Figure S1

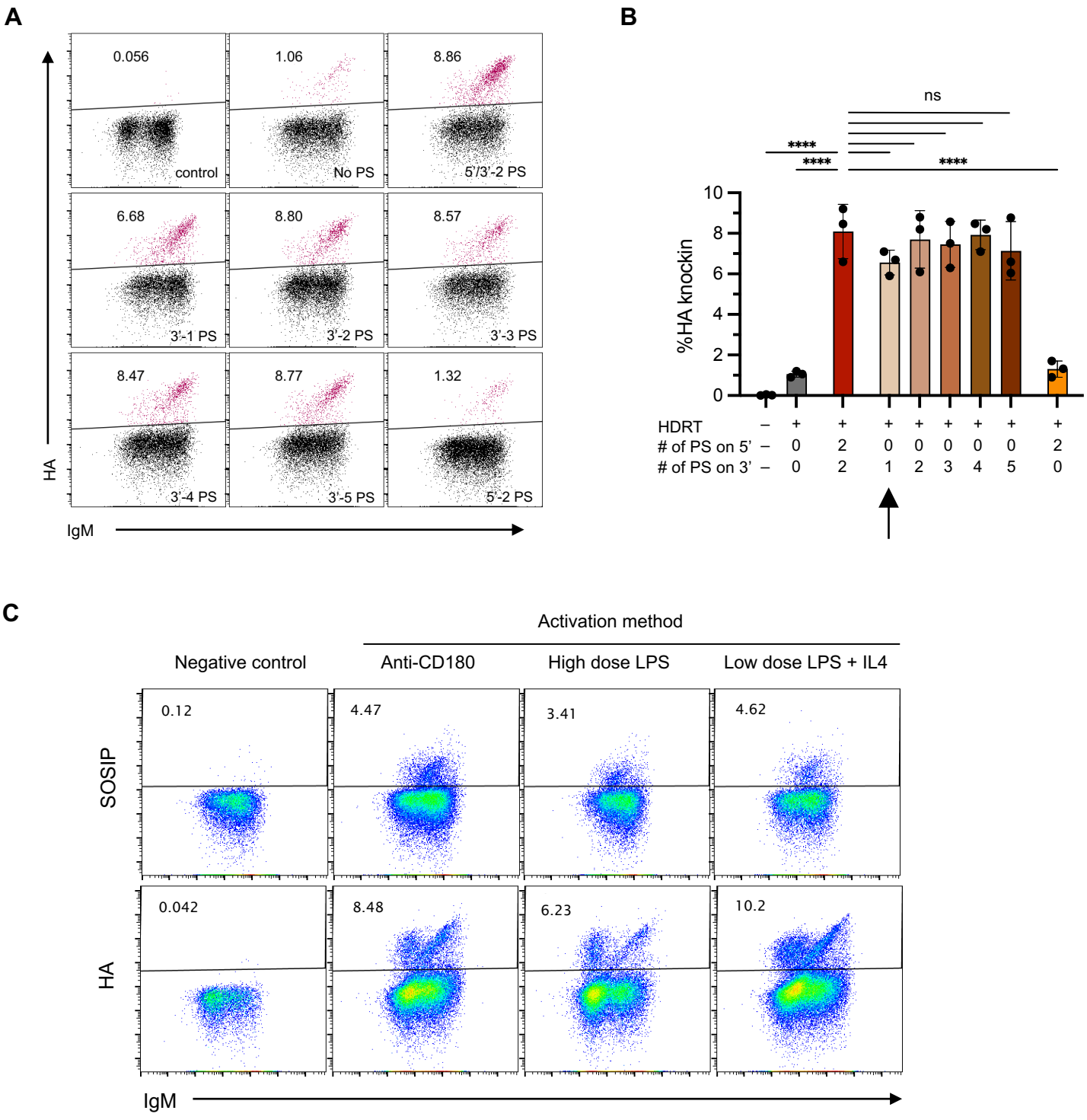

A
0.056
control
1.06
No PS
8.86
5'/3'-2 PS
6.68
3'-1 PS
8.80
3'-2 PS
8.57
3'-3 PS
8.47
3'-4 PS
8.77
3'-5 PS
1.32
5'-2 PS
HA
IgM
B
ns
****
****
****
%HA knockin
10
8
6
4
2
0
HDRT – + + + + + + + +
# of PS on 5' – 0 2 0 0 0 0 0 2
# of PS on 3' – 0 2 1 2 3 4 5 0
C
Activation method
Negative control
Anti-CD180
High dose LPS
Low dose LPS + IL4
SOSIP
0.12
4.47
3.41
4.62
HA
0.042
8.48
6.23
10.2
IgM

