV-1 population and assess their response in mice. The resultant vaccines generated antibodies were found to broadly cross-react diverse Env proteins. Such immune escape guided rationally designed vaccines have the potential to overcome breakthrough infections and improve vaccine efficacy.

77 # Striking image



78

79

80

81

82

83

84

85

Introduction

While combinatorial antiretroviral therapy (cART) has greatly improved the life expectancy of HIV patients, it doesn't cure the infection even with life-long commitment (1, 2). In the absence of a preventative vaccine, HIV-1 continues to be a global public health concern, causing 1.8 million new infections annually (3). After dozens of HIV-1 vaccine failures in the last four decades, the only vaccine trial that showed promise is the phase III RV144 trial conducted in Thailand (4-6). RV144 demonstrated an early efficacy of ~60% reduction in HIV acquisition at 12-months post-vaccination which gradually declined to 31.2% at 42 months (7, 8). Several studies demonstrated the correlation of IgG antibodies specific to the V1V2 variable domain of the HIV envelope protein (Env) to vaccine efficacy (9-11). Notably, protection was not due to their ability to neutralize the virus but most likely, due to their Fc effector function, specifically the antibody-dependent cell cytotoxicity (ADCC) (10). Furthermore, sieve analysis of the breakthrough infections in RV144 vaccinees showed mutations in the V2 domain, specifically the semi-conserved structural core encompassing residues 166 to 183 that seems to be the prime target of vaccine-induced immune pressure (12).

Env is expressed as a 160 kD glycoprotein (gp160) and cleaved by the cellular protease furin into gp120 and gp41 subunits. The membrane-external subunit, gp120, has five conserved regions (C1, C2, C3, C4 and C5) and five variable regions (V1, V2, V3, V4 and V5) that are alternately positioned in the Env sequence with the exception of V1V2 variable regions that assemble as a single domain (13, 14). Three protomers, each composed of non-covalently associated gp120 and gp41 subunits, assemble as a trimeric spike on the viral envelope. The V1V2 domain forms a well-exposed "crown" of the mushroom-shaped spike, hence a frequent target of

the host immune system. Each of the three V1V2 domains consists of a conserved Greek-key motif structure with 4-5 β-strands (A, B, C, C', D) forming an anti-parallel β-sheet and two hypervariable loops that are expected to be flexible and conformationally dynamic (15). The virus takes advantage of region to engage into immune battle with the host defenses undergoing mutations particularly in the hypervariable loops (11, 16-18). Inter-protomeric interactions within V1V2 moieties are also responsible for the stability and infectivity of a native functional trimer.

The Env spike on the surface of HIV undergoes an essential interaction with its primary receptor, CD4 followed by CCR5/CXCR4 co-receptor on the CD4+ T cell for viral entry (19-22). In addition to CD4 and CCR5/CXCR4 receptors, interaction of V2 region on the Env with an integrin, α4β7 has also been implicated as a significant contributor in the pathogenesis of HIV-1 dissemination and gut-reservoir establishment in the infected individuals (23, 24). Recently, it has been shown that V2 domain by virtue of mimicking MadCAM, a natural ligand of α4β7, assists in co-stimulation of CD4+ T cells promoting HIV-1 replication during an acute stage of infection (25). RV144 vaccinees generated non-neutralizing V2 antibodies are also shown to block interaction with α4β7 (26-28).

The V1V2 domain of HIV-1 envelope protein is therefore an attractive target for vaccine design and V2-directed responses have the potential to block HIV acquisition. At the site of exposure, the human host is often exposed to a complex genetic pool of highly diverse and heterogeneous viral quasispecies from an infected donor of which only one (or a few) transmitted founder (T/F) virus can successfully establish a productive infection. However, studies systematically addressing how T/F viruses escape from V2-directed Ab responses resulting in breakthrough infections and hence poor vaccine efficacy, are largely limiting. Furthermore, the

mechanisms by which HIV-1 might exploit the structural and conformational variability of V2 domain as an immune-escape strategy remained poorly understood.

Here, we investigated how a T/F virus escape from V2-directed responses in the acute phase of infection to obtain insights for designing a vaccine that can be broadly effective against diverse strains of HIV-1. Our study identified a mutation in the semi-conserved epitope of the V2 region of the Env of a T/F virus in one of the RV217 study participants that led to virus-escape during an early phase of natural infection. Furthermore, our data demonstrates a switch in V2-epitope conformation owing to this mutation in such a way that the resultant conformation elicited broadly reactive antibody responses. These findings led to a combinatorial vaccine design by incorporating potential escape variants to increase the breadth of antibody responses against the V2 domain.

Results

Overall experimental design

We hypothesized that understanding how the T/F viruses escape host immune pressure at the very early stages of infection might identify variations that if included in a HIV vaccine could stimulate broader immune responses and interfere or block HIV acquisition. Our primary focus is the V1V2 domain of HIV-1 envelope protein as it is one of the key determinants of virus escape. We therefore studied the evolution of V1V2 domain in T/F viruses from four HIV-1 infected participants of RV217 Early Capture HIV Cohort (ECHO) project (Fig 1).

The vaccine design process involved a series of steps. First, we performed phylogenetic analyses of longitudinal Env sequences to identify V1V2-specific variants that correlate with T/F

151 virus escape (Fig 1a). Second, the selected escape variants were fused with the dodecameric
152 bacteriophage T4 terminase protein, gp16, to generate V1V2 epitope-displaying nanoscaffolds.
153 Since gp16 is highly soluble and assembles into stable oligomers, 11-mers and 12-mers (29), this
154 design would symmetrically display multiple V1V2 domains for effective antigen presentation
155 (Fig 1b). Third, these constructs were expressed in GnTi cells to produce gp16-scaffolded V1V2
156 escape mutant domains with native-like high mannose glycosylation (Fig 1c). Fourth, the
157 immunogenicity of the V1V2 variants was evaluated in a mouse model to determine if they could
158 stimulate escape mutant-specific immune responses (Fig 1d). Finally, cocktails of V1V2 variant-
159 targeted immunogen pools were created to determine if such pools would elicit immune responses
160 with increased breadth that recognize diverse HIV-1 Env antigens (Fig 1e). If broadening of V1V2
161 immunity occurred, it would inform on more effective HIV-1 transmission-targeted vaccine
162 designs for clinical trials.

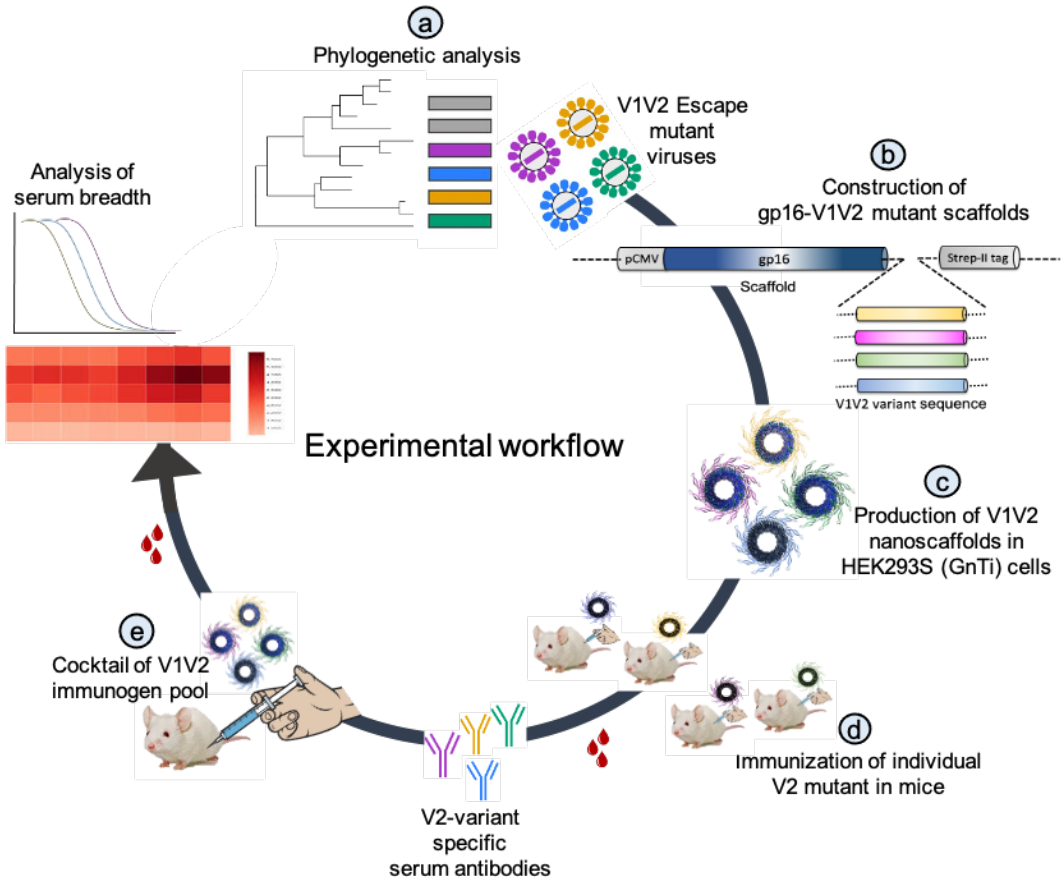


Fig 1. Experimental design. The experimental design consisted of multiple steps labeled a-e. (a) Phylogenetic analysis of longitudinal viral *env* sequences isolated from four RV217 study participants during acute HIV infection. (b) Selection of V1V2-specific escape variants and their fusion to the C-terminus of bacteriophage T4-derived small terminase protein, gp16 to generate gp16-V1V2 nanoscaffolds representing various escape mutations. (c) Expression of gp16-V1V2 escape mutants in GnTi mammalian cells to produce dodecameric nanoscaffold immunogens. (d) Immunization of V1V2 variant nanoscaffolds in mice and evaluation of their immunogenicity. (e) Rational design of combinatorial V1V2 mutant immunogens and evaluation of their immunogenicity and breadth of immune responses.

Phylogenetic analyses identified a striking H173Y escape mutation in V2 domain

A series of HIV-1 viral *env* sequences were isolated from four HIV-1 infected participants of the RV217 ECHO project through single genome amplification (SGA) (Reference IDs: 40007, 40061, 40094, and 40100). Using a sensitive nucleic acid test, each study participant was confirmed of HIV-1 positivity just days after a negative test (Fiebig stage I) and none were put on on antiretroviral therapy (ART) during the timeframe of the study. *Env* sequences were obtained

178 at three time points, wk 1, 4 (~1 month), and 24 (~6 months) following the positive test. That a
179 single T/F virus was responsible for infection in each participant was ascertained by aligning
180 independently isolated *env* sequences from wk 1 plasma, which were nearly identical (30). Plasma
181 viral load with peak, nadir, and set point viremia shown for one representative patient (40007)
182 indicated a typical pattern of early captured infection (Fig S1). Around 30 sequences were analyzed
183 from each patient (median of 10 sequences per visit), with 152 sequences in total across all four
184 patients, over a period of up to 6 months post-infection.

185 To trace viral phylogeny in each of the T/F viruses (hereafter referred to as T/F07, T/F61,
186 T/F94 and T/F100), the longitudinal *env* sequences were translated to protein sequences and
187 aligned. Multiple sequence alignments were then used to construct phylogenetic trees using the
188 respective T/F virus *env* sequence as root for the tree construction. Of the four T/F viruses, T/F61
189 showed a few dominant mutations, occurring in more than 50% of the viruses, in the V1V2 loop
190 region, while T/F100 was found to be the most rapidly diverged virus with mutants appearing at
191 as early as one wk after infection. By wk 4, a major branch of diverging T/F100 viruses harboring
192 mutations in the V1V2 domain appeared. In addition, there were mutations in V5 variable loop
193 and the less conserved α2 helix of C3 constant region. T/F94 viruses also acquired various
194 mutations including deletions in the variable V1V2 and V5 regions. Not surprisingly, most of these
195 mutations are in the surface-exposed variable regions of HIV-1 trimer with hotspots in the loop
196 regions (Fig 2A-B).

197 On the other hand, T/F07 virus did not show any variants until wk 4. At 24-wk post-
198 infection, however, nearly the entire T/F07 virus population shifted to a single variant containing
199 two mutations, one at position 173 that changed histidine to tyrosine (H173Y) and another at
200 position 236 that changed lysine to threonine (K236T). The K236T mutation restored the well-

conserved N-linked glycan at position N234 located at gp120/gp41 interface. Additionally, in some of the variants, a 3-residue deletion in the variable V2 loop (ΔDSV) and a 5-residue deletion (ΔNTTRFL) in the variable V4 loop co-occurred with the H173Y mutation (Fig 2C-D). The H173Y mutation is localized in the relatively well-conserved "C" β-strand positioned at the junction of the cationic first half and the hydrophobic second half of the β strand (Fig 2E). These characteristics, a singular variant and striking shift of viral population, strongly suggested a linkage between the variant and the viral escape. This was in contrast to many mutations observed in the other three T/F viruses in the hypervariable V1 and V2 loops (Fig 2F) which are difficult to track and of little value for vaccine design.

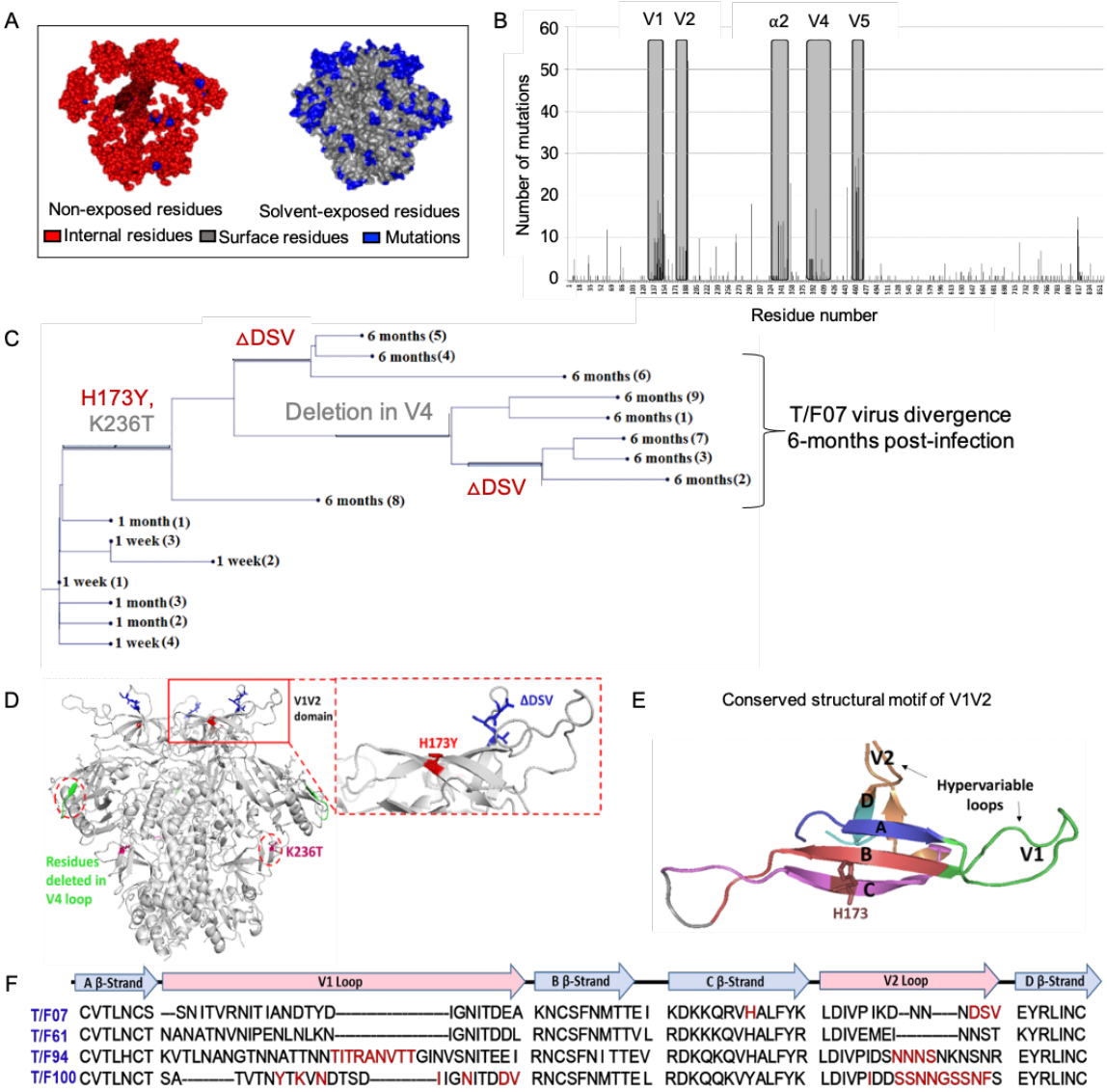


Fig 2. Viral escape mutations in four RV217 participants during the acute stage of HIV infection. (A) Mutations observed in the RV217 participants 40007, 40061, 40094, and 40100 are placed on HIV-1 BG505 pre-fusion trimer structure (PDB ID: 4TVP). Surface model showing dominant mutations (blue) falling on the solvent-exposed regions of the Env (gray) but not on the non-exposed or buried region (red). Modeling was done with PyMol (ver. 1.74) molecular visualization software and surface-exposed residues were defined as all residues that had >5Å^2 exposure to the solvent. (B) Mutational hotspots in the *env* sequence based on the genetic diversification of T/F viruses from all four participants. Total number of mutations in a particular region is plotted on the y-axis against residue positions on the *env* sequence (x-axis), with reference to HXB2 strain. The Env regions (labelled on top) with gray background showed high frequency of mutations. (C) Phylogenetic tree displaying T/F07 virus evolution in 40007 participant. The evolutionary tree was constructed by the neighbor-joining method, rooted to the T/F07 virus sequence. The viruses are designated corresponding to the time post-infection, 1 wk, 1 month (4-wks)) or 6-month (24 wks) at the nodes of the branches. Prominent diverging mutations are labeled on the respective branches of the tree in red (V1V2 region) and gray (another region). (D) Dominant mutations that occurred until 24-wk post-infection in 40007 participant are modeled on a ribbon model of the T/F07 trimer, generated through homology modeling using BG505 trimer (PDB ID: 4TVP) as a template. The zoom in image of V1V2 domain is shown to highlight the positions of V2-specific mutations, H173Y (red) and 3-residue deletion, DSV (blue). Deletion in the variable, V4 region is depicted in bright

green and mutation in the conserved C2 region, K236T substitution, is shown in magenta. (E) A color-coded 4-5 β-stranded (A-D) conserved Greek-key motif structure of V1V2 domain is represented showing residue 173 on the C-strand. (F) Snapshot of V1V2 sequences of T/F viruses from each RV217 participant under study. Major structural features of V1V2 region; semi-conserved four β-strands (A, B, C and D) and hypervariable V1 and V2 loops are labeled on top of the sequence. Prominent V1V2-specific mutation sites (observed in >50% of circulating viruses) until 24-wks are highlighted in red in the respective T/F virus sequence isolated from 40007, 40061, 40094 and 40100 participants.

