

Immunity, Volume 50

Supplemental Information

**An Antigenic Atlas of HIV-1 Escape
from Broadly Neutralizing Antibodies
Distinguishes Functional and Structural Epitopes**

Adam S. Dingens, Dana Arenz, Haidyn Weight, Julie Overbaugh, and Jesse D. Bloom

Supplemental Figures and Files

Data S1 | The excess fraction surviving values plotted across the length of the mutagenized portion of Env for each antibody. Related to Figures 2-5. The underlay indicates contact sites, sites of significant escape, and the overlap between these groups of sites, as in Figure 2B, 2C.

Data S2 | The excess fraction surviving PGT151 and VRC01 plotted across the mutagenized portion of different Envs. Related to Figures 2, S6, and S7.

Data S3 | The computational analysis. Related to Figure 2 and DATA AND SOFTWARE AVAILABILITY in the STAR Methods. A zip file containing an executable Jupyter notebook, an HTML version of the notebook, and all of the necessary input data to run the analysis.

Data S4 | The mutation fraction surviving and excess mutation fractions surviving datasets. Related to Figure 2. A zip file containing three CSV files for each antibody. One contains the mutation fraction surviving estimates, and one contains the excess mutation fraction surviving estimates plotted in the paper. The third contains *differential selection* estimates, which are log-transformed relative enrichment ratios that may be of use for certain purposes, such as examining mutations that are differentially depleted, rather than enriched, upon antibody selection. The differential selection metric is described in detail in (Doud et al., 2017) and at https://jbloomlab.github.io/dms_tools2/diffsel.html. Sites are numbered according to HXB2 reference strain numbering.