Figure S2

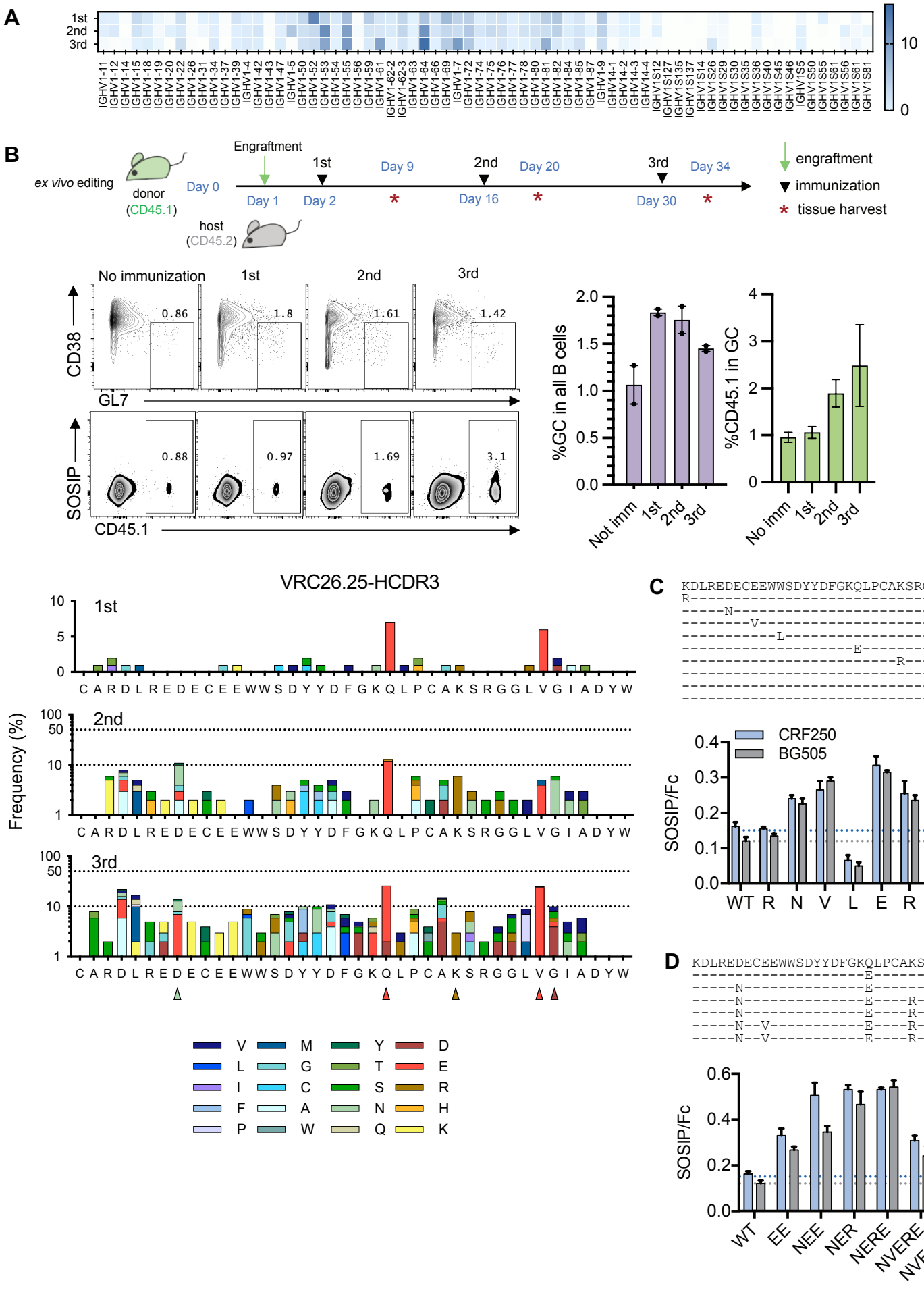
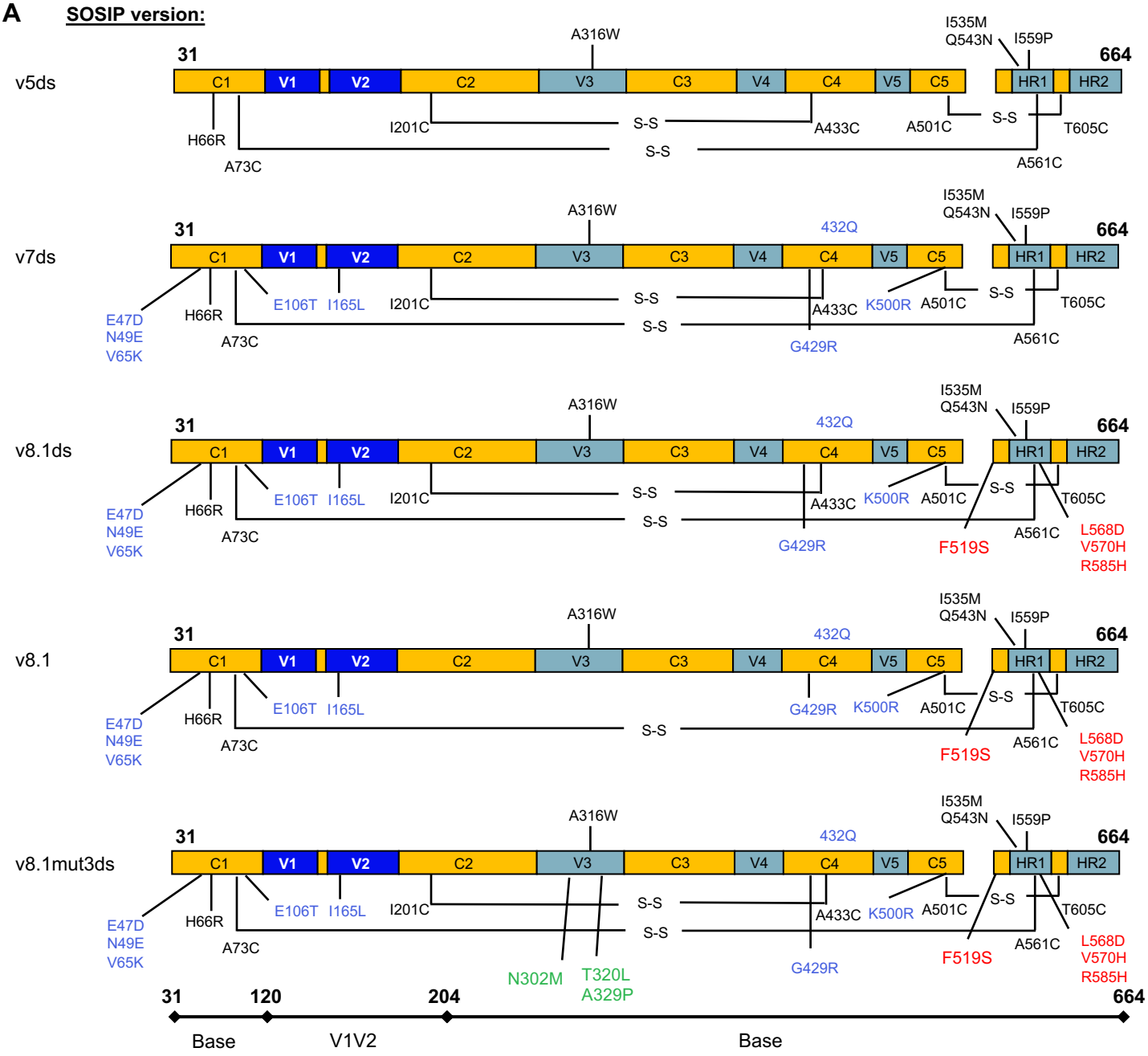