The H173Y mutation in C β-strand of V2 domain is a key determinant of virus escape against host immune pressure

The divergence of nearly the entire T/F07 virus population to H173Y variant made it a strong candidate for a viral escape mechanism. H173 is located in the C β-strand of V1V2 domain, a region that in previous studies was also found to be a critical target for host immune responses by RV144 trial vaccines. To determine if this mutant indeed arose through a strong selection against V2-directed antibody responses, the epitope specificity of antibodies was evaluated in the longitudinal plasma samples. The V1V2 domains, but not the full-length Env proteins, were used for testing in order to exclude the binding responses directed against other regions of the envelope protein. Four V1V2 domain nanoscaffolds were constructed with a C-terminal Strep-tag: gp16-H173 ("wild-type" T/F07) and three 24-wk V2 mutants namely, gp16-H173Y (Y173), gp16-ΔDSV, and gp16-Y173+ΔDSV. These were expressed in HEK-GnTi (GnTi) cells and affinity-purified by StrepTactin chromatography (Fig S2). Binding of 40007 longitudinal plasma samples to these V1V2 variants was assessed by surface plasmon resonance (SPR) assay.

The epitope specificity data revealed remarkable specificity of V2-specific antibodies present at 24-wk post-infection to the original T/F07 virus but not to the H173Y variant. The antibodies bound strongly to V1V2-H173 and V1V2-H173.ΔDSV scaffolds but failed to bind Y173 variant scaffolds V1V2-Y173 and V1V2-Y173.ΔDSV (Fig 3A). These data demonstrated

that the H173Y mutation is a key determinant in epitope switching at the time of virus escape against host immune pressure. The DSV deletion mutation did not appear to play a direct role, though it might have an accessory role (see below).

In addition to this first "wave" of antibodies specific to H173 epitope, there was a second wave after the virus population switched to the resistant Y173 variant. SPR analyses showed that, contrary to the first wave, this second wave of antibodies exhibited increased breadth, recognizing both the V1V2-H173 and V1V2-Y173 variants (Fig 3A; see the peak for 1.4 years). Therefore, the second wave antibodies would be able to restrict both the original T/F07 H173 virus as well as the escaped Y173 variants, further supporting the hypothesis that the V2-region, in particular the C β-strand, is a critical target for mounting host immune pressure during acute HIV infection.

Recapitulation of escape mutant specificity in monoclonal antibodies from RV144 vaccinees

Next, we evaluated the binding of H173 and Y173 variants to C β-strand-specific monoclonal antibodies (mAbs) CH58 and CH59 isolated from RV144 trial vaccinees. These antibodies are specific to the H173 C β-strand present in the gp120 immunogen used in the RV144 vaccine trial. As mentioned above, the moderate protection observed by the RV144 vaccine correlated with such V2 C β-strand-specific antibody responses. Remarkably, while the V1V2-H173 scaffolds bound strongly to both CH58 and CH59 antibodies, the V1V2-Y173 escape variant either failed to bind (CH59) or showed drastically reduced binding (CH58). The DSV deletion again did not show a significant alteration in the binding specificity (Fig 3B-D).

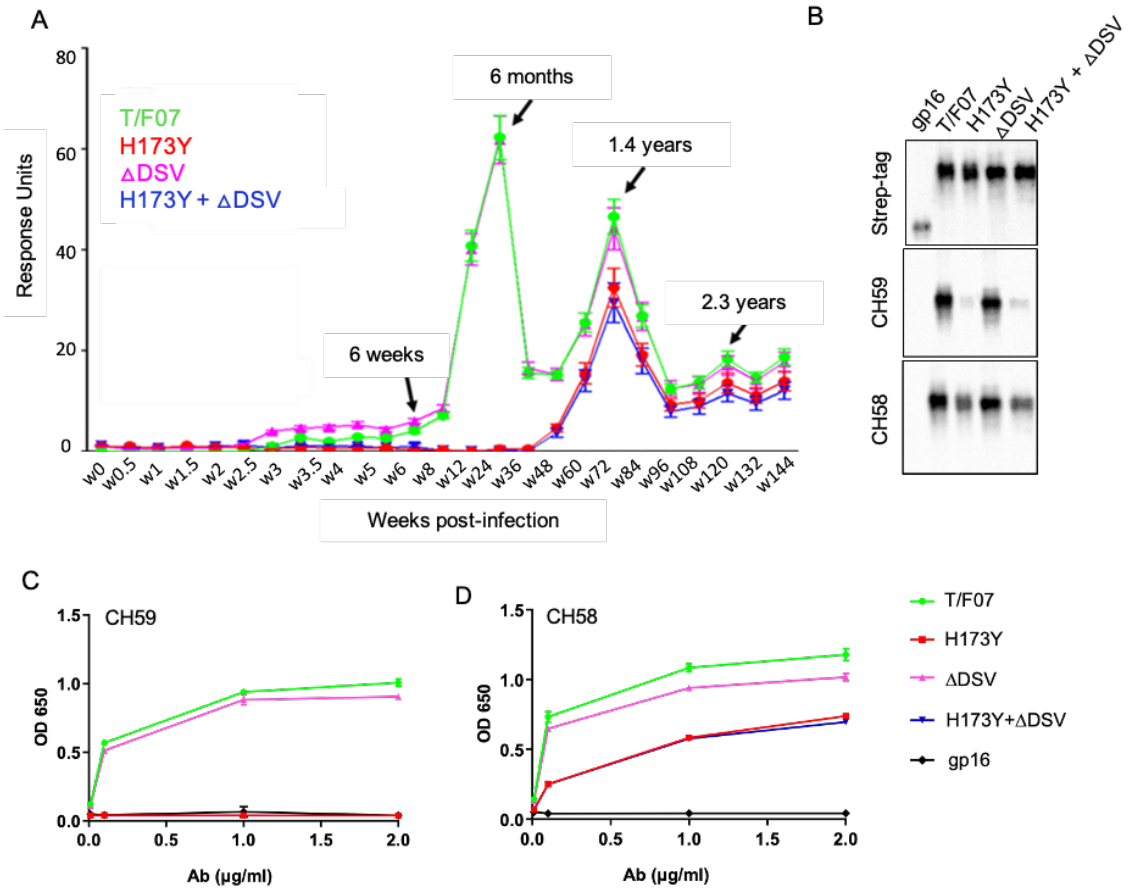


Fig 3. H173Y mutation in T/F07 virus led to virus escape in the 40007 participant. (A) SPR determined binding curves showing reactivity of purified recombinant gp16-T/F07 (green), gp16-H173Y (red), gp16-ΔDSV (magenta) and gp16-H173Y+ΔDSV (blue) V1V2 proteins with a series of longitudinal plasma samples collected from 40007 (shown on x-axis). The samples from wk 0 (pre-infection) until wk 144 (post-infection) were tested. The amount of binding is proportional to the response units (RU) plotted on the y-axis. The corresponding time points for each peak of antibodies are indicated with an arrow. An early wave of V2-specific antibody represented by first peak at 6-month time-point recognized gp16-T/F07 and gp16-ΔDSV [H173 variants] but not to gp16-H173Y and gp16-H173Y+ΔDSV [Y173 variants] implying H173Y mutation mediated viral escape. The second and third peaks of antibodies were found to be reactive to both the variants. (B-D) H173Y mutants poorly react to CH59 and CH58 antibodies. Immunoblot of the purified gp16 T/F07 V1V2 mutants (labeled on the top) showing respective binding with V2 mAbs, CH59 and CH58. Recognition by Strep-tag (purification tag) antibody served as a protein loading control (B). Binding curves of gp16-T/F07 (green), gp16-H173Y (red), gp16-ΔDSV (magenta), gp16-H173Y+ΔDSV (blue) V1V2-proteins, and gp16 scaffold only (negative control) (black) showing reactivity to CH59 (C) and CH58 (D) mAbs, as determined by ELISA.

We then evaluated antibody dependent cell cytotoxicity (ADCC) responses as these were identified as one of the correlates for protection in RV144 vaccinees (31-33). We determined the relative ADCC responses in the 40007 plasma against the C β-strand variants. We constructed

gp120 ectodomain versions of the V1V2 domain variants; gp120-H173 (T/F07), gp120-Y173 and gp120-Y173.ΔDSV (escape mutants), and gp120-92Th023 (control), expressed in GnTi cells, and purified the recombinant proteins. These were then coated on the target cells and ADCC killing by effector cells (healthy PBMCs) was measured by Granzyme activity using wk0, wk4 and wk24 plasma samples of participant 40007. The data showed moderate ADCC responses against H173 T/F07 while resistance to killing was observed for Y173 mutants, with the Y173.ΔDSV double mutant showing more resistance than the Y173 single mutant (Fig 4). This implicates a potential functional role for ΔDSV mutation against ADCC-type immune pressure, though it had no significant impact on recognizing the binding epitope(s). Furthermore, significant resistance was also observed for both the Y173 variants against CH58 mAb mediated ADCC killing, while no resistance was detected against a negative control CH65 Flu antibody.

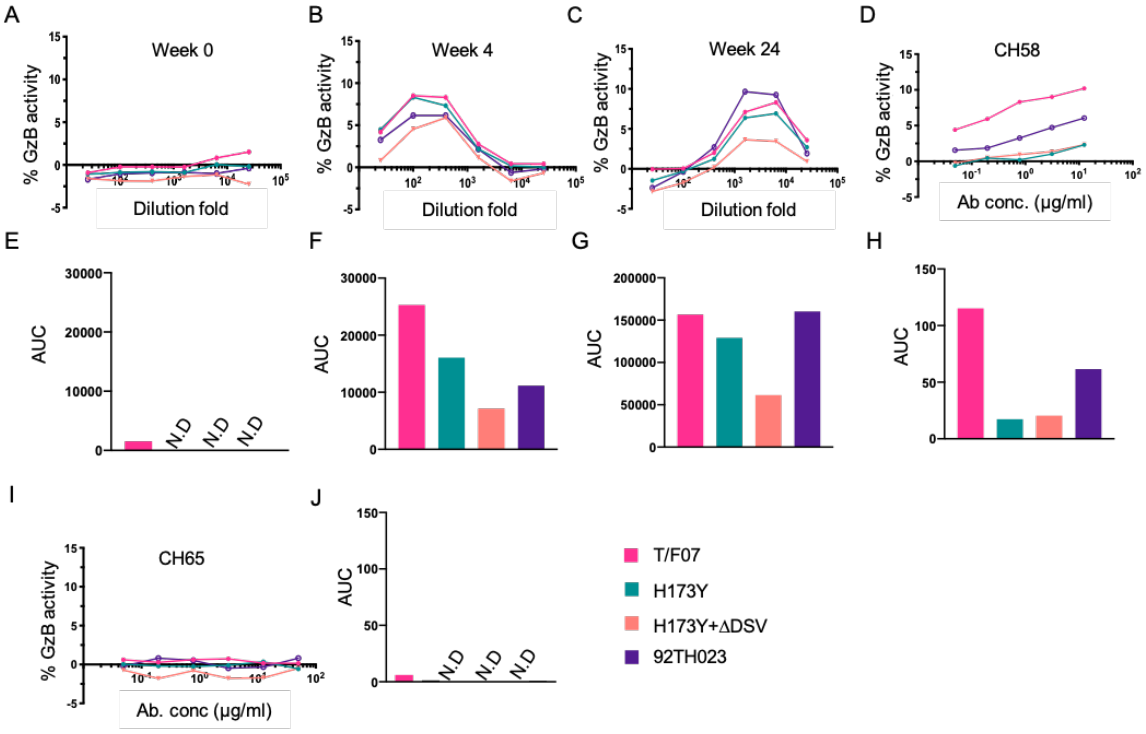


Fig 4. ADCC responses in longitudinal plasma samples of participant 40007. (A-C) ADCC responses at wk 0 (A, pre-infection), 4-wks post-infection (B), 24-wks post-infection (C) of 40007 plasma samples, measured as percentage Granzyme activity (y-axis), against gp120-T/F07 (pink), gp120-H173Y (teal), gp120-H173Y+ΔDSV (orange) and

control antigen, gp120-92TH023 (violet) coated target cells. (E-G) The corresponding area under curve (AUC) values for A-C graphs are shown for each curve. (D and I) CH58 (Positive control) (D) and CH65 (negative control) (I) mediated ADCC responses are shown. (H and J) Corresponding AUC values are plotted for panel D and I respectively. For CH65 and CH58 mAbs, 4-fold serial dilution starting at 50 μg/ml was used as shown on the x-axis. The 40007 plasma samples from each visit were 5-fold serially diluted to determine the ADCC responses.

The above sets of data strongly implicate that the H173Y mutation in the diverged T/F07 mutant population was due to selection against the early V2-directed host antibody responses that restricted the survival of T/F07 HIV-1 virus.

Recapitulation of V2-specific human immune responses in mice

We then hypothesized that the dramatic escape of H173Y mutant viruses might be because the histidine to tyrosine substitution caused a significant structural/conformational change in the C β-strand epitope such that it is no longer recognized by H173-specific antibodies. There is evidence that the C β-strand is conformationally dynamic and that it can take a helical form when bound to certain antibodies (34, 35). This is also consistent with the distinct specificities of human antibodies generated against these variants in participant 40007, i.e., strict specificity of T/F07 plasma for H173 and increased breadth in the case of escaped plasma for both H173 and Y173 (Fig 3). If so, could this be recapitulated through immunogen design, in the absence of viral infection? To address this question, we immunized BALB/c mice with V1V2 nanoscaffolds containing four different V2 variants; H173, Y173, ΔDSV, and Y173.ΔDSV, and analyzed the specificity of the elicited antibody responses (Fig 5A).

V2-specific antibody titers were determined by ELISA using the GnTi-expressed and purified His-tagged gp140 Env proteins containing the respective H173, Y173, ΔDSV, or Y173.ΔDSV mutations as coating antigens. The gp140 Env was used to evaluate V2 epitope specificity in a native context, and it also filtered out the gp16 scaffold-specific and strep-tag-specific antibody titers. The data showed that all the scaffolds elicited robust V2-specific

antibodies that recognized the V2 epitopes in the context of gp140 Env (Fig S3 and Fig 5B). This is noteworthy because in previous reports, the V2 antibodies elicited against other scaffolded V2 immunogens did not recognize the V2 epitopes in gp140 envelope protein, though they recognized the epitopes in the context of the respective scaffold used for immunization (36). The end point titers were on the order of ~10^5 in the immunized mice while the naïve controls showed no significant V2-specific titers. These results demonstrated that the antigenicity and immunogenicity behaviors of the V1V2 scaffolds are consistent with the retention of a native-like conformation.