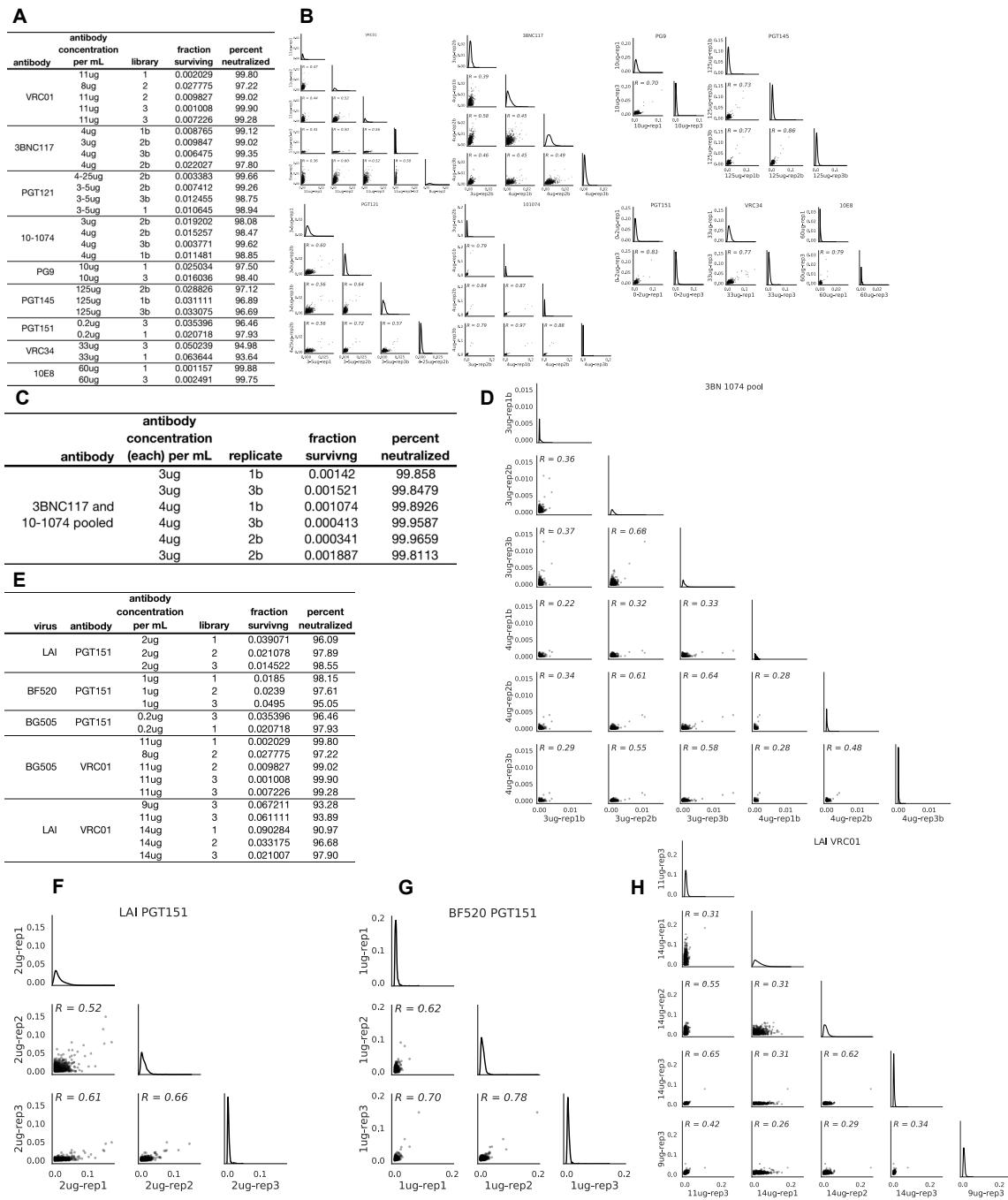


Figure S1 | The fraction surviving measurements and correlation between mutational antigenic profiling biological replicates. Related to Figure 2, 6, S6, and S7.

A. For each biological replicate, the antibody concentration used during the selection, which mutant virus library was used, and the fraction of that library that survived antibody selection is shown. For clarity, the percent neutralized (1- library

fraction surviving) × 100 is also shown. **B.** The correlation between the average excess fraction surviving at each site for each biological replicate, for each antibody. **C.** As in A, but for biological replicates of pooled 3BNC117 and 10-1074. **D.** As in B, but for pooled 3BNC117 and 10-1074. **E.** As in A, but for biological replicates for PGT151 and VRC01 using LAI and BF520 mutant Env libraries. **F.** As in B, but for PGT151 escape using LAI mutant Env libraries. **G.** Same as F, but for BF520 mutant Env libraries. **H.** The same as F, but for escape from VRC01 using LAI mutant virus libraries.