A
1st
2nd
3rd
10
0
B
ex vivo editing
donor
(CD45.1)
host
(CD45.2)
Day 0
Engraftment
Day 1
1st
Day 2
Day 9
2nd
Day 16
Day 20
3rd
Day 30
Day 34
engraftment
immunization
tissue harvest
No immunization
1st
2nd
3rd
CD38
GL7
0.86
1.8
1.61
1.42
SOSIP
CD45.1
0.88
0.97
1.69
3.1
%GC in all B cells
%CD45.1 in GC
Not imm
1st
2nd
3rd
No imm
VRC26.25-HCDR3
Frequency (%)
C A R D L R E D E C E E W W S D Y Y D F G K Q L P C A K S R G G L V G I A D Y W
C
KDLREDECEEWWSDYYDFGKQLPCAKSRGGLVGIADN
R
N
V
L
E
R
E
D
Y
CRF250
BG505
SOSIP/Fc
WT R N V L E R E D Y
D
KDLREDECEEWWSDYYDFGKQLPCAKSRGGLVGIADN
WT
EE
NEE
NER
NERE
NVERE
NVERED

Figure S3

A **SOSIP version:**



B **V1V2 region:**

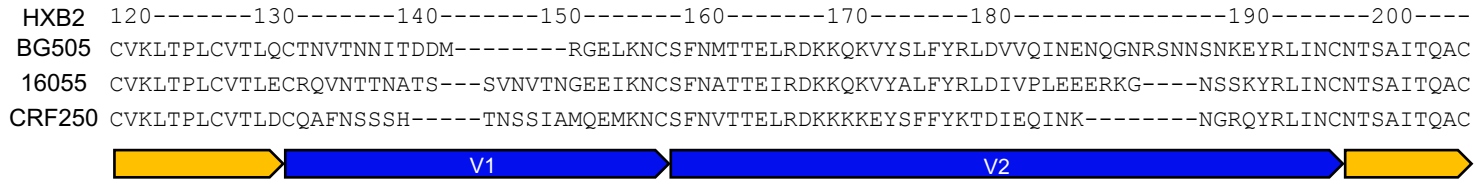
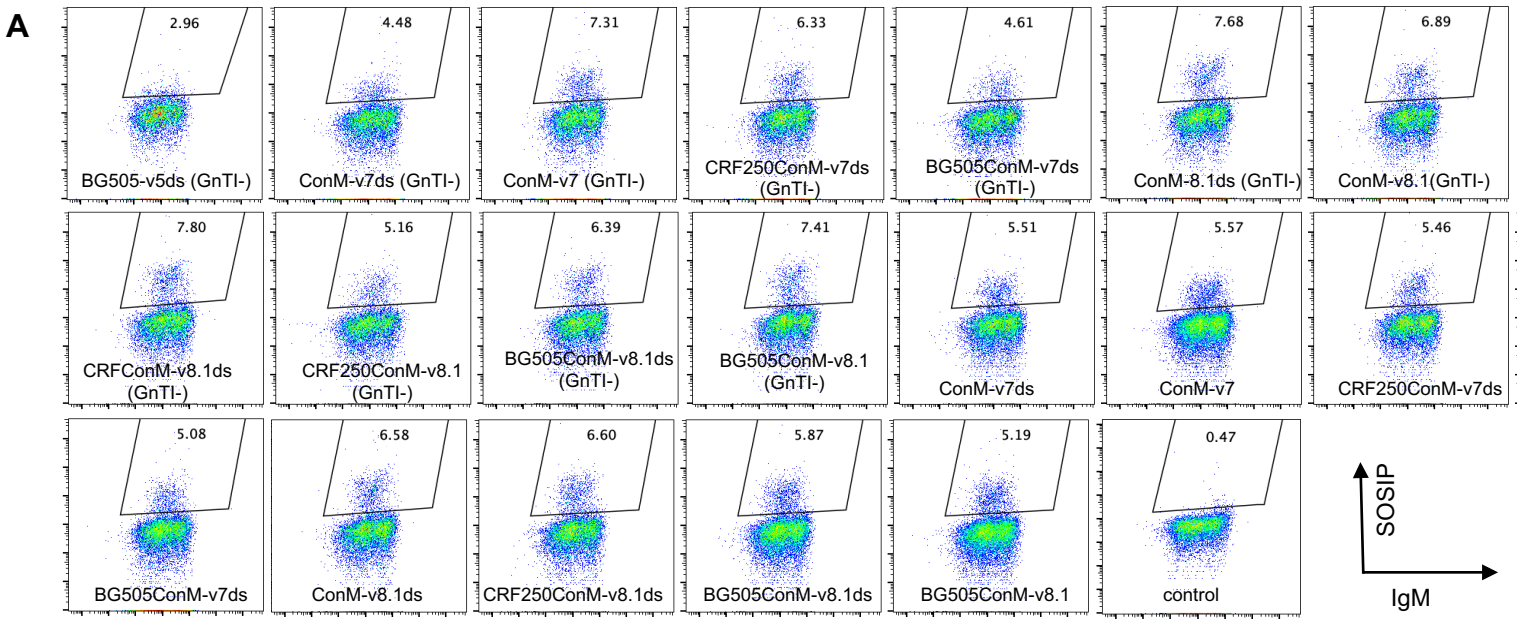
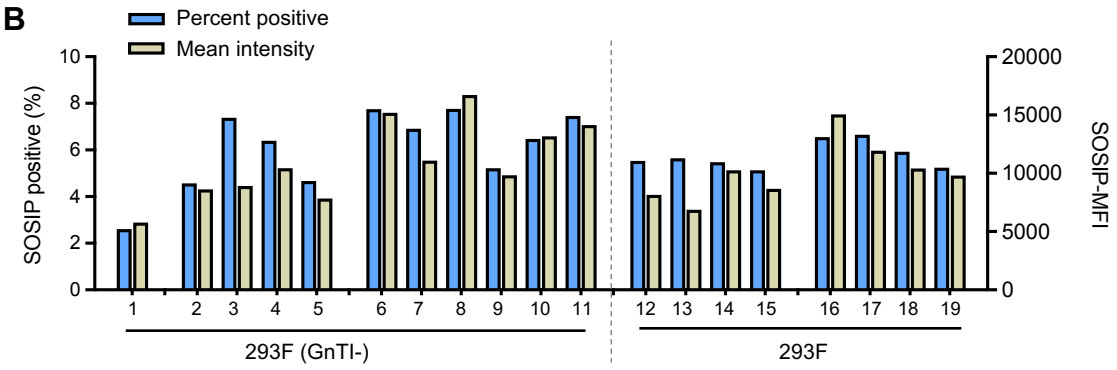
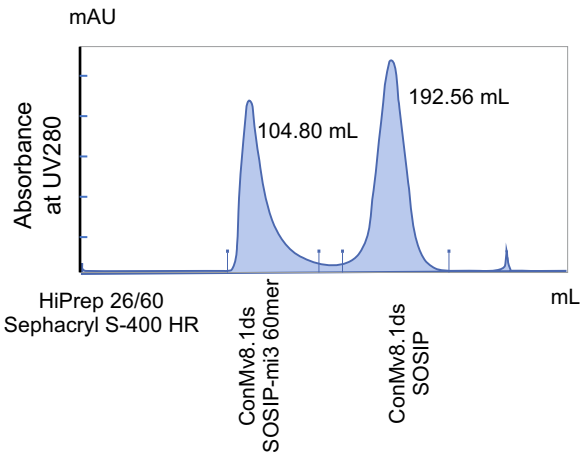