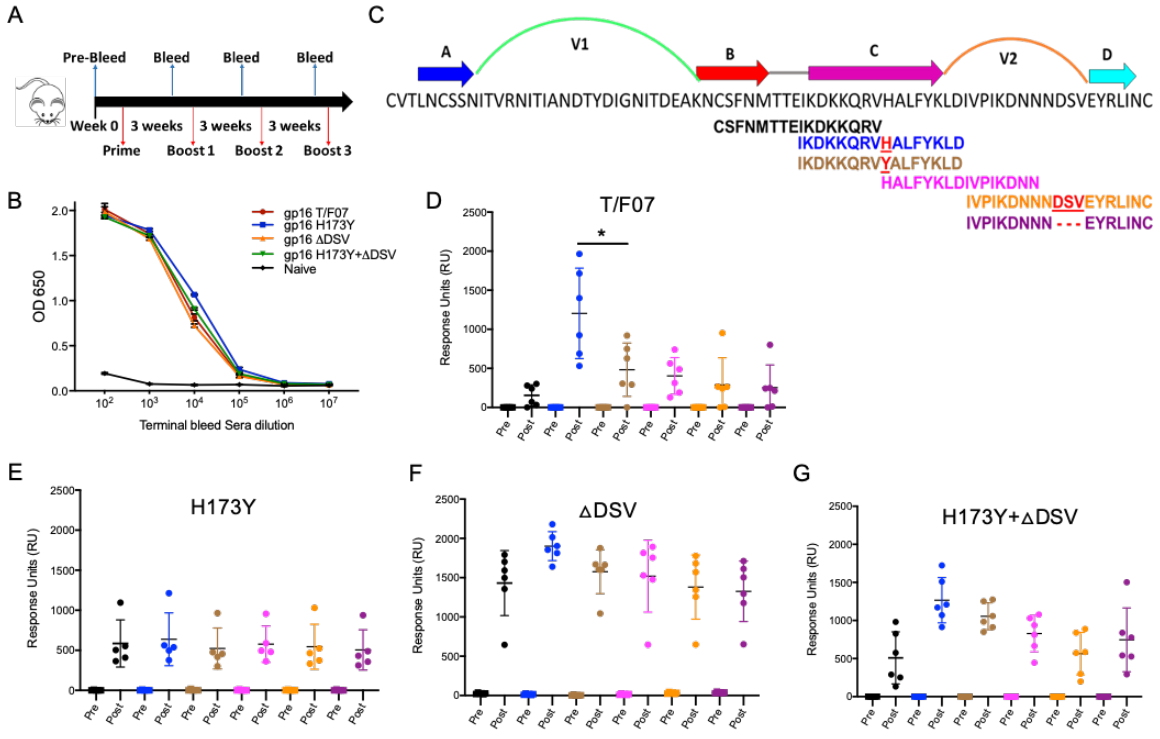


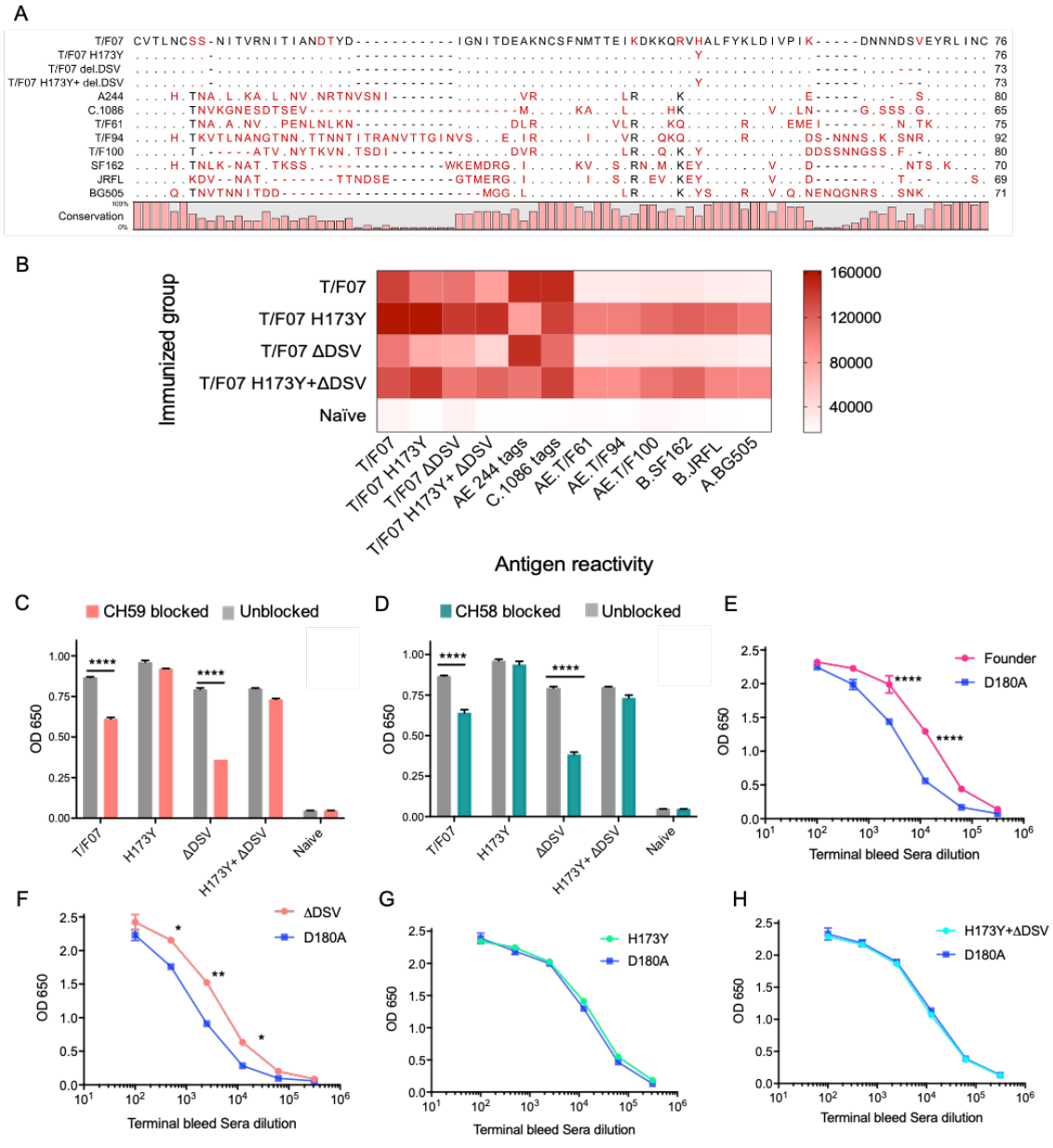
Fig 5. Binding analysis of immunized mice sera with overlapping V2-peptides. (A) Scheme of immunizations. Prime plus 3-boost immunizations were performed at an interval of three wks. Pre-immunization sera were collected as negative controls. (B) V1V2-specific responses in different immunized groups after the last immunization (terminal bleed). Binding curves for each group are color-coded, provided in the legend on the right side of the graph. (C) Schematic showing T/F07-V1V2 sequence labelled on top for A-D strands and hypervariable V1 and V2 loops. The overlapping V2 peptides used for SPR binding analysis with immunized mice sera are shown in different colors. (D-G) SPR binding responses shown for gp16-T/F07 (D), gp16-H173Y (E), gp16-ΔDSV (F), gp16-H173Y+ΔDSV (G) immunized mice terminal (post) and pre-immunization sera (pre) (negative control) with specific V2 peptide represented by the corresponding color of peptide as shown in (C). The binding signal is depicted in terms of response units shown on the y-axis. P-value was determined through unpaired t-test, * = $p < 0.01$.

We then tested the epitope specificity using a series of 15-mer biotinylated peptides spanning the C β-strand by a sensitive Biacore SPR assay (Fig 5C). Remarkably, the T/F07 H173-induced sera reacted strongly with H173 C-strand peptide but poorly with Y173 and other variant peptides (Fig 5D). In contrast, the escape mutant Y173-induced sera showed broad reactivity to both H173 and Y173 peptides as well as to other variant peptides containing C β-strand epitope. However, the level of reactivity of Y173 antibodies to peptides was low overall when compared to the same with gp140 Env proteins probably because these antibodies are conformation-specific (Fig 5E). Thus, the mice sera in principle recapitulated the behavior of H173 and Y173 antibodies produced in a human infection. Furthermore, we observed that the presence of ΔDSV mutation enhanced and broadened the reactivity of the antibodies (Fig 5F-G). These results are consistent with the ADCC assays where the ΔDSV mutation showed enhanced resistance to cell killing.

Distinct specificities of antibodies induced by H173 and Y173 variants

To further define the specificities of H173- and Y173-induced sera, we evaluated their ability to recognize HIV-1 Env proteins from different clades and determine their cross-reactivity. Accordingly, we constructed a series of recombinant clones and purified gp140 Env proteins from different clades including: CRF_AE proteins T/F07-H173, T/F07-H173Y, T/F07-ΔDSV, T/F07-H173Y.ΔDSV, T/F61, T/F94, and T/F100, clade A [BG505], clade B [SF162 and JRFL], and clade C [1086] (Fig 6A). Their reactivity was tested using sera of mice immunized with the four V1V2 nanoscaffolds (Fig S4). The data showed that the T/F07 H173- and ΔDSV-induced antibodies reacted with autologous T/F07 Env proteins or clade C-1086 and clade AE244 strain antigens having C-strand sequence closely resembling to T/F07, whereas the antibodies induced by Y173 and Y173.ΔDSV groups reacted broadly with all the proteins tested, and overall, more strongly than the T/F07-induced sera as shown by the heat map (Fig 6B). These data suggested that the

375 H173 immunogen induced antibodies with narrow specificity to C β-strand whereas the Y173
376 immunogen induced more broadly reactive antibodies.



377

378 **Fig 6. H173 and Y173 V2-variants induce distinct antibody responses.** (A) Sequence alignment of V1V2 region
379 of all the antigens used in the cross-reactivity ELISA experiment. The sequences are compared with T/F07 V1V2
380 sequence. T/F07 matching residues are shown as dots and different residues are highlighted in red. Difference in the
381 lengths of hypervariable loops is shown by dashed lines (gaps). Degree of conservation is depicted graphically at the
382 bottom of the sequence alignment. (B) Heat-map showing cross-reactivity for each immunized group. The map is
383 generated based on AUC values calculated from the binding curves shown in fig S4 for each immunized group. The
384 antigens used in the binding experiment are labelled horizontally at the bottom of the heat-map. The mice groups are
385 labelled vertically on the left-side of the heat map. The color gradient scale on the right side shows the degree of

reactivity with corresponding numerical values. (C-D) In an antibody blocking assay, significant reduction in sera reactivity was observed in gp140 T/F07 coated wells pre-incubated/blocked with purified mAbs, CH59 (C, orange), and CH58 antibodies (D, teal), for gp16-T/F07 or ΔDSV (H173 variants) but not gp16-H173Y or H173Y+ΔDSV (Y173 variants) immunized mice groups, compared with the unblocked wells (gray). Titrated and optimized sera dilution was used for each group in this assay. (E-H) Sensitivity to D180A mutation. Binding curves showing reactivity of gp16-T/F07 (E), H173Y (F), ΔDSV (G) and H173Y+ΔDSV (H) immunized sera to respective autologous T/F07-gp140 (with matching V2 mutations) coating antigens (color-coded curves) versus T/F07 gp140-D180A mutant (blue curves). Binding is determined through ELISA. Triplicate absorbance (OD 650 nm) readings were used to generate binding curves. P-values were determined through unpaired t-test, **** = $p < 0.00001$, *** = $p < 0.0001$, ** = $p < 0.001$ and * = $p < 0.01$.

Next, we tested whether the narrowly specific H173-induced antibodies are similar to CH58 and CH59 mAbs derived from RV144 vaccinees. Since the immunogen used in RV144 trial contains H173 and that the CH58 and CH59 mAbs specifically recognized H173 C β-strand epitope but not the Y173 variant, it is reasonable to hypothesize that H173, but not Y173, might induce CH58/59-like antibody responses. This was tested by blocking assays using CH58 and CH59 mAbs. The H173 gp140 Env antigen was coated on ELISA plates and after blocking by CH58 or CH59 mAbs, were exposed to H173 or Y173 mice sera. Remarkably, the H173 sera, but not the Y173 sera, with or without ΔDSV, showed significant reduction in binding (Fig 6C-D). Furthermore, since CH58 and CH59 binding is sensitive to mutation at D180 residue that is adjacent to C β-strand (35), we evaluated its binding. Consistent with the above antibody blocking data, binding of H173-induced antibodies but not of Y173-induced antibodies was significantly more sensitive to D180A mutation (Fig 6E-H). Overall, these data suggest that the presence of histidine at position 173 favors induction of CH58/59-like antibodies elicited by RV144 vaccine.

The above datasets show that the mouse H173-induced antibodies, like their human counterparts, are narrowly specific to the autologous C β strand and hence sensitive to sequence variation. Conversely, the Y173-induced antibodies are broadly reactive, conformation-requiring, and cross-reactive to diverse V2 domains that differ in length, sequence, and glycosylation.

Furthermore, these antibodies also tolerate variations in C β-strand sequence not only at 173 position but also at other critical positions such as K168 or K169 (12, 35).

Structural analyses indicate conformational switching in virus escape

To determine if the mutational switch involved in virus escape might be due to a conformational switch in C β-strand, we performed Molecular Dynamics (MD) simulations and structural modeling analyses of H173- and Y173-V1V2 domains. For MD simulations, we first modeled H173- and Y173 V1V2 domains based on our recently published cryo-EM structure of CRF_AE T/F100 HIV-1 Env trimer as template (PDB ID: 6NQD), due to extensive sequence similarity (87.6%) (30), using MODELLER 9v7 (37). Both of these V1V2 domains assumed β-stranded conformation matching the template. The V1V2 domains were extracted to run simulations using GROMACS 5.1.2 (38) and trajectories were produced for 100 ns with 2fs time step. We explored how H173 and Y173 strand V1V2 models undergo changes in conformation over time.

In H173/Strand but not in Y173/Strand trajectory, helix content increased over time, though the β-pairing within the domain remained stable. In contrast to H173/Strand trajectory, Y173/Strand trajectory displayed a dynamic and fluctuating β-strand content (Fig 7A, left). We further investigated how V1V2 conformation would change if C-strand was helix in the initial conformation. To construct a helix initial model, helical restraints were put on residues, 167-176 (DKKQRVH/YALF), following the helical conformation of the published crystal structure of V2/C-strand peptide (PDB ID: 4HPO). The helical structures were also modeled using MODELLER 9v7. Different from the strand conformation simulations, trajectories with helix initial conformations showed unstable trajectories for both H173 and Y173 V1V2 domains. The overall β-strand content increased around 30 ns in both the simulations (Fig 7A, right). However,

while helix in Y173/Helix model was fully unwound during simulation, one turn encompassing residues 166-171 remained in H173/Helix until the end of the trajectory (Fig 7B-C). Furthermore, the unwound region of H173/Helix but not Y173/Helix model could still engage into β-pairing. Overall, the MD simulations data suggested distinct conformational dynamics for H173 and Y173 V1V2 domains with the latter being relatively more dynamic. Furthermore, while the conserved β-sheet conformation is thermodynamically favored and hence predominates for V1V2 domain, H173/C-strand region could tolerate helical constraint owing to the stable β-pairing in the rest of the domain.

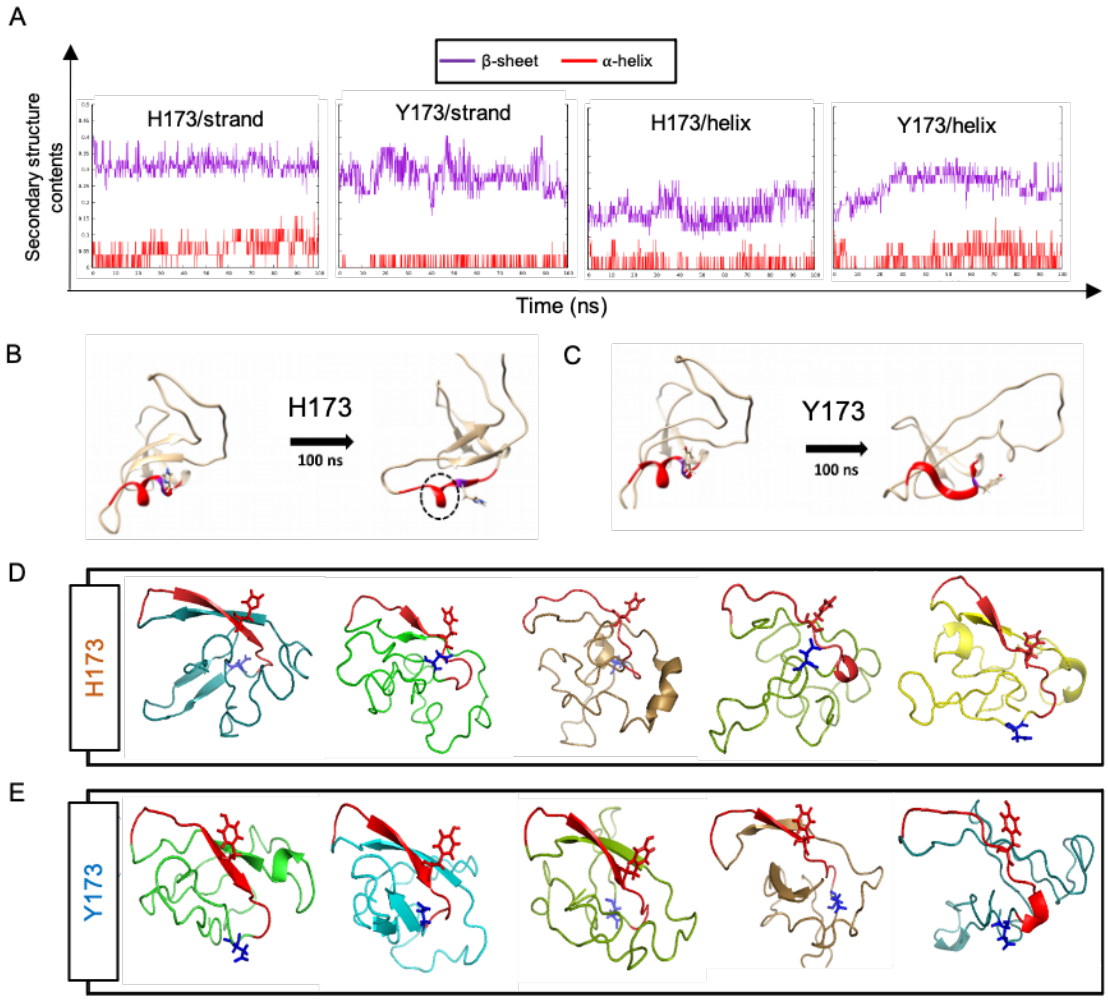


Fig 7. Structural analyses of H173 and Y173-V1V2. (A) Secondary-structure content variation as a function of time (ns) initiated with V2/strand conformation (left) and V2/helix conformation (right) of H173 and Y173 V1V2 models during MD simulation. β-sheet content (purple) and α-helix content (red) changes are depicted over the 100 ns time-course of simulation. (B-C) Initial (left) and final (right) structures of the V2/helix trajectory for H173 (B) Y173 (C) V1V2 models. Histidine and tyrosine side chain is shown in the model with C-strand highlighted in red. A portion of the helix remains stable (encircled by black dashed lines) at the end 100 ns simulation time in H173/helix trajectory. (D-E) Ab initio structure modelling of H173 and Y173 V1V2 domains. Five models were generated through QUARK modelling tool for H173-V1V2 (D) Y173-V1V2 (E). C-strand is colored in red with sidechain shown for residue 173. Sidechain for residue 180 (critical for helix/coil V2-conformation recognizing CH58/59 like antibodies) is shown in blue.