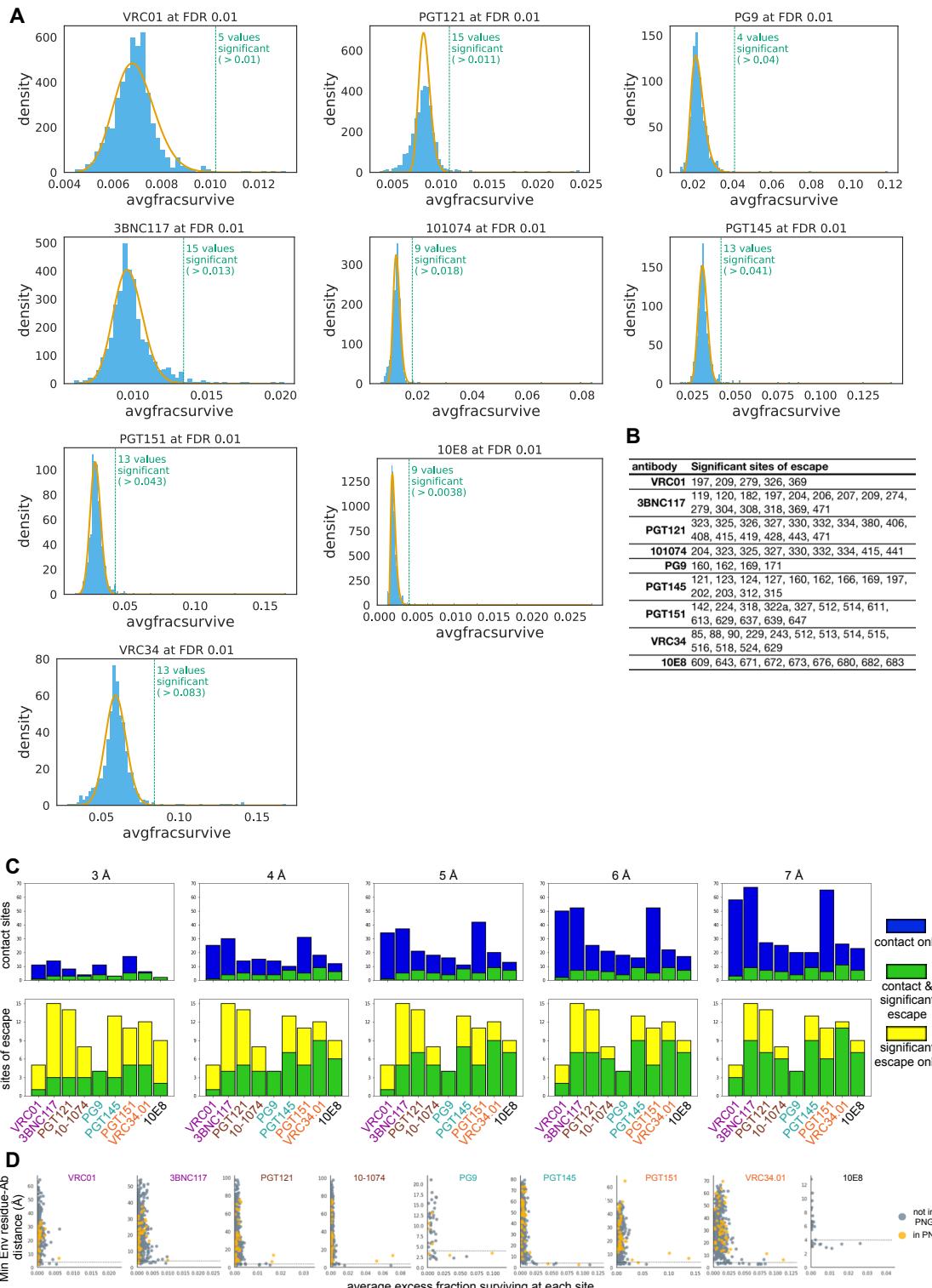


Figure S2 | Identification of significant sites of viral escape and the overlap between each antibody's structural and functional epitope. Related to Figure 2. A.

For each antibody, the distribution of the average fraction surviving at each site is plotted in blue. A gamma distribution fit to the site fraction surviving values using robust regression is overlaid in yellow. A dotted line shows sites that fall beyond this distribution at a fall discovery rate of 0.01, and the number of sites that beyond this cutoff is labeled. Code that performs this analysis is at

https://jbloomlab.github.io/dms_tools2/dms_tools2.plot.html#dms_tools2.plot.findSigSel.

B. A table listing all of the significant sites of viral escape for each antibody.
C. As in Figure 2B and 2C, but using different distance cutoffs between non-hydrogen Env and antibody atoms to determine contact sites. **D.** For each Env residue in our library the minimum distance to the antibody is plotted against that site's excess fraction surviving averaged across mutations. The 4 Å distance cutoff used in Figure 2B and 2C is plotted with a dotted line. Sites that fall within PNGs are colored yellow, while other sites are grey. Only sites that are in the structures used to determine distance cutoffs are plotted (see STAR Methods for details of the structures used).