Figure S4





1. BG505ConM-v5ds (GnTI-)
2. ConM-v7ds (GnTI-)
3. ConM-v7 (GnTI-)
4. CRF250ConM-v7ds (GnTI-)
5. BG505ConM-v7ds (GnTI-)
6. ConM-v8.1ds (GnTI-)
7. ConM-v8.1 (GnTI-)
8. CRF250ConM-v8.1ds (GnTI-)
9. CRF250ConM-v8.1 (GnTI-)
10. BG505ConM-v8.1ds (GnTI-)
11. BG505ConM-v8.1 (GnTI-)
12. ConM-v7ds
13. ConM-v7
14. CRF250ConM-v7ds
15. BG505ConM-v7ds
16. ConM-v8.1ds
17. CRF250ConM-v8.1ds
18. BG505ConM-v8.1ds
19. BG505ConM-v8.1



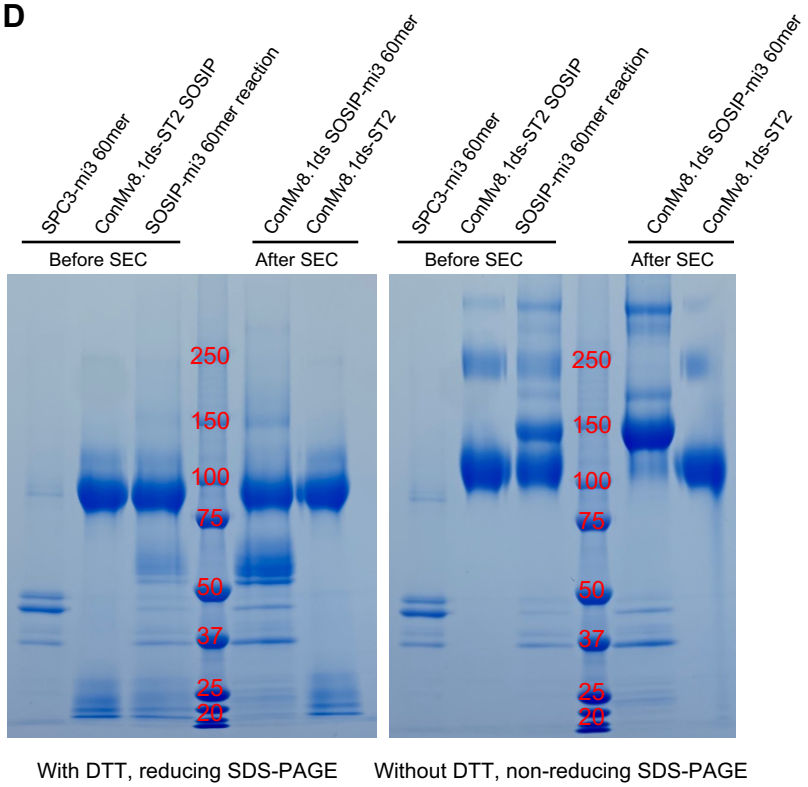
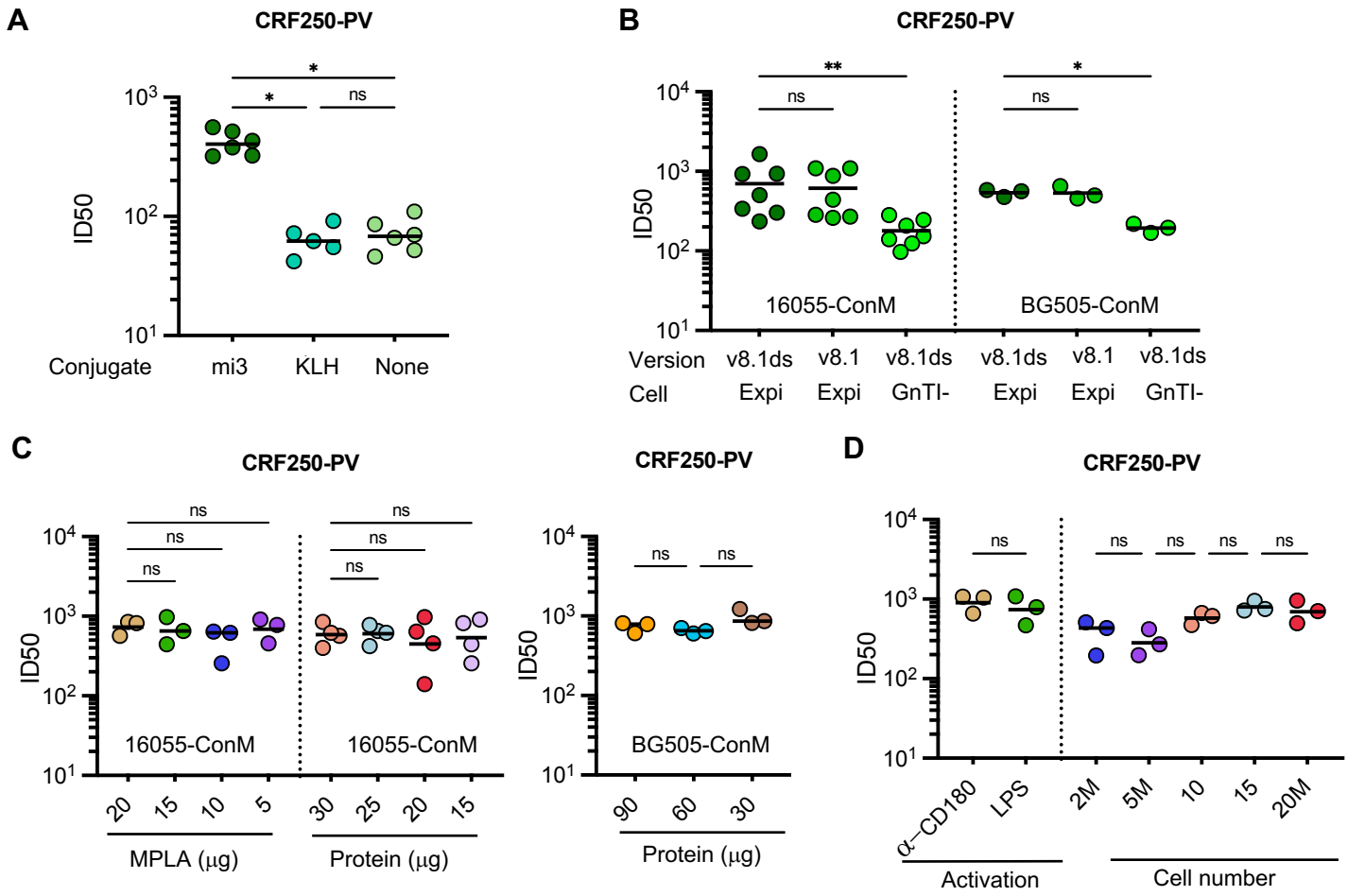