Since the template-based modeling tends to get biased towards the template structure, we next performed structural modeling of H173 and Y173 V1V2 domains using a modelling tool QUARK (39) that generates *ab initio* structure predictions based on physical principles rather than previously resolved structures as templates. Five 3D models depicting possible conformations were generated for each V1V2 variant based on replica-exchange Monte Carlo simulation under the guide of an atomic-level knowledge-based force field. It was found that only one of the five H173-V1V2 models depicted C-strand as β-strand, while the rest showed a coil plus a short β-strand or a full coil conformation (Fig 7D). In contrast, four of the five Y173-V1V2 models depicted C-strand in β-stranded conformation while only one model showed this region to assume coil conformation (Fig 7E). Overall, these models indicate that H173Y mutation has the potential to cause a structural change in the C β-strand thereby altering the conformational dynamics of the V1V2 domain. Furthermore, notably, models of both the variants depicted the C strand as β-strand to some degree which represent the conserved Greek key motif captured by all the resolved Env trimer structures. However, the degree of β-stranded character was identified to be much higher in Y173 than in H173 variant. Hence, it is plausible that the observed antibody evasion and virus escape that occurred in participant 40007 was due to a structural transition in this V2 epitope, from helix to β-strand owing to H173Y mutation.

A combinatorial approach to V2 domain vaccine design

The H173Y mutation leading to structural transition, viral escape, and induction of broadly reactive antibody species provided a conceptual basis to design an ensemble of V2-conformation variants that can potentially induce even more broadly cross-reactive V2 antibodies that might be difficult for the virus to overcome without compromising survival fitness. Therefore, we developed a combinatorial approach to create an ensemble of V2 variants that might also mimic the natural diversity of HIV-1.

First, we sought to identify the potential escape sites in the semi-conserved C β-strand, deliberately excluding the loop region that primarily generates strain-specific immune responses. HIV sequence database was explored to extract 100 Env sequences from each of the major geographically prevalent HIV subtypes including A, B, C, and AE. These sequences were individually aligned using CLC Main Workbench and the partly conserved C strand sequence alignment was extracted to generate a consensus logo (Fig 8A). Subsequently, sites of highest variability were identified in the consensus logogram with the rationale that a less conserved or highly variable site is likely to be linked to viral escape. This resulted in the prediction of four residues-at amino acid positions 166, 169, 170, 171 as highly variable. Then, primers were designed such that the most common variants in the natural HIV-1 population, lysine (K), arginine (R), or glutamine (Q), were incorporated at each of these positions to generate a combinatorial library (Fig 8B). The resultant library encompassed 54 V2-variants representing the V2 conformations prevalent in the HIV population. The rationale behind choosing naturally selected mutations is not only to represent diverse HIV strains but to also ensure the structural and functional integrity of the V1V2 domain, thereby presenting only the most relevant epitope diversity to the immune system.

Three different combinatorial gp16-V1V2 nanoscaffold libraries were constructed in the background of H173, Y173, and Y173.ΔDSV templates. Deletion of V1 or V2 loop has previously been shown to modulate immunogenicity of the Env protein and subsequent antibody responses (40-42). Hence, we constructed a fourth library using Y173 template in which a 15-amino acid residue mutational hotspot in the V1 loop (SNITVERNITIANDTYD) was replaced with a flexible linker (AGGAS), the length of which was optimized through *insilico* structural modeling to have a minimal effect on the V1V2 backbone conformation. This Y173.ΔV1 library is also supposed to eliminate certain immunodominant residues in the V1 loop and enhance V2-directed antibody responses. These four combinatorial libraries plus two original H173 and Y173 V2 immunogens as controls, all as gp16 nanoscaffolds were constructed and expressed in GnTi cells and the recombinant proteins were purified. To minimize epitope distraction, the Strep-tags were removed using HRV3c protease cleavage site engineered into each of the constructs and the protease was separated by size-exclusion column chromatography (Fig S5). The pure tag-less nanoscaffolds were then used for mouse immunizations and induction of V2-specific antibodies was evaluated.

Combinatorial immunogens broaden V2 antibody responses

BALB/c mice were immunized at wks 0, 3, 6, and 12 with pure V2 immunogen libraries as described above (Fig 5A). Naïve mice (PBS/no antigen) and mice immunized with gp16-scaffold alone (without V2) served as negative control groups. Sera were collected after the final boost and analyzed by a series of immunological assays.

V1V2-specific antibody titers were quantified by ELISA using the respective purified proteins as coating antigens. V1V2 antibodies were detected after prime immunization and enhanced by several fold with each successive boost for all the groups except for the negative

control groups where no V2-specific responses were detected. The terminal bleed sera having the maximum antibody titers were then used for detailed epitope specificity studies (Fig S6).

To determine the breadth of responses, a series of ~20 heterologous recombinant gp140 and gp120 Env proteins from diverse HIV subtypes A, B, AE and C were used as coating antigens for ELISA assays. Many of these were purified from GnTi cells while some were obtained from NIH Reagent Program (Fig S7A-C and Table 1). These Env proteins differed in sequence composition of V1V2 domain C β-strand with significant differences in the length, sequence, and glycosylation of their hypervariable V1 and V2 loops (Fig S7D). All of these recombinant proteins were used as coating antigens to determine the cross-clade antibody responses generated by different combinatorial libraries (Fig S8).

Consistent with the first mouse study, the Y173-induced antibodies showed greater cross-reactivity than the H173-induced sera, as shown by their binding to majority of the heterologous Env proteins. In contrast, H173-induced sera strongly reacted with Env antigens from A244, 93TH75, 1086 and CM235 strains that contained similar C β-strand sequence as the T/F07 virus, while showing moderate to poor reactivity with the rest of the Env antigens as depicted in the heat map (Fig 8C). Upon grouping the antigens into T/F07 matched and mis-matched C β-strand sequence, H173 antibody responses were more sensitive to mismatches in the C β-strand region than the Y173 antibodies (Fig 8D). Cumulative response towards all the antigens was also compared for H173 and Y173 groups by plotting the summarized response based on AUC values (%) (Fig 8E). These data clearly showed that with single residue change, H173Y, there occurred significant improvement in recognition of diverse Env antigens and hence breadth of V2 reactivity.

The combinatorial H173-induced antibodies, like the H173-induced antibodies, were significantly inhibited by CH58 antibodies while no significant inhibition was observed for sera

induced by Y173 or Y173 combinatorial groups (Fig 8F). These data suggest that histidine at 173 position plays a dominant role in influencing the conformation even in the context of a library of V2 variants. Furthermore, since CH58 recognizes V2 region as α-helix, it is consistent with our analyses described above that the H173 immunotype has more propensity for helical conformation over β-strand (Fig 8G).

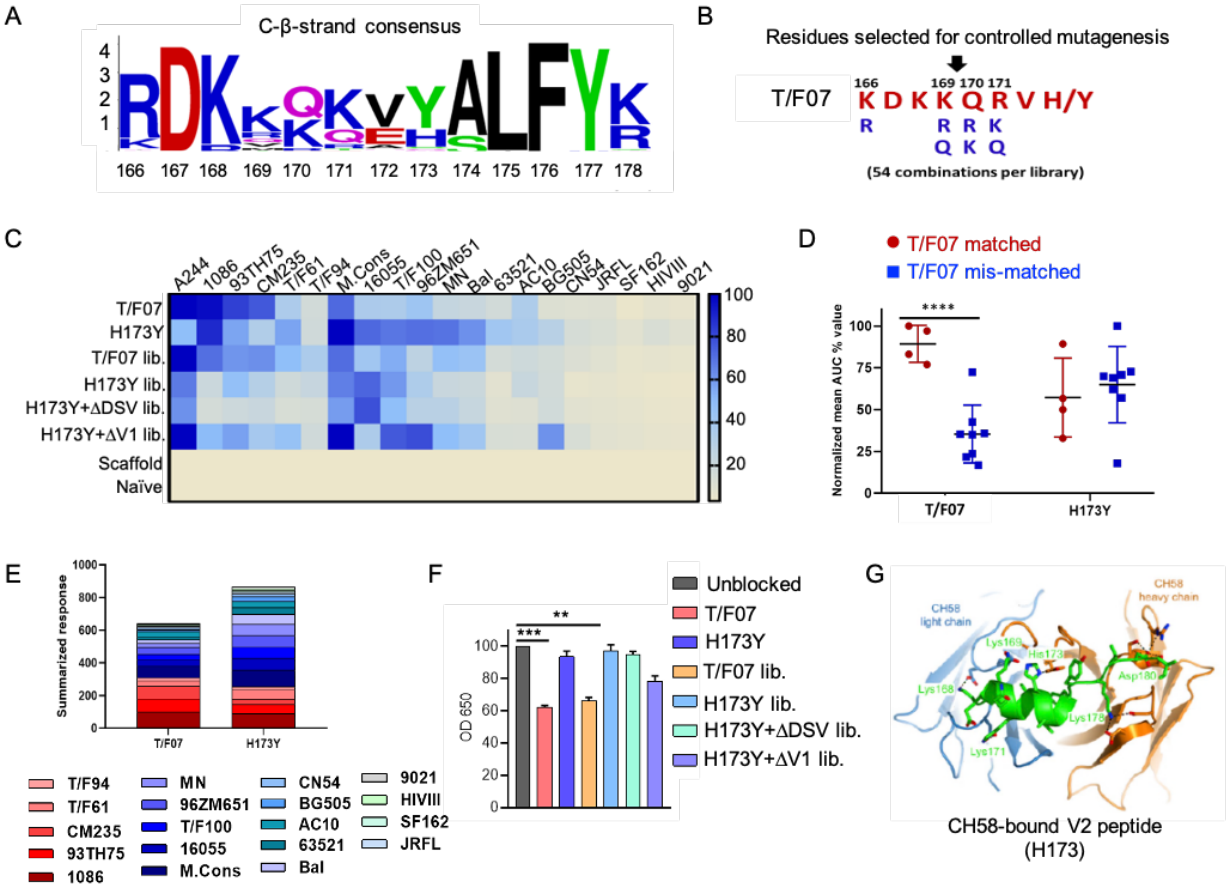


Fig 8. Combinatorial and non-combinatorial V2-immunogens induce differential breadth and antibody responses. A) Consensus logo of C-ß-strand (residue 166 to 178) of the V1V2 region generated using 100 *env* sequences each of clade A, B, C, D and AE. The size of the letter depicting residue in the logo is proportional to its conservation across viral sequences under analysis. B) Four highly variable residue positions (166, 169, 170 and 171) were selected for combinatorial mutagenesis with additional substitutions shown in blue to generate V2-combinatorial libraries. C) Heat-map showing binding of sera from different immunized mice groups to diverse HIV-1 Env antigens. Percent AUC values derived from the binding curves and normalized to autologous antigen binding response were used to generate the heat-map. The antigens used in the binding experiment are labelled horizontally at the top of the heat-map. The immunized mice groups are labelled vertically on the left-side of the heat map. The color gradient scale on the right side shows the degree of reactivity with corresponding AUC percent values. D) Effect of C-strand sequence variation on the binding of T/F07 and H173Y immunized groups' sera. Scatter plot showing significant

difference in the reactivity of gp16-T/F07 but not H173Y immunized sera with Env antigens having C-strand sequence closely matching (red spheres) versus mismatching (blue squares) to T/F07. Normalized mean AUC (%) values are plotted on the y-axis estimated from the binding curve for each antigen represented as sphere or square in the graph (each sphere or square represent different antigen). E) Cumulative AUC (%) values are plotted for T/F07 and H173Y groups to display the summarized or total response from all the diverse antigens in the library. Each antigen is depicted as a small rectangle colored with respect to the key provided at the bottom of the graph. Binding to each antigen corresponds to the area of rectangle in the bar graph. F) CH58 antibody blocking assay data. Significant reduction in sera reactivity was observed for T/F07 and T/F07 library groups when gp140 T/F07 coated wells were pre-incubated/blocked with purified mAb, CH58 (grey) compared to unblocked (pink). No statistically significant inhibition of binding signal was observed for any H173Y-based immunogen groups. P values determined using unpaired t-test, ** = $p < 0.01$ and *** = $p < 0.001$ (significant difference). G) Crystal structure of V2 peptide with H173 (bright green) bound to mAb CH58 (orange) (PDB ID: 4HPO). The image is adapted from Liao et al. (Liao et al., 2013). CH58 recognize C-strand of V2 region as an α-helix.

Finally, unlike the H173 group, the combinatorial H173 group showed relatively broader responses, with much improved binding to Env antigens such as T/F61, T/F94, T/F100, MN and Bal. Hence, H173 library-based vaccine candidates led to enhanced recognition of other Env proteins. With respect to Y173 and Y173.ΔDSV combinatorial libraries, strong binding was seen with all Y173 antigens and moderate reactivity towards H173 bearing antigens. On the other hand, notably, the sera induced by Y173.ΔV1 combinatorial library showed strong overall binding to both H173 and Y173 antigens. Though the breadth of reactivity for this group was comparable to a non-combinatorial Y173 group, it represented a distinct binding profile because of its stronger reactivity to most of the Env antigens such as A244, T/F100, 96ZM651, and BG505 than the H173Y group containing the V1 loop (Fig 8C). Thus the Y173.ΔV1 libraries containing both the Y173 mutation and V1 loop deletion might be considered the most effective design for the greatest breadth and strength of the V2 domain antibodies.

Discussion

The modest 31.2% efficacy of the only successful RV144 HIV vaccine trial in Thailand was correlated with the induction of V2-directed antibodies (10, 17, 35, 43). Analysis of breakthrough infections of vaccinees showed mutations in the V2 domain of the circulating

viruses, presumably selected for their ability to survive under the host immune pressure. If the antibody responses had increased breadth, vaccine efficacy would have been greater, and it would have minimized the emergence of viral escape mutants. Hence, increasing the breadth of vaccine-induced V2 response remained as one of the critical goals of HIV-1 vaccine design. Here, we report detailed analysis of a genetic shift of a transmitted/founder HIV-1 virus through viral escape that guided the design of combinatorial V2 immunogens for increased breadth and cross-reactivity.

Phylogenetic analyses of T/F viruses in longitudinal samples of RV217 ECHO trial participant 40007 identified a remarkable, near-complete genetic shift of virus population at 24-wk post-infection. The viruses carried a histidine to tyrosine substitution at position 173 of the semi-conserved C β-strand of the V2 domain. This shift coincided with a "wave" of H173-epitope specific antibodies produced by the host immune system. That the Y173-epitope showed no detectable reactivity to these antibodies while the H173 epitope reacted strongly gave a clear indication that the Y173 substitution was an escape mutant selected for its survival under strong immune pressure. Furthermore, the Y173 epitope either did not bind, or bound poorly, to CH58 and CH59 mAbs isolated from RV144 vaccinees. These antibodies are also directed to the same C β-strand and recognize H173 as a critical residue. In fact, two of the three vaccine immunogens (A244 and 93TH023) used for RV144 trial has histidine at 173 position, explaining the H173 specificity of these vaccine induced CH58 and CH59 antibodies.

Neither the 40007 T/F sera nor the RV144 vaccine sera or the CH58 and CH59 mAbs, exhibited strong virus neutralizing activity, but the latter exhibited strong ADCC activity which correlated with reduced infection in immune correlate analysis (31-33). Similarly, in our current study, the Y173 escape mutant which also consisted of an additional DSV tripeptide deletion in the adjacent variable V2 loop showed significant resistance to ADCC mediated killing when

613 compared to the H173 epitope. The DSV deletion alone otherwise had no effect on the binding of
614 the C β-strand epitope to antibodies in ELISA or SPR binding assays, yet it was co-selected along
615 with the Y173 mutation. This leads to a compelling argument that the strong immune pressure
616 exerted by the host might be due to the ADCC activity of the elicited antibody responses, which
617 is consistent with previous reports suggesting the importance of ADCC responses in HIV-1
618 infected individuals particularly the elite controller, and protection imparted by vaccines in non-
619 human primates (44-48).