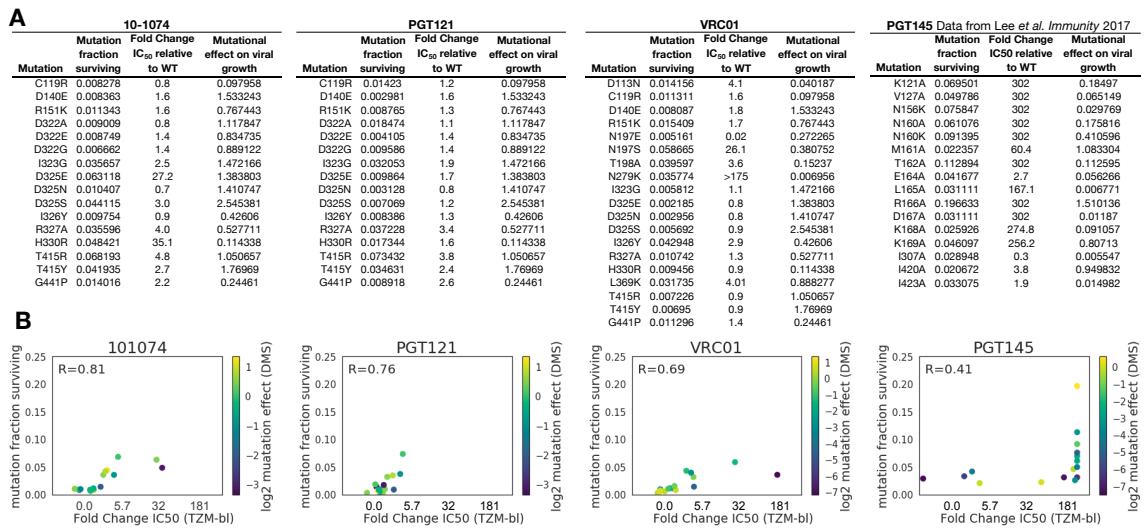


Figure S3 | Mutational antigenic profiling results are validated by TZM-bl

neutralization assays. Related to Figure 2. A. The tables list mutations that were selected for validation in neutralization assays. Data for PGT145 was taken from (Lee et al., 2017). For each mutation, the table gives the mutation's fraction surviving antibody, fold change in IC₅₀ in the TZM-bl neutralization assay, and the mutation's effect on viral growth. The mutation's effect on viral growth is calculated from our prior deep mutational scanning of Env for viral growth in cell culture, in the absence of any immune selection (Haddox et al., 2018). The mutational effect is the ratio of the preference for that mutant amino acid relative to the wildtype amino acid at that site. If the mutational effect is >1, then that mutation grows better than wildtype in cell culture; if it is <1, that mutant grows worse than wildtype. **B.** For each antibody, we plot the correlation between the mutation fraction surviving and the fold change in IC₅₀ relative to wildtype from TZM-bl assays. Points are colored according the log₂ transformed values displayed in A). Where there are discrepancies between mutational antigenic profiling and TZM-bl assays, these mutants (e.g. N279K for VRC01; H330R for 10-1074) often have log₂ mutational effects on viral growth <<0 (darker blue), indicating they are deleterious for viral growth in cell culture.

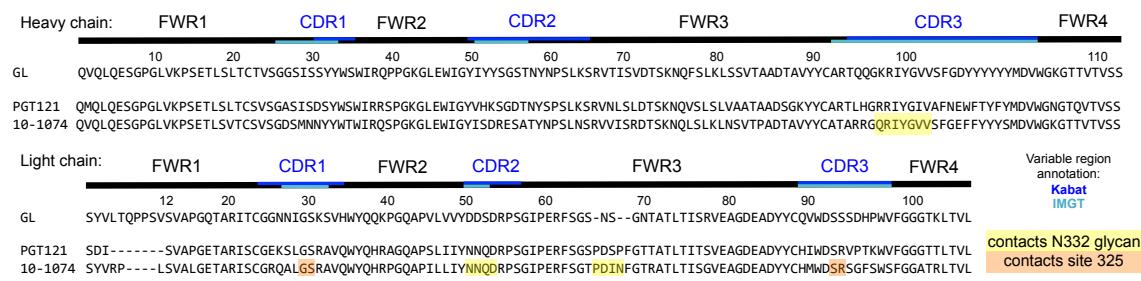


Figure S4 | Sequence alignment of PGT121 and 10-1074 antibody chains. Related to Figure 3. The heavy (top) and light chain (bottom) genes of the clonal variants PGT121 and 10-1074 are aligned, along with their inferred germline (GL). Kabat and IMGT variable regions annotations are shown. There are several sites in Env where the effects of mutations differ markedly between PGT121 and 10-1074, most prominently sites 325 and the N332 PNG. We used the 10-1074 / Env co-crystal structure 5TZ3 to identify sites in 10-1074 that contact these Env sites, and have highlighted them in the alignment above PGT121 contact sites are not labeled, as there is not a high-resolution structure of available. The germline inference, alignment, and variable loop annotation are from Mouquet *et al.* 2012, which also discusses additional structural differences between these antibodies (such as differences in the light chain loops of the unliganded antibody structures).