Figure S5



E

		1st immunization		2nd immunization		3rd immunization	
Psudoviruses	Cade	engraft	unengraft	engraft	unengraft	engraft	unengraft
BG505	A	<20	<20	<20	<20	23.5	<20
ZM233	C	<20	<20	95.2	<20	193.7	<20
CRF250	AG	213	<20	992	<20	4390	<20
16055	C	102	<20	421	<20	1120	<20
25710	C	73.2	<20	122	<20	297.5	<20
x1632	G	<20	<20	<20	<20	26.6	<20
CNE8	AE	<20	<20	<20	<20	<20	<20
WITO	B	<20	<20	<20	<20	<20	<20
% neutralized		3/8	0/8	4/8	0/8	6/8	0/8

ID50
>1000
300-1000
100-300
20-100
<20

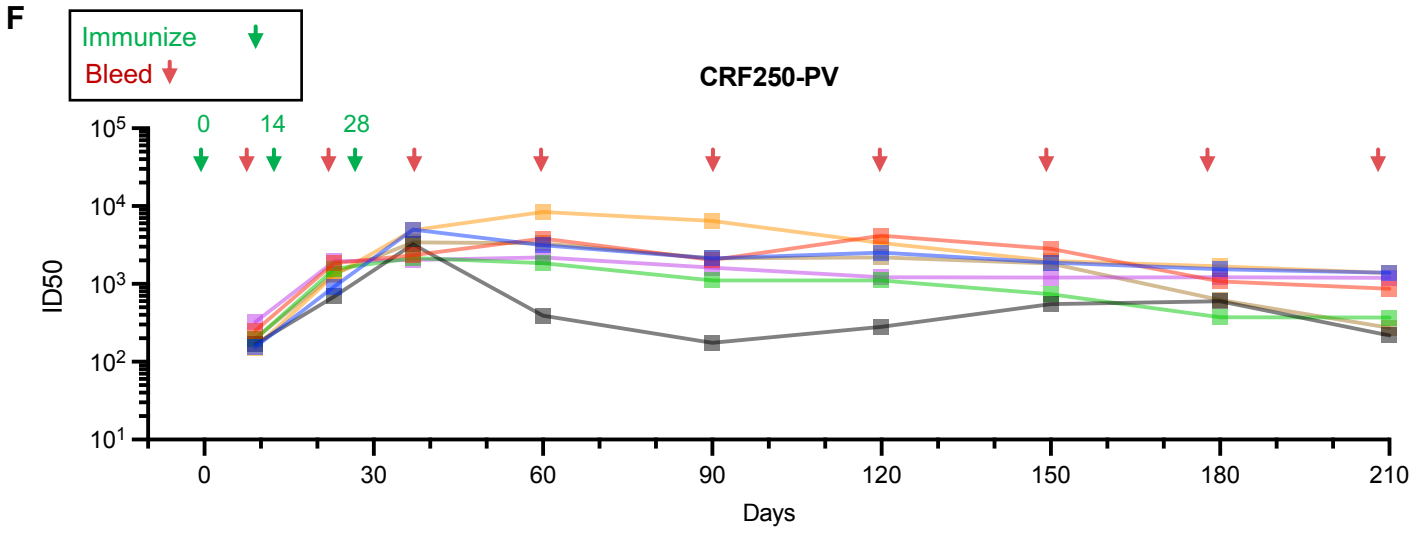


Figure S6

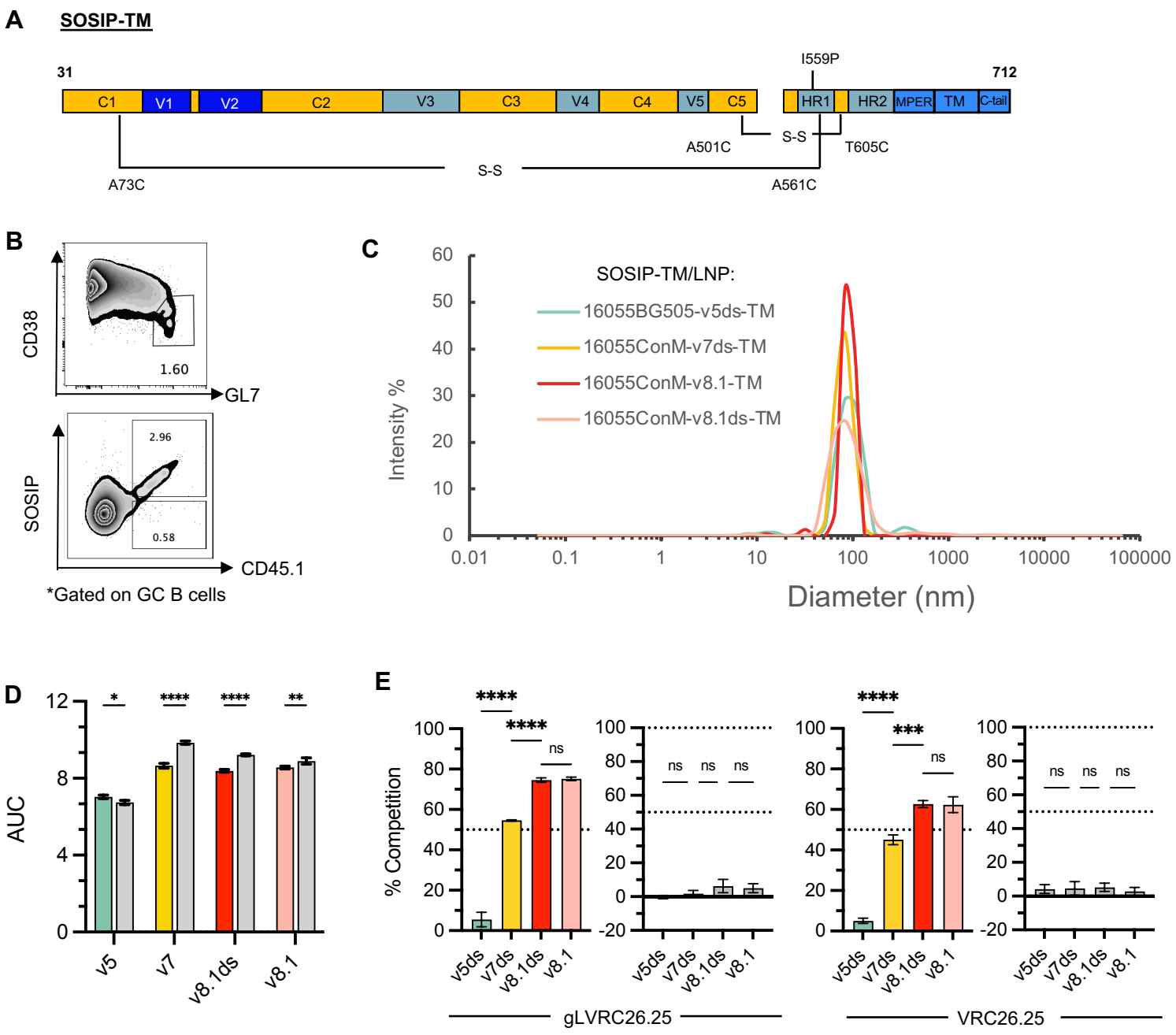
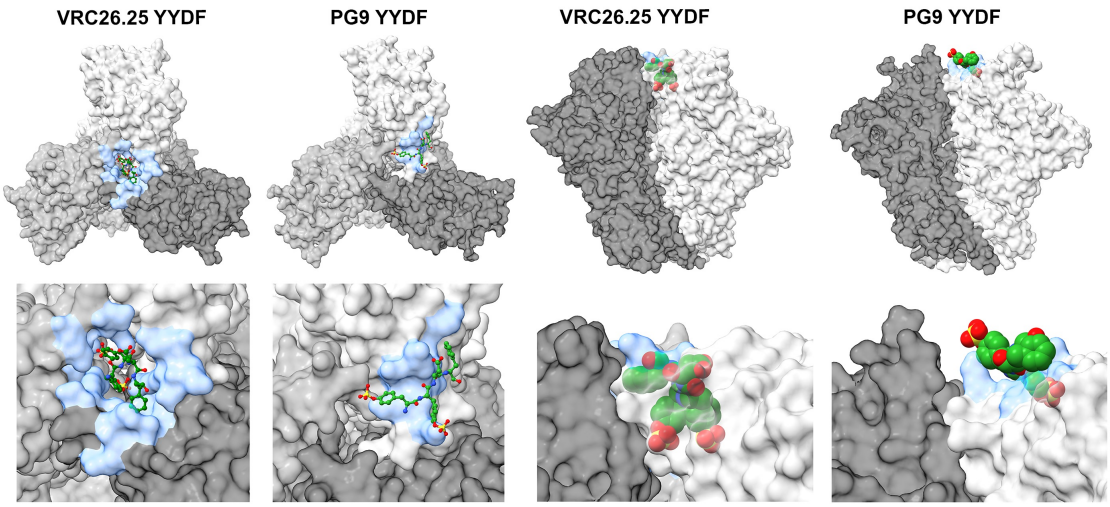

A
SOSIP-TM
31
712
I559P
C1
V1
V2
C2
V3
C3
V4
C4
V5
C5
HR1
HR2
MPER
TM
C-tail
A501C
S-S
T605C
A73C
S-S
A561C
B
CD38
1.60
GL7
SOSIP
2.96
0.58
CD45.1
*Gated on GC B cells
C
SOSIP-TM/LNP:
16055BG505-v5ds-TM
16055ConM-v7ds-TM
16055ConM-v8.1-TM
16055ConM-v8.1ds-TM
Intensity %
Diameter (nm)
D
AUC
v5
v7
v8.1ds
v8.1
E
% Competition
v5ds
v7ds
v8.1ds
v8.1
gLVRC26.25
VRC26.25