620 Intriguingly, a second wave of antibodies with greater breadth emerged following the H173
621 to Y173 genetic shift and disappearance of the H173-specific first wave antibodies. Unlike the
622 latter, the second wave antibodies bound equally well to both H173 and Y173 C β strand epitopes,
623 as analyzed by multiple assays including the sensitive SPR assay. Notably, the H173 residue is at
624 the center of the C β strand that contains hydrophilic and solvent-exposed residues flanking on one
625 side and hydrophobic and buried residues on the other side. A H173Y mutation could therefore
626 have significant structural consequences. Previous X-ray structures showed that the H173 peptide
627 epitope assumes an α-helix when bound to CH58 and CH59 antibodies as opposed to a β-strand in
628 the envelope proteins with tyrosine at residue 173. Furthermore, prevalence of H or Y residues at
629 this position among numerous Env sequences in the database argues for functional importance of
630 this site in HIV-1 evolution. Thus, it appears that the H173Y genetic shift in T/F07 HIV-1 virus
631 may have caused a significant structural/conformational change in the C β strand epitope, which
632 led to induction of the second wave antibodies with distinct specificity, with greater breadth than
633 the first wave antibodies.

634 Remarkably, the differential immunological response observed in the human immune
635 system was basically recapitulated in a mouse model. First, the gp16-V1V2 nanoscaffolds elicited

antibody responses in mice that recognized the respective V2 epitopes in the context of gp120 or gp140 envelope protein structures. This is significant not only because antibodies induced by most other scaffolded V2 domains do not bind well to the envelope proteins but also that it indicates that the gp16-scaffolded V2 epitopes are displayed in a native-like conformation. Second and more important, the H173 and Y173 variants induced antibodies with distinct specificities, similar to that observed in the human system. While the H173-induced antibodies strongly reacted with the autologous peptide but not with the Y173 peptide, the Y173-induced antibodies exhibited broad reactivity to both the peptides. This differential response was also observed in their reactivity towards diverse HIV-1 Env proteins. While the H173 antibodies were sensitive to mismatches or sequence diversity in the C β-strand, the Y173 antibodies showed cross-reactivity to a variety of Env proteins. Third, the H173 antibodies but not the Y173 antibodies competed with the RV144 CH58 and CH59 mAbs for binding to the C β strand epitope.

The picture that emerges from the above observations, combined with the previously reported studies, is that the C β-strand region is conformationally dynamic and probably undergoes a structural transition when H 173 changes to Y, from a β-strand conformation as part of conserved Greek motif to an α-helix/helix-coil conformation (28, 34, 35). Our MD simulations and *ab initio* structural modeling of H173 and Y173 V1V2 domains further suggest that the H173 region is rigid and assumes a helical/coil conformation whereas the Y173 domain is more dynamic, preferring to be β-stranded and part of a β-sheet core. Thus, it is imperative to include both these structural forms in any vaccine design for increasing the breadth of antibody responses. In fact, a pentavalent vaccine containing five Env proteins including both the H173 and Y173 variants gave better protection in non-human primates when compared to single immunogen (33).

Given the above considerations, we rationalized that a vaccine to be highly effective, it should not only contain the H173 and Y173 variants, but it should also include additional most commonly found mutations that could fine-tune the conformation, some of which might have arisen as escape mutants against host immune pressures in past infections. Furthermore, presenting these variants as gp16 scaffolds would be ideal because, not only does this eliminate nonspecific distraction to other nonessential epitope sites of the Env but also that the gp16-scaffolded V1V2 domains, as discussed above, elicit antibodies that recognize the C β-strand epitopes in a native context. Therefore, a combinatorial vaccine design was developed by including both the H173 and Y173 variants, each in addition carrying combinations of the most commonly found substitutions in the C β-strand as informed by sequence analyses. These in some respects mimic the natural V2 conformations humans are exposed at the site of entry. Assessment of the antibody responses in the mouse model showed that, indeed, these enhanced the breadth significantly in all the libraries containing cocktails of variants when compared to the respective single immunogen controls. However, the breadth is greater in the context of Y173 when compared to H173, particular when Y173 was combined with a partial deletion of V1 loop. The greater breadth and cross-reactive responses resulted from the Y173-ΔV1 library was likely due to reducing the immunodominance of V1 epitopes as well as generation of breadth favoring conformational variants generated by Y173 switch. Consistent with these data is the recent report that responses directed to V1 loop interferes with binding of protective V2-directed antibodies to Env and promotes virus acquisition in SIV vaccinated macaques (25). Furthermore, we found that the H173 combinatorial library groups induced antibodies similar to CH58 mAb as shown by the CH58 blocking data. Induction of such a response despite additional mutations in the C β-strand argues for a dominant role the H173 residue plays to assume helical conformation as is also recognized by CH58 antibody (35).

These data lead to a model (Fig 9) implicating the conformational dynamism in the V2 region responsible for virus escape or breakthrough infections. This could potentially be curtailed by combinatorial vaccine designs such as the Y173-ΔV1 library of immunogens as future vaccine candidates to induce antibody responses with substantial increase in breadth and cross reactivity to enhance vaccine efficacy and minimize virus escape.

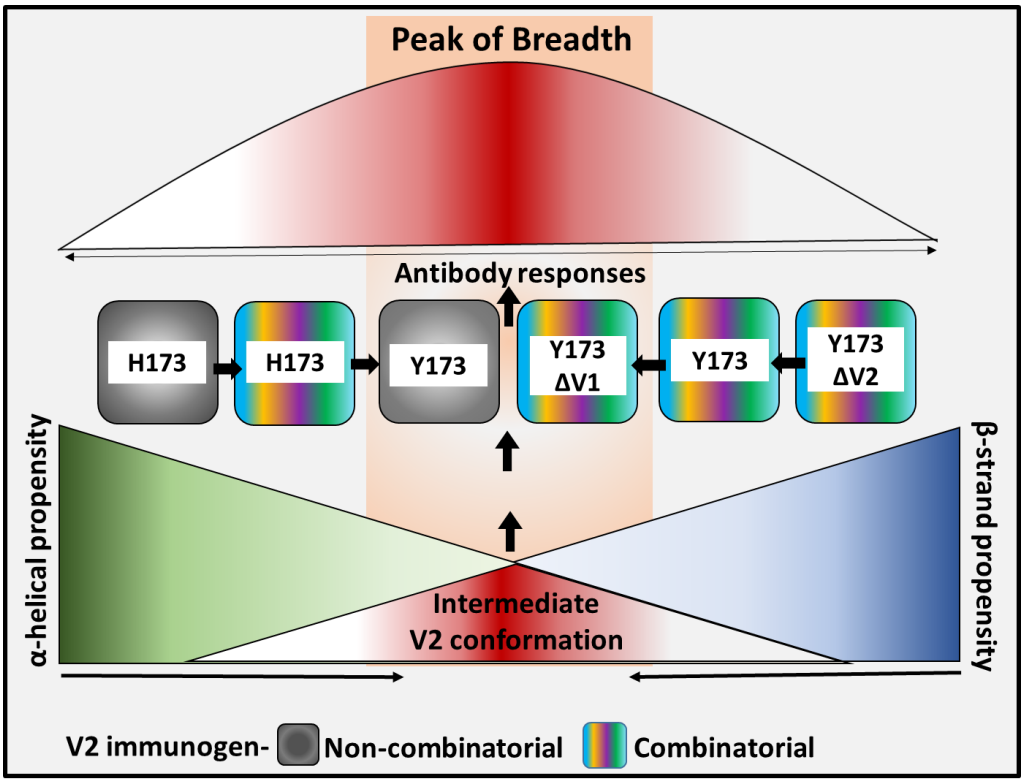


Fig 9: Model of the study. The viral escape observed owing to H173Y mutation during natural infection, and induction of distinct antibody responses by V2 immunogens: H173, Y173 and H173/Y173-based combinatorial libraries, implies that residue 173 has a prominent role in governing the dynamics of the V2 region. The current study shows that V2 immunogens can be categorized into 3 conformational variants based on their propensity to assume: helical/coil conformation, β-strand conformation and possibly a more dynamic intermediate conformation switching between beta/helix conformations. The antibody responses of these V2 variants represent a specific pattern of reactivity recognizing H173 or Y173 or both antigens. Each of the combinatorial and non-combinatorial V2 immunogens studied here are aligned with respect to their plausible conformation and generated extent of breadth, in the above schematic diagram.

Materials and Methods

Ethics statement

All mice were maintained in the pathogen-free animal facility at the Catholic University of America, Washington, D.C. All animal protocols conducted for the current study were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at the Catholic University of America.

Viral Load Analysis

Viral loads were determined in patient 40007 longitudinal samples from wk 0 up to wk 144 as described previously (11, 49). Briefly, viral RNA was isolated from cell-free plasma using the QIAamp viral RNA isolation kit (Qiagen) and quantitated. Quantitative reverse transcription-PCR was conducted in a two-step process. First, RNA was reverse transcribed followed by treatment with RNase H (Stratagene) for 20 min at 37°C and then cDNA was quantified using specific amplification primers, dyes and probes. All reactions were carried out on a 7300 ABI real-time PCR system with TaqGold polymerase (Applied Biosystems) according to the manufacturer's protocols.

Single genome amplification (SGA)-derived envelope sequencing

SGA sequencing was performed at viral sequencing core in Walter Reed Army Institute of Research (WRAIR) and conducted as described previously (11, 50). Briefly, viral RNA was extracted from the plasma of the infected RV217 participants using the QIAamp Viral RNA Mini Kit (QIAGEN, Valencia, CA, USA) and complementary DNA (cDNA) was synthesized using the SuperScript III RT kit (Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's instructions. cDNA was amplified as a full genome or 2 half genomes overlapping

by 1.5 kb as previously described using SGA strategy, which was then end-point diluted in 96-well plate, such that to yield less than 30% amplification product. Env specific primers were used to amplify *env* gene from the HIV genome.

Phylogenetic analysis

Multiple sequence alignment and construction of phylogenetic trees were done using CLC Main Workbench (ver. 7.6.1) software. SGA derived *env* sequences obtained from various time points (1-, 4-, 24-wk post-infection) were aligned with the respective T/F virus sequence using following parameters; Gap Open Cost = 10.0; Gap Extension Cost = 1.0; and Alignment Sensitivity = Very Accurate. Phylogenetic trees were constructed using Neighbor-Joining method, Jukes Cantor protein distance measure, and 100 bootstrap replicates. For mutational frequency analysis, the total number of mutations at each residue of the Env sequence was determined. Numbering of each residue is consistent with the HXB2 strain sequence used as reference. For consensus logo construction, Env sequences of diverse subtypes were fetched from HIV sequence database, aligned using CLC Main Workbench. The logo was generated using an online tool, WebLogo (https://weblogo.berkeley.edu/logo.cgi).

Structural Modeling

3D model of a gp140-T/F07 was generated using homology modeling server, SWISS-MODEL. BG505 gp140 (PDB ID: 4TVP) trimer was used as a template. Mutations were mapped on the modeled trimers using PyMol (ver. 1.74) molecular visualization software (51). Ab initio structural modeling of T/F07- and Y173-V1V2 domains was conducted using QUARK online tool.

743 **Molecular dynamics (MD) simulation**

744 First, T/F07-H173 and -Y173 trimer models were generated by MODELLER 9v7 using T/F100
745 Env cryo-EM structure (PDB ID: 6NQD) as template for modeling. After predicting trimer
746 structures, V1V2 domains were extracted and used for MD simulation. GROMACS 5.1.2 was used
747 to run four MD simulations: H173/Strand, Y173/Strand, H173/Helix, and Y173/Helix as initial
748 models. AMBER99SB-ILDN force field was employed for protein. Dodecahderon periodic box
749 was solvated by TIP3P water molecules. The size of solvation box was set to 1.0 nm from the
750 V1/V2 domain. Sodium ions were added to neutralize the system. After solvation, steepest decent
751 minimization with 50000 steps was applied. Then, the systems were equilibrated at 300 K during
752 100 ps (NVT equilibrium) and at 1.0 bar during 100 ps (NPT equilibrium). After equilibration,
753 100 ns MD trajectory was produced for each system. The equilibration and MD production were
754 done with 2 fs time step, applying LINCS algorithm.

755 **Plasmids construction**

756 gp16 V1V2 scaffolds and combinatorial libraries

757 The gp140-TF07 *env* sequence (spanning gp120 and the gp41 ectodomain up to amino acid 664)
758 was codon optimized and synthesized using GeneArt Strings gene synthesis (Life Technologies).
759 This synthetic fragment was used as template to amplify T/F07-V1V2 sequence corresponding to
760 the residues 117-206 of the Env. The V1V2 sequence was cloned into pCDNA3.1(–) mammalian
761 expression vector engineered to harbor codon optimized bacteriophage T4 terminase, gp16 with
762 an N-terminal Gaussia luciferase (GLuc) signal peptide for secretion of recombinant protein into
763 the media, and a Twin Strep-tag II sequence (WSHPQFEK) for affinity purification at the C-
764 terminus. The amplified V1V2 fragment was cloned downstream to gp16 followed by Twin-strep

765 tags. After the construction gp16-T/F07 V1V2 clone, point mutation and/or deletion was
766 introduced using site-directed mutagenesis to construct gp16-H173Y, -ΔDSV and -H173Y+ΔDSV
767 V1V2 scaffolds. For the construction gp16-V1V2 combinatorial libraries, these scaffolds were
768 used as templates. Controlled mutagenesis was performed using randomized primer pool (IDT)
769 with desired mutations or substitutions at the selected positions in the C-strand to generate V1V2
770 mutant libraries. Amplified V1V2 library fragments were then cloned into the same pcDNA 3.1 (-
771) vector used previously but with an exception that this vector was further engineered to harbor
772 HRV3c protease cleavage site positioned before the Twin-strep tags for tag removal. For the
773 construction of gp16-H173Y_ΔV1 combinatorial library, H173Y template with a stretch of V1
774 loop residues (SNITVERNITIANDTYD) deleted and replaced by an optimized short-linker
775 (AGGAS) was used.

776 gp140 and gp120 clones for heterologous antigen library

777 Codon-optimized gp140 *env* sequences (spanning gp120 and the gp41 ectodomain up to amino
778 acid 664) for T/F07, T/F61, T/F94, T/F100, BG505, JRFL and SF162 harboring trimer stabilizing
779 SOSIP mutations (52) and six arginine (R6) furin cleavage site (replacing native REKR cleavage
780 site at the junction of gp120 and gp41) were synthesized (53). The gp140 genes were then cloned
781 into pCDNA 3.1(–) vector that was engineered to harbor an N-terminal Cluster of Differentiation
782 5 antigen (CD5) signal peptide for secretion of the recombinant proteins into the media and 8X-
783 Histidine tag for affinity purification. Additional mutations/deletions were introduced in gp140-
784 T/F07 using site directed mutagenesis kit (NEB). The clones' sequences were verified through
785 sequencing (Retrogen, Inc.). The gp120 expression vectors were obtained from NIH AIDS
786 Reagent Program (Table 1). The furin-expressing plasmid, Furin:FLAG/pGEM7Zf(+), was

obtained from Dr. Gary Thomas (Vollum Institute, Portland, Oregon). The furin fragment from this plasmid was subcloned into pcDNA3.1(–) (Life Technologies).

Cells and Media

HEK293S GnTI- (ATCC CRL-3022) suspension cells used for expression of HIV Env proteins were maintained in FreeStyle 293 expression medium (Life Technologies), supplemented with 1% heat-inactivated fetal bovine serum (FBS, Quality Biologicals). All cells were grown in suspension in a Multitron Pro orbital shaker (Infors HT) incubator at 37°C in 8% CO2, 80% humidified atmosphere.

Transfection

Plasmid DNAs for transfection were purified using Plasmid Midi kit (Qiagen) as per manufacturer's instructions. Transfections were carried out as described previously (54). Briefly, GnTi cells were grown to 1 x 10^6/ml cell density for transfection. Prior to transfection, cells were centrifuged at 100 rpm for 5 minutes followed by full replacement of media with the fresh Freestyle293 media lacking FBS. The final cell density was adjusted to 2 x 10^6 /ml in half or 50% of the final volume of transfection. The cells were then placed in the shaker incubator for 1 hour at 37°C in 8% CO2, 80% humidified atmosphere. After incubation, DNA (1 µg/ml final transfection volume) was added followed by addition of linear polyethylenimine (PEI25k, Polyscience, Inc.) (1mg/ml) at a 3:1 ratio (PEI:DNA) to the cell suspension. For gp140 expression, cells were co-transfected with furin plasmid DNA to produce cleaved gp120 and gp41 subunits that then associate non-covalently to yield native Env proteins. After 12 h of transfection, HyClone SFM4HEK293 medium (GE Healthcare) supplemented with 1% FBS (v/v) and protein expression enhancing sodium butyrate (55) solution (SIGMA-ALDRICH) to a final concentration of 2 nM

were added to the cells to make up to the final volume of transfection. After 5 days of transfection, the supernatant was harvested by centrifuging the cells, and filtered using a 0.2 µm filter (Corning, Inc.).

Protein purification

Secreted twin strep-tagged gp16-V1V2 proteins in the harvested and filtered supernatant were supplemented with BioLock biotin blocking solution (IBA Lifesciences) at 5 µl/ml to mask the biotin present in the supernatant. After 30 min of incubation, the supernatant was loaded onto a 1 ml StrepTactin column (Qiagen) at a flow rate of 0.7 ml/min in the ÄKTA prime-plus liquid chromatography system (GE Healthcare). Non-specifically bound proteins were washed off by passing at least 20 column volumes of the wash buffer (300 mM NaCl, 50 mM Tris-HCl, pH 8) or until the absorbance reached the baseline level. Bound gp16-V1V2 proteins were eluted with StrepTactin elution buffer (5 mM d-Desthiobiotin, 300 mM NaCl, 50 mM Tris-HCl, pH 8) at a flow rate of 1 ml/min. Eluted peak fractions were buffer exchanged into 100 mM NaCl, 50 mM Tris-HCl, pH 8 buffer. Protein fractions were stored with 10% glycerol at -80°C until use for antigenicity and immunogenicity studies. GnTi expressed His-tagged gp140s and gp120s were purified from the harvested and clarified supernatant using Ni-NTA agarose beads (Qiagen) following manufacturer's instructions.