Antibody	Study	ClinicalTrials.gov Identifier (NCT Number)	Study description	Overlap between reported sites of viral evolution during bnAb immunotherapy* and significant sites of viral escape from our mutational antigenic profiling
10-1074	Casket et al 2017	NCT02511990	Single infusion of 10-1074 to HIV infected individuals	325, 332, 334^A
3BNC117	Caskey et al 2015, Schoofs et al 2016	NCT02018510	Single infusion of 3BNC117 to HIV infected individuals	209, 279, 308, 318, 471^B
3BNC117	Scheid et al 2016	NCT02446847	Multiple (2-4) infusions of 3BNC117 after the discontinuation of antiretroviral therapy	274^C
VRC01	Lynch et al 2015a	NCT01950325	One or two infusions of VRC01 in antiretroviral treated and untreated HIV-infected patients, respectively	none ^D
VRC01	Bar et al 2016	NCT02463227 and NCT02471326	Multiple (3-8) infusions of VRC01 after the discontinuation of antiretroviral therapy	279^E ; reanalysis of these data revealed potential selection at 326**

* The sites in this table were determined in each study individually, using disparate methods, explained below.

A: Discussed in main figures and text.

B: Used LASSIE (Longitudinal Antigenic Sequences and Sites from Intrahost Evolution) to identify sites selected within the 24-week time frame, using a selection cutoff changing $\geq 80\%$ amino acid frequency compared to baseline. Indels are omitted from this group.

C: Discussed in text based on sequence alignments.

D: Used a neutralization-based epitope prediction (NEP) algorithm to predict mutational differences that could be associated with VRC01 selection. The top 5 highest scoring sites for the four patients examined are reported here.

E: Two separate criteria were used by the two clinical trials reported in Bar et al 2016. In one, the VRC01 antibody footprint sequence was analyzed using LASSIE, and we then identified sites within the examined regions that differed $\geq 50\%$ amino acid frequency compared to baseline. In the other study, HIV env sequences were analyzed using a neutralization-based epitope prediction (NEP) algorithm. Changes in amino-acid residues that occurred within or next to the VRC01 epitope were reported.

**Examination of *env* genotype from this study in the context of our VRC01 escape profile revealed a potential additional site of viral escape not discussed in Bar et al. 2016. In Patient V10, T326 was fixed shortly after VRC01 treatment, but T326 was present in only 2/49 sequences at later timepoints when antibody levels may have waned (I326, present in the remaining sequences, is 96.8% conserved in the LANL filtered web alignment) (Bar et al., 2016). No pretreatment sequences were available.

Table S1 | Overlap between mutational antigenic profiling sites of escape and sites of viral evolution that occurred *in vivo* during bnAb immunotherapy. Related to Figures 3 and 4.