Figure S7

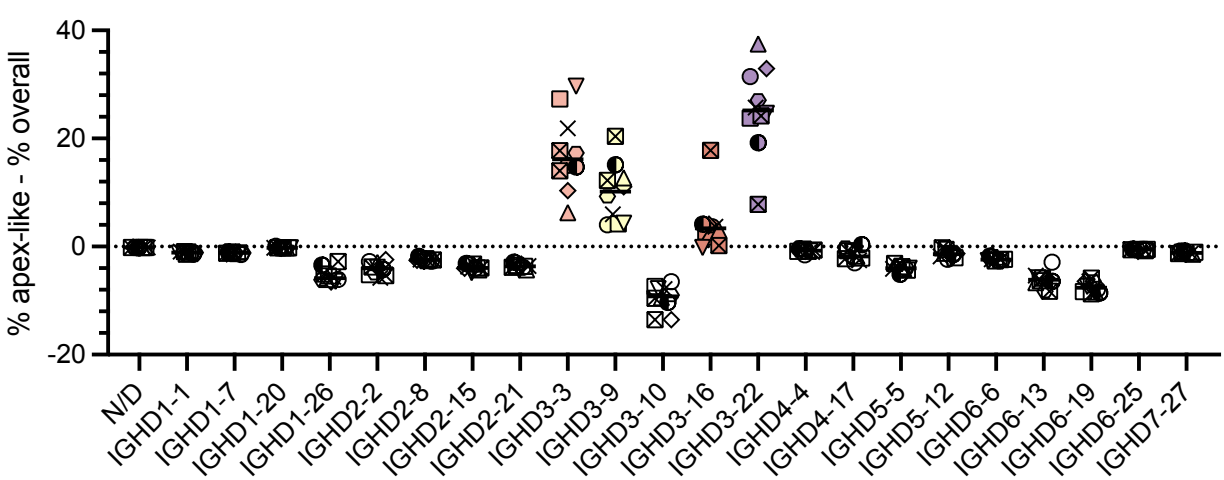
A

```
PG9          CVREAGGPDYRNGYNYYDFYDGYYNYHYMDVW
PG16         CAREAGGPIWHDDVKYYDFNDGYYNYHYMDVW
VRC26.25   CAKDLREDECEEWWSDYYDFGKQLPCAKSRGGLVGIADNW
VRC26 UCA  CAKDLGESENEEWATDYYDFSIGYPGQDPRG.VVGAFDIW
IGHD3-3                  YYDFWSGYYT
IGHD3-9                  YYDILTGYYN
IGHD3-16                 YYDYVWGSYAYT
IGHD3-22                YYYDSSGYYY
```

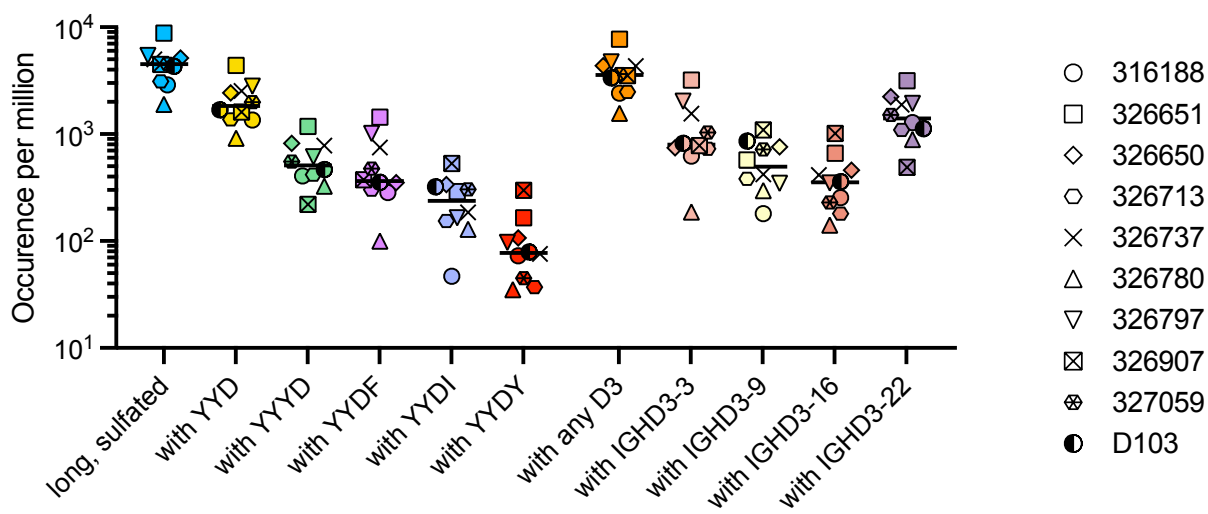
B



C



D



SUPPLEMENTARY FIGURES

Figure S1. Optimization of HCDR3 editing, related to Figure 1 (**A**) The HA tag-encoding homology-directed repair template (HDRT) used in Figure 1C was modified with one to three phosphorothioate (PS) bonds at the 5' or 3' end or both, as indicated. These HDRT were used to introduce the HA tag into the HCDR3 region of primary murine B cells. Cells were then analyzed by flow cytometry with an anti-HA (vertical axis) and anti-mouse IgM antibody (horizontal axis). The percentage of HA-positive cells (dark red) is indicated in each panel. (**B**) A summary of three independent experiments similar to that shown in panel A is shown. Arrow indicates the two 3' PS modification used in all other figures. (**C**) Cells isolated from mouse spleens were activated for 36 to 42 hours with anti-CD180 (4 μg/ml), high dose LPS (50 μg/ml), or low dose LPS (10 μg/ml) and IL-4 (10 μg/ml). Cells were then electroporated with Mb2Cas12a RNP and HDRT encoding an HA-tag or the VRC26.25 HCDR3, replacing endogenous murine HCDR3s, and incubated in the same conditions for another 18 hours. Cells were then analyzed by flow cytometry with an anti-mouse IgM antibody (horizontal axis) and either an APC-labeled anti-HA tag antibody or an APC-labeled SOSIP trimer. Cells activated with anti-CD180 and electroporated with control templates served as a staining control. Numbers indicate the percentage of successfully edited cells. High dose LPS was used in all other figures except where explicitly stated.