Strep-tag removal from gp16-V1V2 immunogens

For the second mice immunization study, the Twin-strep tags were cleaved off the immunogens using HRV3c protease. The recombinant proteins eluted after StrepTactin affinity chromatography, were buffer exchanged with 1X HRV3C protease buffer to remove the desthiobiotin present in the elution buffer. 1 µL of protease was added per 20 µg of the purified

protein (1:20) and incubated at 4°C for 16 hours. Digested protein was passed twice through StrepTactin spin column (IBA). Uncleaved strep-tagged protein bound to the column while the flow-through containing desirable cleaved fraction was collected. The cleaved protein was then loaded onto the size-exclusion chromatography column for fractionation using 100 mM NaCl, 50 mM Tris-HCl, pH 8 buffer. Owing to a large difference in the native dodecameric gp16-V1V2 (~336 kD), HRV3C protease (47.8 kD) was separated from the final immunogens. The gp16-V1V2 fractions were pooled, concentrated and stored at -80°C until use.

Biochemical analyses and quantitation of proteins

Purified proteins were run on SDS-PAGE to quantify and assess for non-specific protein contamination. SDS-PAGE analyses were performed using 4-20% gradient Tris-glycine gels (Life Technologies) or home-made 12% gels in the presence of DTT (reducing conditions) or absence of DTT (non-reducing). All gels were stained with Coomassie blue R-250 solution. Band intensities were measured using Bio-Rad Gel Doc XR+ System and Image Lab software. BSA standards were used to generate a standard curve for quantification. Deglycosylation was also performed to sharpen the bands for accurate quantitation as HIV Env proteins are glycosylated and hence appear fuzzy on gels. For deglycosylation, 1 µl (500 Units) of PNGase F (New England BioLabs, Inc.) was used to deglycosylate 10 µg of the protein in the presence of 5 mM DTT and mild detergents by incubating at room temperature for 1 hour according to manufacturer's recommendations.

Western Blotting

Proteins separated by SDS-PAGE were transferred to a PVDF membrane using the Trans-Blot® Turbo RTA Mini PVDF Transfer Kit (Bio-Rad Laboratories, Inc.). Membranes after activating

with methanol were blocked with bovine serum albumin (Amresco, LLC). For Strep-Tag II detection, HRP-conjugated StrepMAB-Classical MAb was used at 1:3000 dilution in PBS. Purified mAbs, CH58 and CH59 were used as primary antibodies at 1:5000 dilution in PBS and rabbit anti-human Ab HRP conjugate (Santa Cruz Biotechnology) was used as secondary antibodies at 1:10,000 dilution in PBS. Signal from HRP-conjugated antibodies was detected using Clarity™ Western ECL Blotting substrate (Bio-Rad Laboratories, Inc.). Band intensities were measured using Bio-Rad Gel Doc XR+ System and Image Lab software.

Enzyme Linked Immunosorbent Assay (ELISA)

StrepTactin ELISA

StrepTactin ELISA was performed to determine CH58 and CH59 binding to gp16-V1V2 proteins. These specialized plates are pre-coated with StrepTactin to capture strep-tagged antigens. Since the antigen does not directly bind to the plate surface it is maintained in native conformation that improves antibody recognition. To perform this assay, StrepTactin coated microplates (IBA Life Sciences) were coated with 1 µg/ml Strep-tagged proteins in a volume of 100 µl per well of buffer (25 mM Tris-HCl, pH 7.6, 2 mM EDTA, and 140 mM NaCl) and incubated for 2 h at 4^0C. Following three washes with PBST (0.05% Tween-20 in 1X PBS), 100 µl of serially diluted Abs (10-0.001 µg/ml) in PBS were added to the wells and the plates were incubated for 1 h at 37°C. After three washes with PBST, the plates were incubated with 100 µl of rabbit anti-human Ab HRP conjugate at 1:3,000 dilution in PBS for 30 min at 37°C. The plates were then washed three times with PBST and the peroxidase substrate was added to develop the color reaction (TMB Microwell Peroxidase Substrate system, KPL). The reaction was terminated by adding 100 µl of BlueSTOP solution (KPL) and OD650 was recorded using VersaMax ELISA Microplate Reader (Molecular Devices).

Conventional ELISA

96-well Nunc ELISA plates were coated with 100 ng/well antigen diluted in 1X PBS to a concentration of 1µg/ml, for overnight at 4°C. After 12 hours, the plates were washed thrice with 1X PBST (1X PBS + 0.05% Tween), followed by blocking with 5% BSA in 1X PBS for 1h at room temperature (RT). After incubation, plates were washed thrice with 1X PBST, followed by addition of 100 µL of primary antibody or serum dilution for 1h at 37°C. After incubation, the plates are washed three times same as before and anti-mouse secondary antibody was added, followed by 30 mins incubation at RT. The remaining procedure is same as described above for Strep-Tactin ELISA.

Antibody blocking assay

Antibody blocking assay was conducted by modifying the ELISA protocol. Briefly, 96-well Nunc ELISA plates were coated with 100 ng/well gp140-T/F07 antigen diluted in 1X PBS to a concentration of 1µg/ml, for overnight at 4°C. After 12 hours, the plates were washed thrice with 1X PBST, followed by blocking with 3% BSA in 1X PBS for 1h at room temperature (RT). Plates were washed thrice with 1X PBST, followed by addition of 100 µL of CH58 or CH59 antibody dilution (1µg/ml) for 1h at 37°C. Wells not preincubated with CH58/CH59 antibodies served as unblocked positive control. The antibodies blocked and unblocked (control) plates were washed three times same as before and incubated with optimized mice sera dilutions (1:2000) from various immunization groups for another 1h at 37°C. After washing three times, the plates were incubated with anti-mouse HRP conjugated secondary antibody for 30 mins at RT. The plates were then washed followed by TMB peroxidase substrate addition (KPL) and read as described above for ELISA.

898 **Surface Plasmon Resonance (SPR) Binding Assay**

899 Longitudinal plasma samples of RV217 participant-40007 were analyzed by SPR for the presence
900 of V2-specific Abs. SPR measurements were made with a Biacore 4000 system (GE Healthcare,
901 Uppsala, Sweden). The assay was conducted as described previously (11). Briefly, plasma samples
902 were heat-inactivated at 56°C for 45 mins followed by centrifugation at 16,000 × g at 4 °C for 20
903 min and the supernatants were used for SPR analyses. Recombinant gp16-T/F07 and other escape
904 mutant proteins were immobilized onto CM5 or CM7 series sensor chips. Plasma samples diluted
905 1:100 in running buffer (10 mM HEPES, 300mM NaCl and 0.005% Tween 20, pH 7.4) were
906 injected onto the immobilized chip. The detection of antigen-antibody complexes captured on the
907 chip surface was then enhanced with a 200 s injection of 30 μg/mL secondary sheep anti-human
908 IgG antibody (Binding Site, Birmingham, United Kingdom). For evaluating the binding of Abs
909 present in immunized mice sera with V2-peptides, the sensor chips were first prepared using a
910 standard amine-coupling method as previously described (11) for coupling of streptavidin. Then,
911 N-linked biotinylated overlapping V2-peptides synthesized by JPT Peptide Technologies GmbH
912 (Berlin, Germany) were captured onto the streptavidin immobilized chips. Sample or buffer was
913 injected at a flow rate of 10 μL/min at 25 °C. The data analysis was performed using Biacore 4000
914 Evaluation Software v4.1 (GE Healthcare, Uppsala Sweden). The reported RU for the IgG specific
915 values are represented as a difference between the average value of a 5 s window taken 60 s after
916 the end of the anti-IgG injection and the average value of a 5 s window taken 10 s before the
917 beginning of the anti-IgG injection. The response units were double subtracted by the RU of the
918 unmodified surface and buffer.

919

920

GranToxiLux (GTL) Antibody dependent cell cytotoxity (ADCC) assay

ADCC assays were performed as described previously (56). Human CD4+ T lymphoblasts, CEM.NKR, were used as target cells to coat recombinant gp120 proteins. The amount of coating gp120 was optimized through competition by binding of the Leu3A (anti-CD4) antibody (clone SK3; Catalog no. 340133; Final dilution 1:5; BD Bioscience, San Jose, CA, USA). Cryopreserved peripheral blood mononuclear cells, PBMCs from a healthy donor, thawed and rested overnight in R10 media, were used as effector cells (source of effector NK cells). The following day, target cells were coated with titrated amount of T/F07 and its V1V2-mutant gp120s for 75 minutes at 37^0C. After incubation, coated target cells were mixed with effector cells in 30:1 ratio in 96-well V-bottom plate, followed by addition on granzymeB (GzB). Finally, 4-folds serially dilutions of heat-inactivated plasma samples of RV217 participant- 40007, from three visits, v0 (pre-infection) for the baseline, and 4-wks (v9) and 24-wks (v14) post-infection or purified antibodies, CH58 and CH65 were added to the respective wells. After 1 hour incubation at 37°C and 5% CO_2, the plate was centrifuged and washed with wash buffer. After washing, the cells were resuspended in wash buffer and the plate was read using the BD LSRII or BD LSRFortessa with the High Throughput Sampler (HTS) with a minimum of 1250 events, to detect the activity of granzyme B (GzB) released by the effector population into target cells. The viable target cells with activated GzB substrate represented the actual population recognized by the effector cells and reported as %GzB activity. The results are reported after background subtraction of the signal acquired from target cells incubated with effector cells in the absence of plasma/antibodies.

Mice immunization

6-wks old female BALB/c mice were received from Jackson's laboratory for immunization experiments. The immunization was initiated after 2 wks of quarantine. The weights of the mice

944 were taken periodically from the start of the quarantine period to assess growth and health of mice
945 throughout the experiment. 20 µg of the antigen complexed with Alhydrogel 2% (Invivogen) as
946 adjuvant was injected intramuscularly per mouse using a 22-23 gauze needle syringe. Three
947 boosters were given after prime/first immunization at an interval of 3 wks and tail bleed was
948 performed to collect sera before each immunization. Mice were also bled before the first
949 immunization to collect pre-immunized (pre-bleed) sera for the negative control. Terminal bleed
950 was performed through cardiac puncture under general anesthesia followed by cervical dislocation
951 to euthanize the animals.

952

953

954 ## Funding

955 This work was supported by the National Institute of Allergy and Infectious Diseases NIH grants
956 AI111538, AI102725, and AI081726 to V.B.R. and by a cooperative agreement (W81XWH-07-
957 2-0067 and W81XWH-11-0174) between the Henry M. Jackson Foundation for the Advancement
958 of Military Medicine, Inc. and the U.S. Department of Defense. DK acknowledges supports by the
959 National Institutes of Health (R01GM133840 and R01GM123055) and the National Science
960 Foundation (CMMI1825941, MCB1925643, and DBI2003635).

961

962

963

964

References

1. Mocroft A, Phillips AN, Gatell J, Ledergerber B, Fisher M, Clumeck N, et al. Normalisation of CD4 counts in patients with HIV-1 infection and maximum virological suppression who are taking combination antiretroviral therapy: an observational cohort study. Lancet. 2007;370(9585):407-13.
2. WHO. HIV treatment and care.
3. UNAIDS. Global HIV & AIDS statistics — 2019 fact sheet 2019 [Available from: https://www.unaids.org/en/resources/fact-sheet.
4. Esparza J. A brief history of the global effort to develop a preventive HIV vaccine. Vaccine. 2013;31(35):3502-18.
5. Esparza J. What Has 30 Years of HIV Vaccine Research Taught Us? Vaccines (Basel). 2013;1(4):513-26.
6. Miedema F. A brief history of HIV vaccine research: stepping back to the drawing board? AIDS. 2008;22(14):1699-703.
7. Gilbert PB, Berger JO, Stablein D, Becker S, Essex M, Hammer SM, et al. Statistical interpretation of the RV144 HIV vaccine efficacy trial in Thailand: a case study for statistical issues in efficacy trials. The Journal of infectious diseases. 2011;203(7):969-75.
8. Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N Engl J Med. 2009;361(23):2209-20.
9. de Souza MS, Ratto-Kim S, Chuenarom W, Schuetz A, Chantakulkij S, Nuntapinit B, et al. The Thai phase III trial (RV144) vaccine regimen induces T cell responses that preferentially target epitopes within the V2 region of HIV-1 envelope. J Immunol. 2012;188(10):5166-76.
10. Haynes BF, Gilbert PB, McElrath MJ, Zolla-Pazner S, Tomaras GD, Alam SM, et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. N Engl J Med. 2012;366(14):1275-86.
11. Trinh HV, Gohain N, Pham PT, Hamlin C, Song H, Sanders-Buell E, et al. Correction: Trinh, H.V., et al. Humoral Response to the HIV-1 Envelope V2 Region in a Thai Early Acute Infection Cohort. Cells 2019, 8, 365. Cells. 2019;8(6).
12. Rolland M, Edlefsen PT, Larsen BB, Tovanabutra S, Sanders-Buell E, Hertz T, et al. Increased HIV-1 vaccine efficacy against viruses with genetic signatures in Env V2. Nature. 2012;490(7420):417-20.
13. Land A, Braakman I. Folding of the human immunodeficiency virus type 1 envelope glycoprotein in the endoplasmic reticulum. Biochimie. 2001;83(8):783-90.
14. Leonard CK, Spellman MW, Riddle L, Harris RJ, Thomas JN, Gregory TJ. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. J Biol Chem. 1990;265(18):10373-82.
15. Pan R, Gorny MK, Zolla-Pazner S, Kong X-P. The V1V2 Region of HIV-1 gp120 Forms a Five-Stranded Beta Barrel. Journal of Virology. 2015;89(15):8003.
16. Hioe CE, Kumar R, Upadhyay C, Jan M, Fox A, Itri V, et al. Modulation of Antibody Responses to the V1V2 and V3 Regions of HIV-1 Envelope by Immune Complex Vaccines. Front Immunol. 2018;9:2441.
17. Rao M, Peachman KK, Kim J, Gao G, Alving CR, Michael NL, et al. HIV-1 variable loop 2 and its importance in HIV-1 infection and vaccine development. Curr HIV Res. 2013;11(5):427-38.
18. Sagar M, Wu X, Lee S, Overbaugh J. Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. J Virol. 2006;80(19):9586-98.
19. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature. 1998;393(6686):648-59.

20. Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, et al. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature. 2014;514(7523):455-61.
21. Klasse PJ. The molecular basis of HIV entry. Cellular Microbiology. 2012;14(8):1183-92.
22. Wilen CB, Tilton JC, Doms RW. Molecular mechanisms of HIV entry. Adv Exp Med Biol. 2012;726:223-42.
23. Arthos J, Cicala C, Martinelli E, Macleod K, Van Ryk D, Wei D, et al. HIV-1 envelope protein binds to and signals through integrin alpha4beta7, the gut mucosal homing receptor for peripheral T cells. Nat Immunol. 2008;9(3):301-9.
24. Arthos J, Cicala C, Nawaz F, Byrareddy SN, Villinger F, Santangelo PJ, et al. The Role of Integrin alpha4beta7 in HIV Pathogenesis and Treatment. Curr HIV/AIDS Rep. 2018;15(2):127-35.
25. Goes LR, Sajani A, Sivro A, Olowojesiku R, Ray JC, Perrone I, et al. The V2 loop of HIV gp120 delivers costimulatory signals to CD4(+) T cells through Integrin alpha4beta7 and promotes cellular activation and infection. Proc Natl Acad Sci U S A. 2020;117(51):32566-73.
26. Peachman KK, Karasavvas N, Chenine AL, McLinden R, Rerks-Ngarm S, Jaranit K, et al. Identification of New Regions in HIV-1 gp120 Variable 2 and 3 Loops that Bind to alpha4beta7 Integrin Receptor. PLoS One. 2015;10(12):e0143895.
27. Lertjuthaporn S, Cicala C, Van Ryk D, Liu M, Yolitz J, Wei D, et al. Select gp120 V2 domain specific antibodies derived from HIV and SIV infection and vaccination inhibit gp120 binding to alpha4beta7. PLoS Pathog. 2018;14(8):e1007278.
28. van Eeden C, Wibmer CK, Scheepers C, Richardson SI, Nonyane M, Lambson B, et al. V2-Directed Vaccine-like Antibodies from HIV-1 Infection Identify an Additional K169-Binding Light Chain Motif with Broad ADCC Activity. Cell Reports. 2018;25(11):3123-35.e6.
29. Sun S, Gao S, Kondabagil K, Xiang Y, Rossmann MG, Rao VB. Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages. Proc Natl Acad Sci U S A. 2012;109(3):817-22.
30. Ananthaswamy N, Fang Q, AlSalmi W, Jain S, Chen Z, Klose T, et al. A sequestered fusion peptide in the structure of an HIV-1 transmitted founder envelope trimer. Nat Commun. 2019;10(1):873.
31. Mayr LM, Decoville T, Schmidt S, Laumond G, Klingler J, Ducloy C, et al. Non-neutralizing Antibodies Targeting the V1V2 Domain of HIV Exhibit Strong Antibody-Dependent Cell-mediated Cytotoxic Activity. Scientific reports. 2017;7(1):12655-.
32. Bonsignori M, Pollara J, Moody MA, Alpert MD, Chen X, Hwang KK, et al. Antibody-dependent cellular cytotoxicity-mediating antibodies from an HIV-1 vaccine efficacy trial target multiple epitopes and preferentially use the VH1 gene family. J Virol. 2012;86(21):11521-32.
33. Bradley T, Pollara J, Santra S, Vandergrift N, Pittala S, Bailey-Kellogg C, et al. Pentavalent HIV-1 vaccine protects against simian-human immunodeficiency virus challenge. Nature Communications. 2017;8(1):15711.
34. Wibmer CK, Richardson SI, Yolitz J, Cicala C, Arthos J, Moore PL, et al. Common helical V1V2 conformations of HIV-1 Envelope expose the alpha4beta7 binding site on intact virions. Nat Commun. 2018;9(1):4489.
35. Liao HX, Bonsignori M, Alam SM, McLellan JS, Tomaras GD, Moody MA, et al. Vaccine induction of antibodies against a structurally heterogeneous site of immune pressure within HIV-1 envelope protein variable regions 1 and 2. Immunity. 2013;38(1):176-86.
36. Hessell AJ, Powell R, Jiang X, Luo C, Weiss S, Dussupt V, et al. Multimeric Epitope-Scaffold HIV Vaccines Target V1V2 and Differentially Tune Polyfunctional Antibody Responses. Cell Rep. 2019;28(4):877-95 e6.
37. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol. 1993;234(3):779-815.

38. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. Journal of computational chemistry. 2005;26(16):1701-18.
39. Xu D, Zhang Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. Proteins. 2012;80(7):1715-35.
40. Saunders CJ, McCaffrey RA, Zharkikh I, Kraft Z, Malenbaum SE, Burke B, et al. The V1, V2, and V3 regions of the human immunodeficiency virus type 1 envelope differentially affect the viral phenotype in an isolate-dependent manner. J Virol. 2005;79(14):9069-80.
41. Sanders RW, Schiffner L, Master A, Kajumo F, Guo Y, Dragic T, et al. Variable-loop-deleted variants of the human immunodeficiency virus type 1 envelope glycoprotein can be stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits. J Virol. 2000;74(11):5091-100.
42. Silva de Castro I, Gorini G, Mason R, Gorman J, Bissa M, Rahman MA, et al. Anti-V2 antibodies virus vulnerability revealed by envelope V1 deletion in HIV vaccine candidates. iScience. 2021;24(2):102047.
43. Kim JH, Excler JL, Michael NL. Lessons from the RV144 Thai phase III HIV-1 vaccine trial and the search for correlates of protection. Annu Rev Med. 2015;66:423-37.
44. Excler J-L, Ake J, Robb ML, Kim JH, Plotkin SA. Nonneutralizing functional antibodies: a new "old" paradigm for HIV vaccines. Clin Vaccine Immunol. 2014;21(8):1023-36.
45. Bruel T, Guivel-Benhassine F, Lorin V, Lortat-Jacob H, Baleux F, Bourdic K, et al. Lack of ADCC Breadth of Human Nonneutralizing Anti-HIV-1 Antibodies. J Virol. 2017;91(8).
46. Dupuy FP, Kant S, Barbe A, Routy JP, Bruneau J, Lebouche B, et al. Antibody-Dependent Cellular Cytotoxicity-Competent Antibodies against HIV-1-Infected Cells in Plasma from HIV-Infected Subjects. mBio. 2019;10(6).
47. Mabuka J, Nduati R, Odem-Davis K, Peterson D, Overbaugh J. HIV-specific antibodies capable of ADCC are common in breastmilk and are associated with reduced risk of transmission in women with high viral loads. PLoS Pathog. 2012;8(6):e1002739.
48. Isitman G, Lisovsky I, Tremblay-McLean A, Kovacs C, Harris M, Routy JP, et al. Antibody-Dependent Cellular Cytotoxicity Activity of Effector Cells from HIV-Infected Elite and Viral Controllers. AIDS Res Hum Retroviruses. 2016;32(10-11):1079-88.
49. Robb ML, Eller LA, Kibuuka H, Rono K, Maganga L, Nitayaphan S, et al. Prospective Study of Acute HIV-1 Infection in Adults in East Africa and Thailand. New England Journal of Medicine. 2016.
50. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci U S A. 2008;105(21):7552-7.
51. Schrödinger LLC. The PyMOL Molecular Graphics System, Version 1.74. 2010.
52. Sanders RW, Derking R, Cupo A, Julien J-P, Yasmeen A, de Val N, et al. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. PLoS pathogens. 2013;9(9):e1003618-e.
53. Binley JM, Sanders RW, Master A, Cayanan CS, Wiley CL, Schiffner L, et al. Enhancing the proteolytic maturation of human immunodeficiency virus type 1 envelope glycoproteins. J Virol. 2002;76(6):2606-16.
54. AlSalmi W, Mahalingam M, Ananthaswamy N, Hamlin C, Flores D, Gao G, et al. A New Approach to Produce HIV-1 Envelope Trimers: BOTH CLEAVAGE AND PROPER GLYCOSYLATION ARE ESSENTIAL TO GENERATE AUTHENTIC TRIMERS. J Biol Chem. 2015;290(32):19780-95.
55. Reeves PJ, Kim JM, Khorana HG. Structure and function in rhodopsin: a tetracycline-inducible system in stable mammalian cell lines for high-level expression of opsin mutants. Proc Natl Acad Sci U S A. 2002;99(21):13413-8.
56. Pollara J, Hart L, Brewer F, Pickeral J, Packard BZ, Hoxie JA, et al. High-throughput quantitative analysis of HIV-1 and SIV-specific ADCC-mediating antibody responses. Cytometry A. 2011;79(8):603-12.

Supplementary information

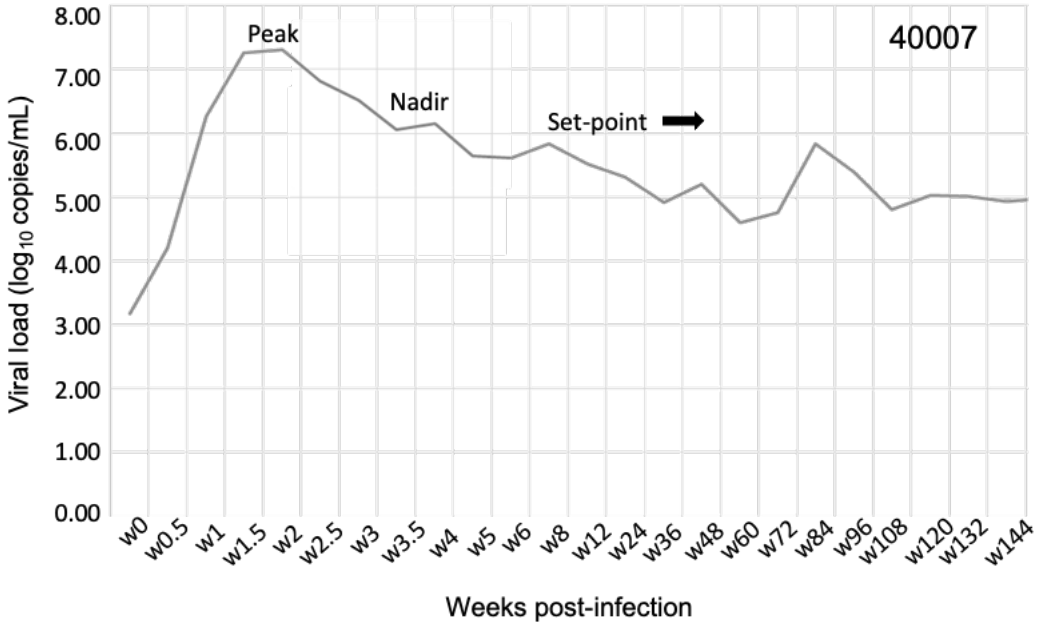


Fig S1. Longitudinal viral load analysis over a course of infection in 40007 participant. Viral load values (log_{10}) are plotted on y-axis, against the number of weeks since the first HIV-positive reaction with nucleic acid test on the x-axis. A typical pattern of curve with a peak, nadir and set point in viral load was observed in this patient representing early captured infection. The set point is the viral load of a person infected with HIV, which stabilizes after a period of acute HIV infection.

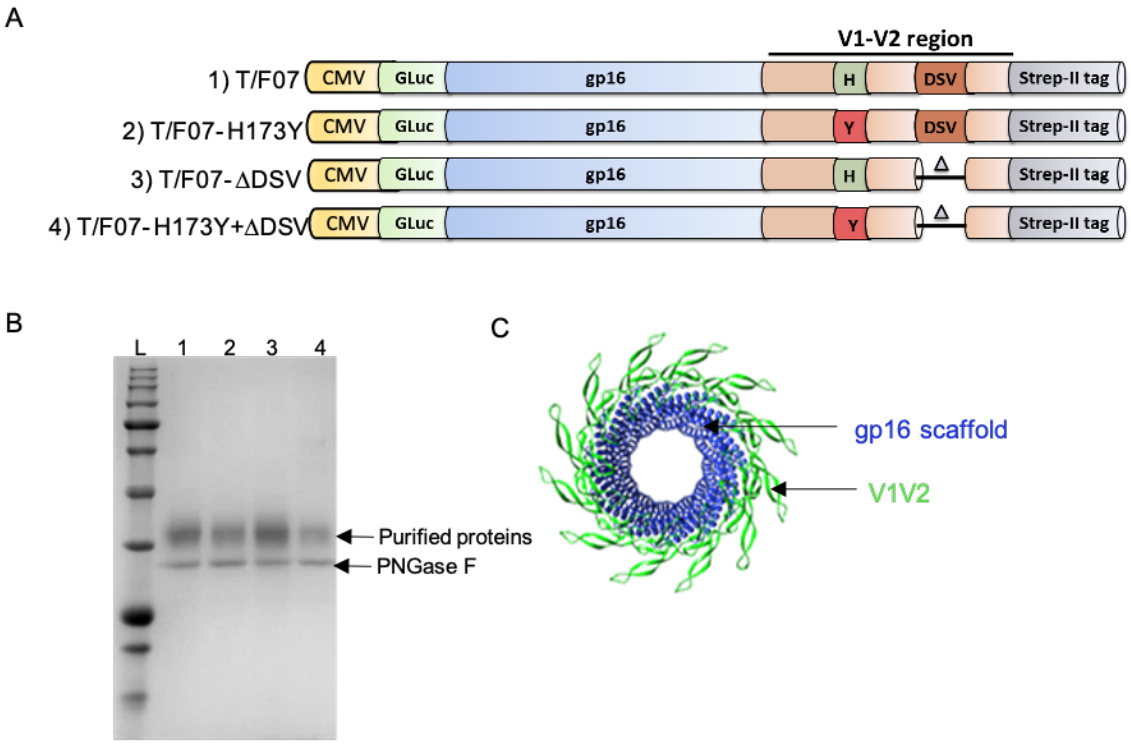


Fig S2. gp16-V1V2 construct design and purification (A) gp16-V1V2 constructs corresponding to the T/F07 virus sequence and V1V2-specific mutations accumulated until 24-weeks of infection in participant 40007. The V1V2 sequence (light orange) amplified from the T/F07 Env (gp160) sequence was fused in frame to the C-terminus of gp16 scaffold (blue). Each construct was cloned under CMV promoter and contained an N-terminal signal peptide (GLuc) for secretion of these recombinant proteins into the medium, and a C-terminal Twin Strep-tag II (gray) for affinity purification. 24-week V2-mutations were introduced in the C β-strand (H173Y, red) and the hypervariable V2 loop (3 residue deletion-ΔDSV) indicated by a small triangle in the parental T/F07 construct to generate single and double mutants. (B) SDS-PAGE gel profile of gp16-V1V2 variant scaffolds of T/F07 expressed in HEK293S (GnTi) cells and purified through StrepTactin-based affinity chromatography. These recombinant glycoproteins were deglycosylated by PNGase F (band labeled on the gel) to obtain sharper bands on the gel for the purpose of quantification. (C) Dodecameric model of gp16-V1V2 showing gp16-scaffold in blue and fused V1V2 domain in bright green.

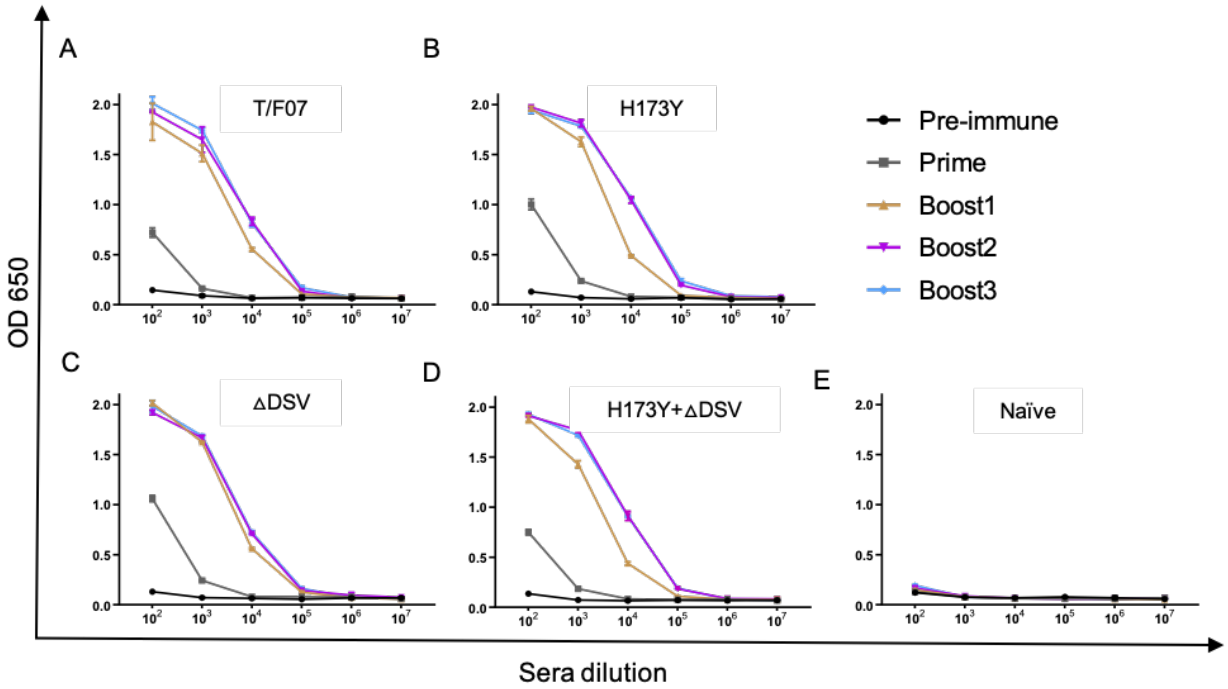


Fig S3. V1V2-specific antibody titers. (A-E) Comparative titers of V1V2-antibodies ranging from first (prime) to the last immunization (Boost 3) in mice groups immunized with gp16- T/F07 (A) H173Y (B) ΔDSV (C) H173Y+ΔDSV (D), and Naïve group (no antigen control) (E) are shown. Respective pre-immune sera (collected before first immunization) from each group were also kept as negative control. The antibody titers are determined through ELISA. Respective purified recombinant soluble gp140-T/F07, -H173Y, -ΔDSV and -H173Y+ΔDSV Env glycoproteins were used as coating antigens (1μg/ml) matching the V1V2 region. Triplicate absorbance (OD 650 nm) readings are used to generate binding curves for each sample. A color-coded key is provided on the top-right corner for each immunization specific curve.

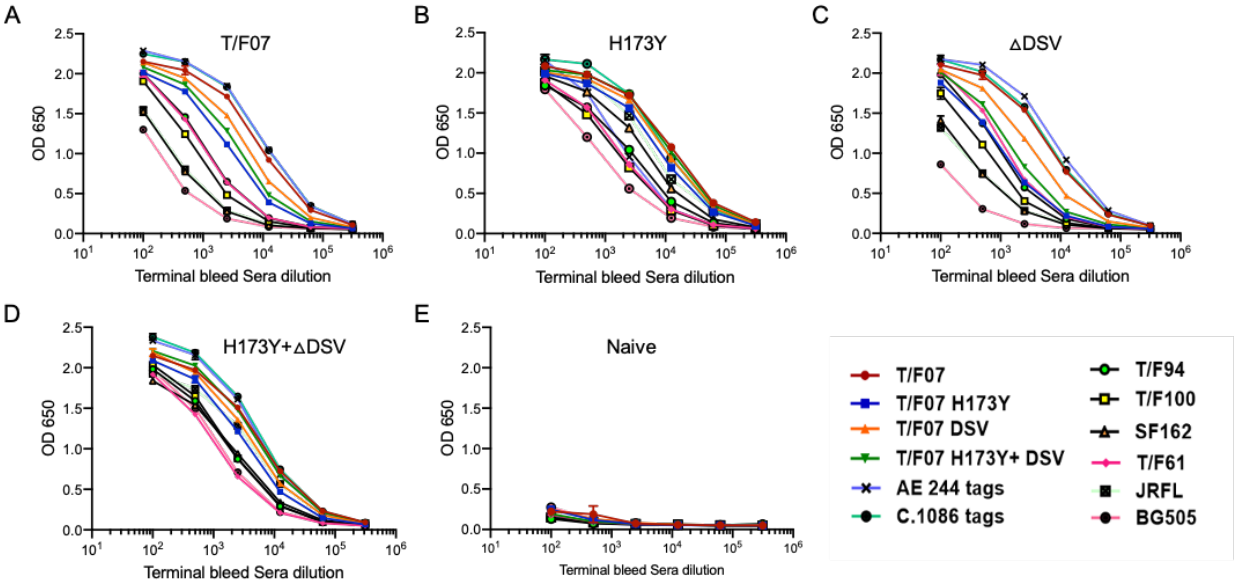


Fig S4. (A-E) Breadth was evaluated by assessing binding with a set of purified autologous and heterologous Env proteins (soluble gp140s and V1V2 tags). Respective antigen binding curves are plotted against the serial dilutions of the terminal bleed sera from each group immunized with T/F07 (A), H173Y (B), ΔDSV (C), H173Y+ΔDSV (D), and PBS (Naïve, negative control) (E). Binding curves are color-coded with respect to antigen coated as shown in the legend at the bottom of the graph. The binding was determined through ELISA and triplicate absorbance (OD 650 nm) readings were used to generate binding curves.