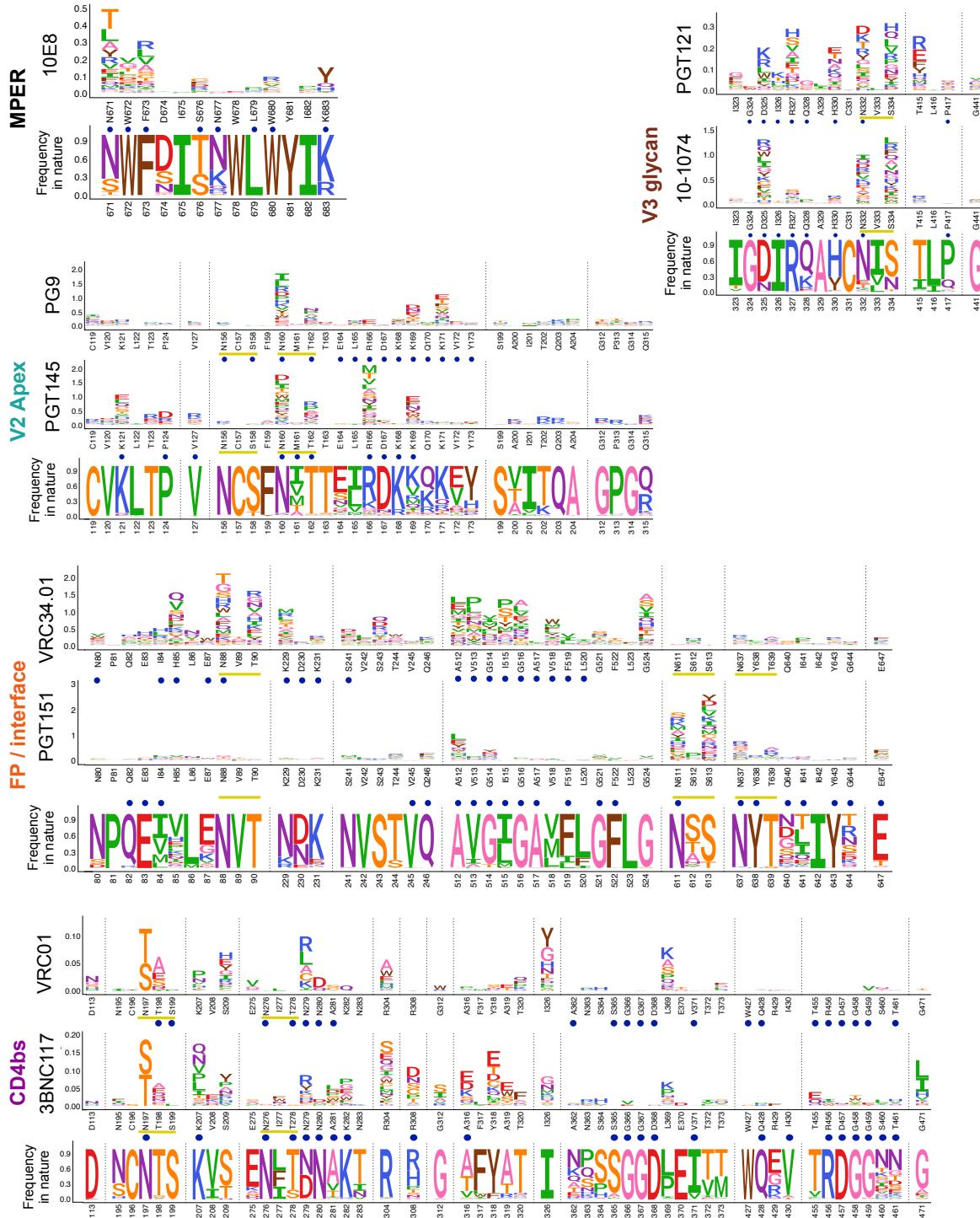


Figure S5 | The natural sequence variation of escape mutations. Related to

Figure 7. For each epitope, the excess fraction surviving is shown for each

antibody, as in Figures 3-5. Blue circles indicate antibody structural contacts and yellow underlines indicate glycosylation motifs. On bottom, the frequency of amino acids in nature is also plotted. Natural sequence variation is based on LANL's group M filtered Web Alignment (Foley et al., 2017).

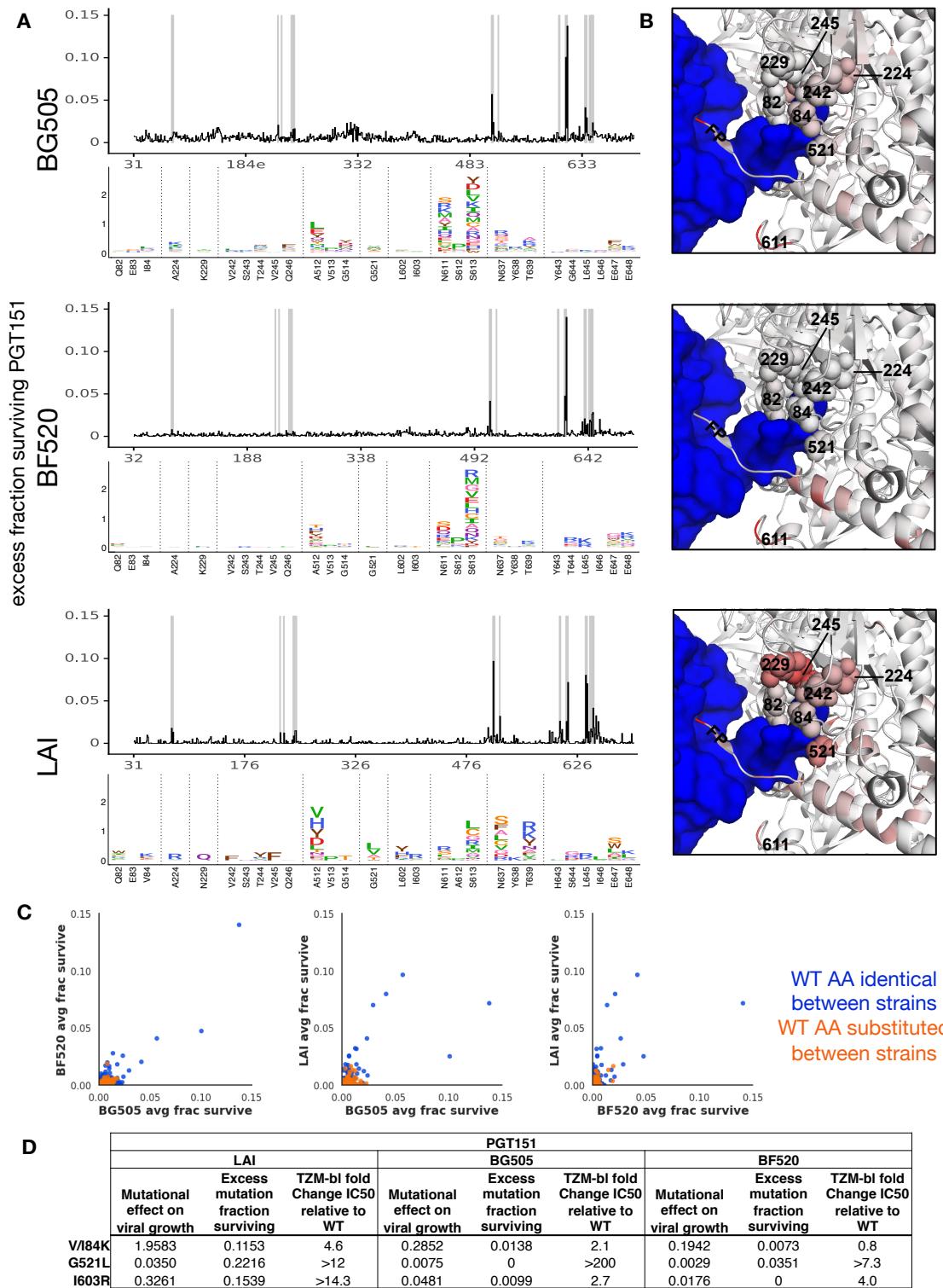


Figure S6 | Differences in PGT151 escape across Envs. Related to Figure 2.

A. The line plots shows each site's average excess fraction surviving PGT151 for BG505, BF520, and LAI Env. Beneath, logoplots show the mutation level escape

profile for sites highlighted in grey in the line plots. The entire mutation-level escape profile for each Env is available in File S2. **B.** View of the PGT151 epitope colored according to escape in each Env. PGT151 Fab is colored blue, and Env is colored according to the maximum mutation fraction surviving at each site for each strain (PDB:5FUU). **C.** The correlation between BG505, BF520, and LAI excess fraction surviving PGT151 for each site. Sites are colored blue if the wildtype amino acid is identical between strains, orange if they differ. **D.** The mutational effect, excess mutation fraction surviving, and fold change in PGT151 IC₅₀ from TZM-bl assays using point mutant pseudoviruses of validated point mutants. As in Figure S4, the mutational effect is measured via deep mutational scanning of each Env under selection for viral replication in cell culture in the absence of any immune selection (Haddox et al 2016, Haddox et al 2018). The mutational effect is the ratio of the preference of that mutant relative to the wildtype amino acid preference at that site. If the mutational effect is >1, then that mutation replicates better than wildtype in cell culture; if it is <1, that mutant replicates worse than wildtype. Where there are discrepancies between mutational antigenic profiling and TZM-bl assays, that mutation has an effect <<1, indicating it is deleterious for viral replication in cell culture.