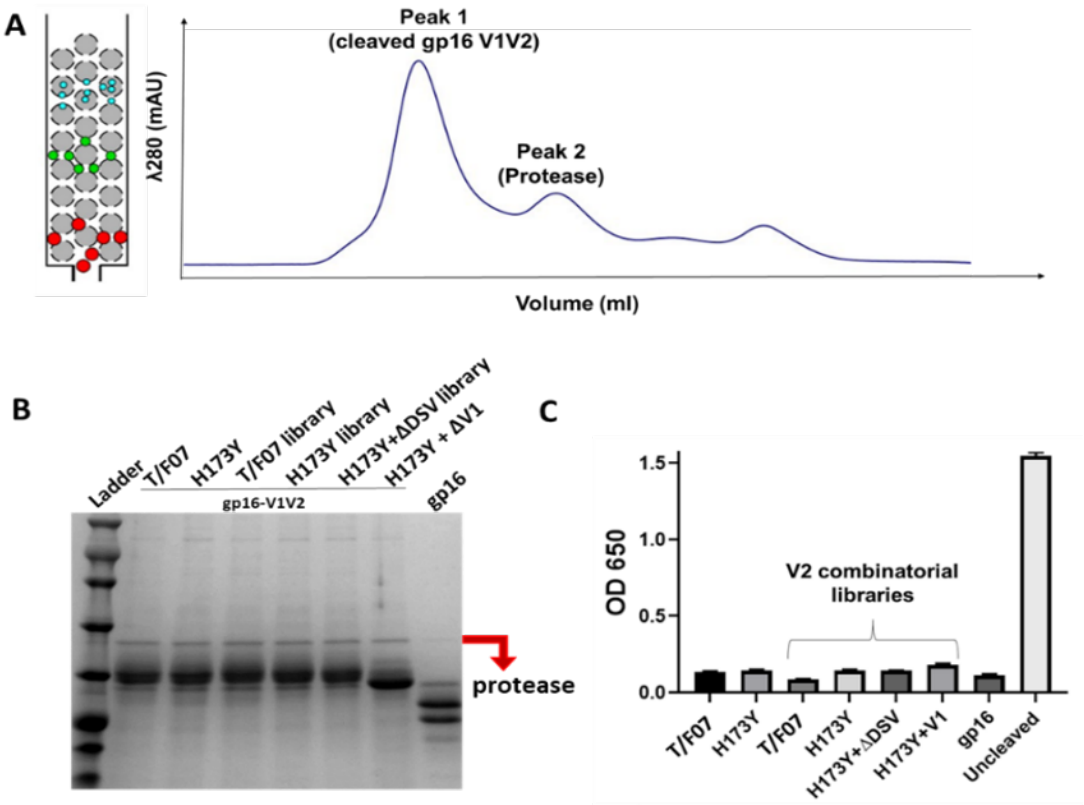


Fig S5. Purification of Twin-strep-tag cleaved V1V2 immunogens. (A) Size exclusion chromatography (SEC) fractionation profile of cleaved gp16-V1V2 proteins (peak 1) depicting separation of HRV3C protease (peak 2). Elution volume is plotted on the x-axis while the y-axis shows UV absorbance of the fractions. (B) Reducing SDS-PAGE profile of purified and concentrated gp16-V1V2 immunogens. Presence of very small fraction of the protease (marked by a red arrow) was detected. Appearance of smeary pattern or doublet bands are due to the presence of glycoforms. (C) ELISA based detection of α-strep tag response in the final preparation of twin-strep tag cleaved immunogens for quality control. Negligible amount of α-strep tag response was observed for all the immunogens. Uncleaved (with twin-strep tag) gp16-V1V2 was used a positive control.

1179

Table 1. List of reagents ordered from NIH Reagent Program

Catalog #	Description
4961	HIV-1 BaL gp120 Recombinant Protein
7749	HIV-1 CN54 gp120 Recombinant Protein
10080	HIV-1 96ZM651 gp120 Recombinant Protein
11556	HIV-1 JR-CSF Fc-gp120 Recombinant Protein
11784	HIV-1 IIIB gp120 Recombinant Protein
12063	HIV-1 UG037 gp140 Recombinant Protein
12064	HIV-1 CN54 gp140 Recombinant Protein
12569	HIV-1 gp120 Recombinant Protein (AE.A244 D11gp120)
12570	HIV-1 gp120 Recombinant Protein (B.MN D11gp120)
12571	HIV-1 gp120 Recombinant Protein (B.9021 D11gp120)
12572	HIV-1 gp140 Recombinant Protein (B.6240 gp140C)
12574	HIV-1 gp120 Recombinant Protein (B.63521 D11gp120 mutC)
12576	HIV-1 gp120 Recombinant Protein (M.CON-S D11gp120)
12581	HIV-1 gp140 Recombinant Protein (C.1086 gp140C)
13055	HIV-1 AC10.29 gp120 Avi His Recombinant Protein
13342	HIV-1 93TH975 gp120 Recombinant Protein
12567	HIV-1 Env V1V2 Recombinant Protein (AE.A244 V1V2.tags)
12568	HIV-1 Env V1V2 Recombinant Protein (C.1086 V1V2.tags)
8660	HIV-1 96ZM651.8 gp140 Optimized Expression Vector
12806	HIV-1 CM235 gp120 Expression Vector (pCI.CM235.gp120)
12957	HIV-1 AC10.29 gp120 Avi His Optimized Expression Vector
13348	HIV-1 BaL gp120 His Expression Vector
13349	HIV-1 93TH975 gp120 His Expression Vector
13350	HIV-1 CN54 gp120 His Expression Vector
12551	CH59 mAb
12550	CH58 mAb

1180

1181

1182

1183

1184

1185

1186

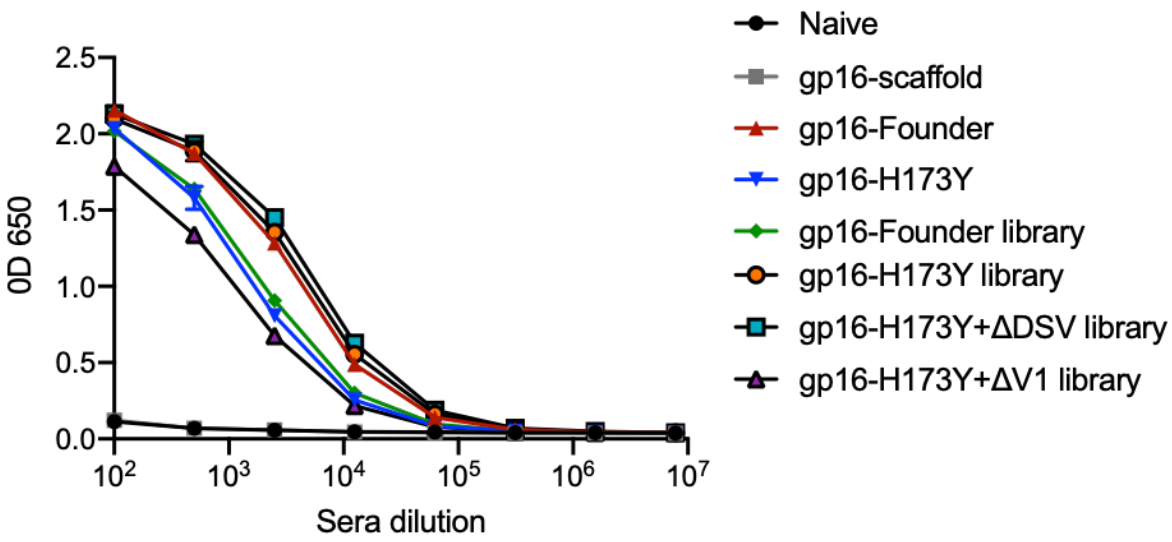


Fig S6. V1V2-specific binding responses in the terminal bleed sera. V1V2-antibodies were detected in mice groups immunized with buffer (no antigen, Naïve group), gp16-scaffold only (no V1V2 control), gp16-V1V2- T/F07, H173Y, T/F07 library, H173Y library, H173Y+ΔDSV library and H173Y+ΔV1 library. A color-coded key is provided on the right side of the graph for each binding curve. Both naïve and gp16-scaffold only groups showed no non-specific reactivity towards the coating antigen. The antibody titers are determined through ELISA. Respective purified recombinant soluble gp140-T/F07, -H173Y, -ΔDSV, -H173Y+ΔDSV and –H173Y+ΔV1 Env glycoproteins were used as coating antigens (1μg/ml) matching the V1V2 region (parental template mutations for combinatorial libraries). Triplicate absorbance (OD 650 nm) readings are used to generate binding curves.

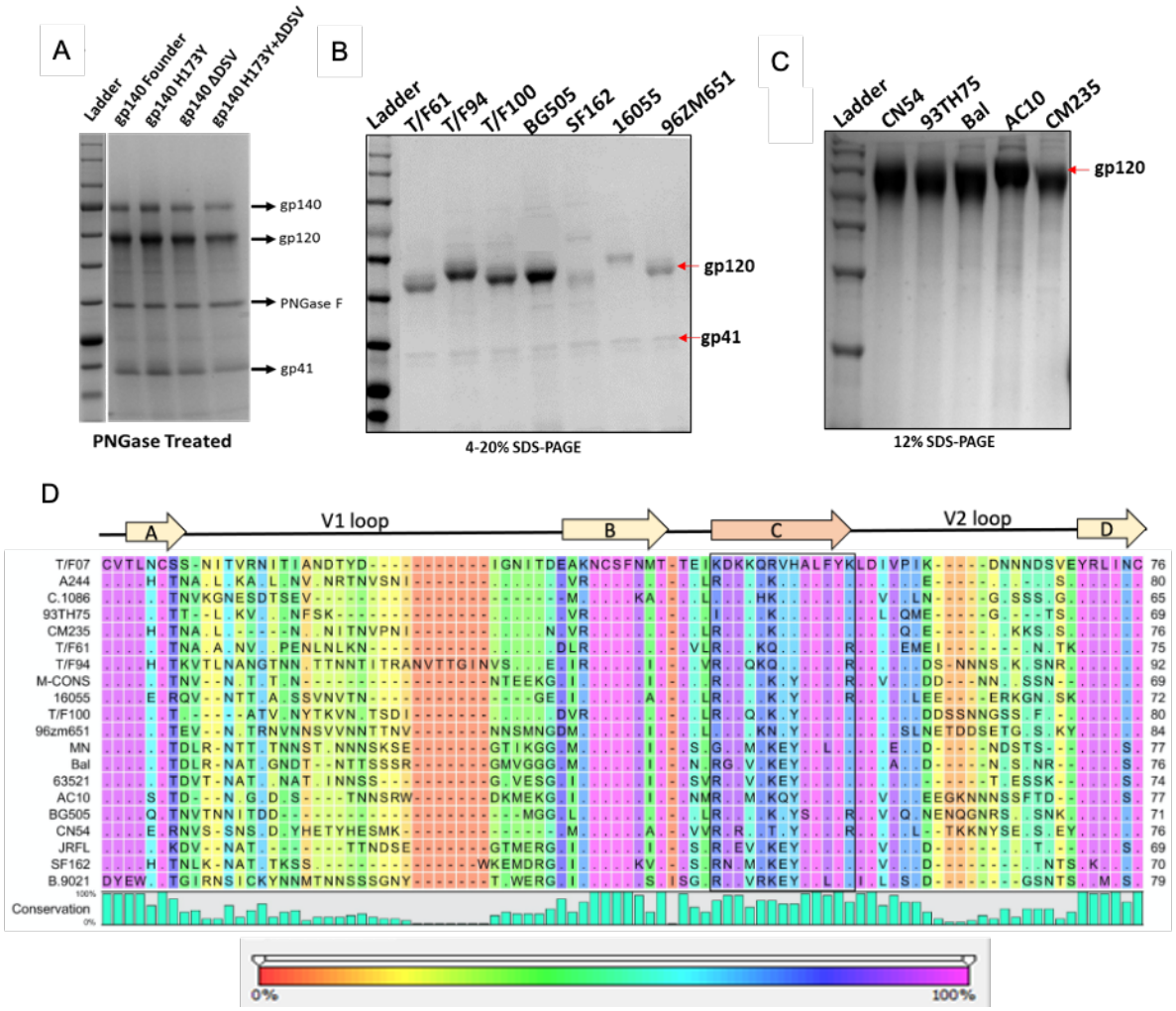


Fig S7. Antigens purified to determine cross-reactive responses and breadth. (A-C) Reducing SDS-PAGE gel profile of GnTi expressed recombinant His-tagged Env proteins, gp140-T/F07 and its V2 mutants (A); gp140s (cleaved into gp120 and gp41 subunits) (B) and gp120s (C) of different HIV-1 subtypes used as heterologous Env antigens. (D) V1V2 sequences of the diverse HIV-1 subtypes included in the heterologous Env protein library used to determine breadth. Degree of conservation (0-100%) at each residue position is depicted graphically at the bottom of the alignment. Variability in the V1V2 region of the chosen Env antigens is shown with background color gradient (red to pink) showing conservation on a scale of 0-100%.

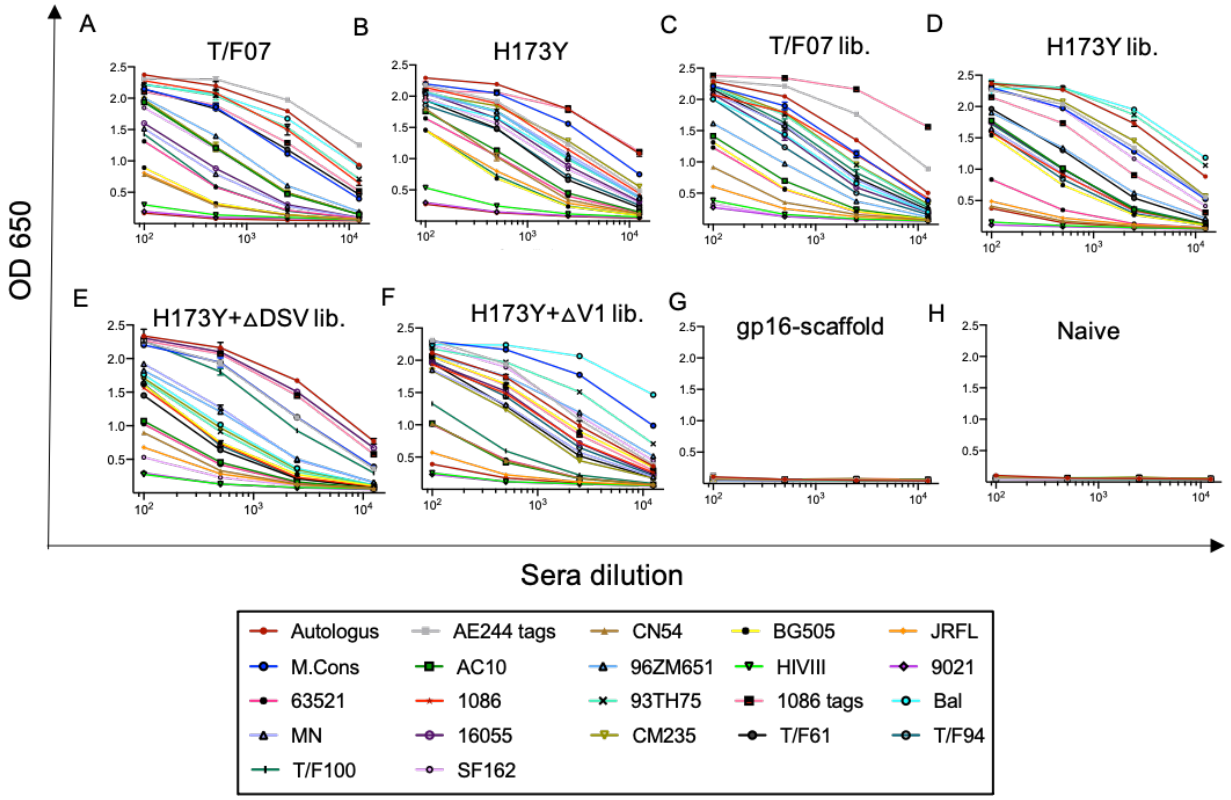


Fig S8. Breadth analysis of V2 combinatorial library immunogens using heterologous Env antigen library. (A-H) ELISA generated binding curves showing the reactivity of sera of mice groups immunized with T/F07 (A) H173Y (B), and combinatorial V2 libraries (lib.); (C) T/F07 (D) H173Y (E) H173Y+ΔDSV (F) H173Y+ΔV1 groups (G). gp16-scaffold (H) and Naïve (I) groups sera were used as negative controls. Experiment was performed with 5-fold serially diluted pooled sera from each group in triplicates. Curves are color-coded for respective antigen provided in the legend at the bottom of the panel.