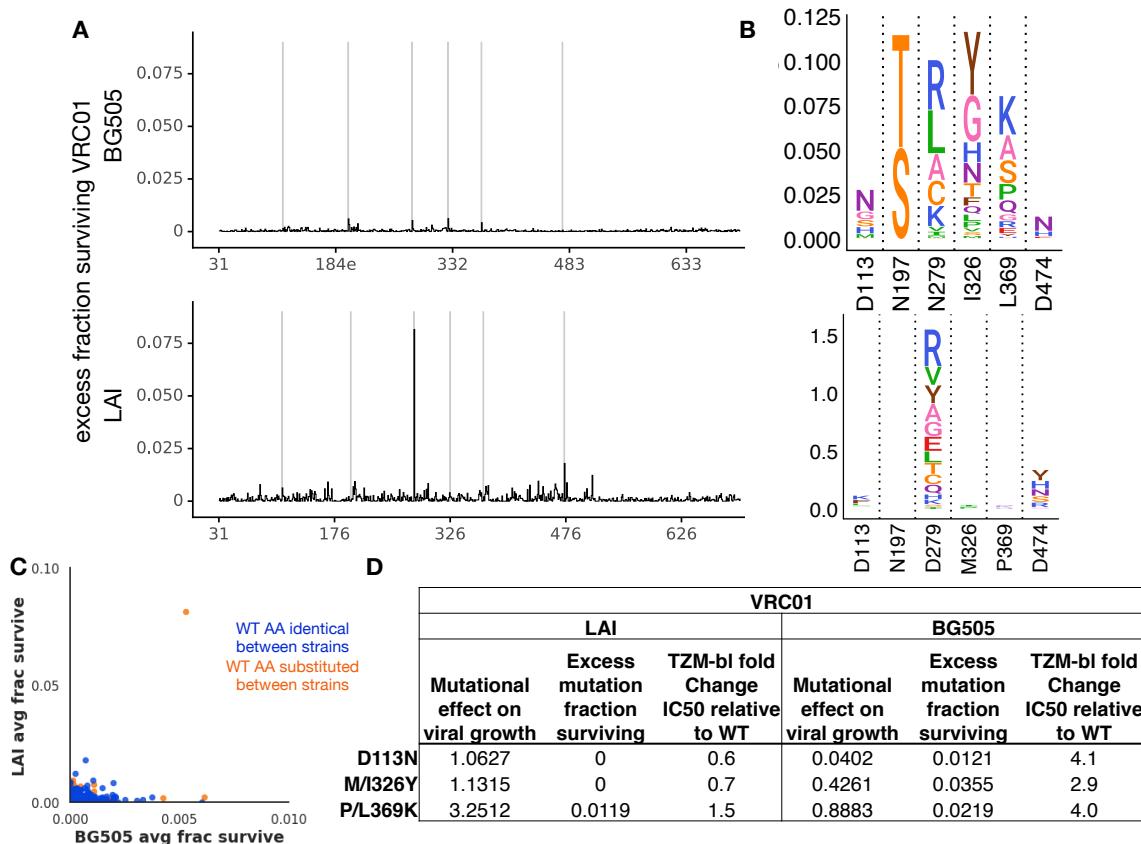


Figure S7 | Differences in VRC01 escape across Envs. Related to Figure 2.

A. The line plots shows each site's average excess fraction surviving VRC01 for BG505 and LAI Env. **B.** Logoplots show the mutation level escape profile for sites highlighted in grey in A. The entire mutation-level escape profile for both Envs is available in File S2. **C.** Same as in S6C, but for VRC01 escape in BG505 and LAI. **D.** Same as in S6D, but for validated VRC01 escape mutants.