Figure S2. NGS analysis of donor murine B cells, related to Figure 2 and 3 (**A**) NGS analysis of IgH genes of donor CD45.1 cells isolated as in **Fig. 2E**, except that the distribution of VH1 family genes is shown for donor cells that expressed the native murine HCDR3. Figure shows results combined from donor B cells harvested from two mice for each time point. Note that an enrichment of VH1-64 was observed, similar to the result of successfully edited donor cells. (**B**) A time course experiment similar to that shown in **Fig. 2A**, except that a less immunogenic KLH-conjugated CRF250-v7ds SOSIP was used. Two mice were sacrificed at the indicated time points following each immunization. B cells isolated from spleen and lymph nodes, or from engrafted mice without immunization, were analyzed by flow cytometry. A representative flow figure is shown for each group. Germinal center (GC) B-cells identified as GL7+/CD38- were gated and quantified (top

panels). These gated cells were then analyzed for the frequency of CD45.1+, indicating they derived from donor mice. The mean percentage of GC- and CD45.1-positive cells is shown from two mice harvested after each indicated immunization is shown. Donor B cells (CD45.1+) were sorted and pooled for sequencing. Bar graphs at the bottom panels show the frequency of HCDR3 mutations present in successfully edited clones. The mutations also observed in **Fig. 3A** are marked with triangles. Note that the first two and last three residues shown derive from murine VH1-family and JH4 segments, respectively, rather than the VRC26.25 HCDR3. (**C**) At the top, nine individual mutations identified from **Fig. 3A** are shown beneath the amino acid sequence of the input VRC26.25 HCDR3. Unmodified VRC26.25 or VRC26.25 variants bearing these mutations were expressed as transmembrane antibodies on 293T cell surface and analyzed by flow cytometry with ConM SOSIP proteins (v8.1ds) modified with the CRF250 or BG505 V1V2 regions, as indicated, conjugated to APC. Cells were co-stained with FITC-labeled anti-human Fc. The ratio between these two signals is shown. (**D**) A figure similar to that in panel **C** except that the indicated combinations of mutations were analyzed. The neutralization profiles of the highest binding variants (NEE, NER, and NERE) are shown in **Fig. 3B** and **C**.

Figure S3. Designs of SOSIP variants, related to Figure 4 (**A**) Linear representations of the indicated SOSIP versions (v5, v7ds, v8.1ds, v8.1 and v8.1mut3ds). Constant and variable regions of Env gp120 and the helical repeat (HR) regions of Env gp41 are indicated, Mutations shown in blue indicate modifications from v5 SOSIP sequence to v7 SOSIP, red indicates modifications from v7 to v8.1, and green indicated modification from v8.1 to v8.1mut3. An additional ‘ds’ disulfide bond is indicated with the extra bracket linking I201C to A433C. The positions of “V1V2” region and the “base” region were labeled with HXB2 numbering. These SOSIP variants were generated using the base from BG505, CRF504, or ConM HIV-1 isolates, which in many cases were modified by a V1V2 region from a different isolate. Thus, 16055-ConM-v8.1 indicates a ConM-v8.1 SOSIP modified with the V1V2 region of the 16055 isolate. (**B**) Sequence of the V1V2 region of Envs used to modified on different bases.

Figure S4. ***In vitro*** **down-selection of SOSIP variants, related to Figure 4 and 5** (**A**) The indicated SOSIP variants were characterized for their ability to bind primary murine

B cells edited to express the VRC26.25 HCDR3. Each variant is labeled according to the HIV-1 isolate contributing its V1V2 region (BG505, CRF250, ConM), the remainder of the engineered Env ectodomain (BG505, ConM), and the SOSIP version used (v5, v7ds, v8.1, v8.1ds). Protein was produced in Expi293F cells or the same cells lacking the acetylglucosaminyltransferase enzyme (GnTI-). Numbers indicate the percentage of SOSIP-binding cells observed in the indicated gate. (**B**) The percentage of SOSIP-binding cells and mean fluorescent intensity of these cells analyzed in panel **B** are represented. (**C**) An SEC-purification profile of Expi293F expressed ConM-v8.1ds SOSIPs conjugated on mi3-60mers using a Sephacryl S-400 HR HiPrep 26/60 column. The elution fractions of the SOSIP-mi3 and unconjugated SOSIP are indicated. (**D**) SDS-PAGE analysis of the indicated conjugated and unconjugated SOSIPs before and after SEC purification, stained by Coomassie blue. The molecular weight of marker is included. ST2 indicates the presence of a SpyTag2 at the SOSIP C-terminus; SPC3 indicates a SpyCatcher3 domain at the N-terminus of the mi3 60-mer.

Figure S5. Optimization of a murine model engrafted with VRC26.25 HCDR3-edited B cells, related to Figure 4 and 5 (**A**) 16 μg of the ConMv8.1ds SOSIP trimers conjugated to KLH, the mi3 60-mer, or as a free trimer (none). (**B**) mi3-conjuates of the 16055-ConM-v8.1ds or BG505-ConM-v8.1ds SOSIP variants were produced in Expi239F cells or the same cell lacking the enzyme GnTI (GnTI-) which adds complex glycans on SOSIP proteins. (**C**) Varying amounts of MPLA adjuvant and protein antigens were tested. Two protein antigens (16055-ConM-v8.1ds and BG505-ConM-v8.1ds) were tested in different ranges of dosage. (**D**) Different *ex vivo* activation methods (anti-CD180, high-dose LPS) of B cells and number of donor B cells transferred to the recipient mouse were tested. For panel **A-D,** immunization efficiency was determined by the neutralizing response in mice engrafted with HCDR3-edited B cells. Neutralization assays used sera harvested one week after the second immunization, with the same schedule shown in **Fig. 1E**. Statistical differences were determined by one-way ANOVA (ns, not significant, *, $p<0.05$ and **, $p<0.01$, and ***, $p<0.001$). (**E**) Mice immunized with 16055-ConM-v8.1ds from **Fig**. **5C** were further characterized. A table summarizing ID_{50} values of pooled sera from neutralization assays against the indicated pseudoviruses derived from tier 2 HIV-1 isolates from the indicated clades. Sera was collected from mice with or without

engraftment of edited B cells one week after each immunization. (**F**) Sera from engrafted mice characterized in panel **E** was further harvested each month at the time points indicated, and the ID_{50} of neutralizing sera against the CRF250 pseudovirus was plotted with days.

Figure S6. Potent immune responses elicited by mRNA-LNP expressing SOSIP-TM variants, related to Figure 6 (**A**) A linear representation of a SOSIP-TM construct. The membrane proximal region (MPER), the transmembrane domain (TM) and a cytoplasmic domain truncated at residue 712 (CT) are added to the SOSIP constructs shown in **Fig. S3A**. SOSIP-TM nomenclature is identical to that used for soluble SOSIP proteins. (**B**) For the experiment in **Fig. 6A**, mice were sacrificed after two immunizations to monitor the *in vivo* response of HIV-1 Env-specific donor B cells. Top panel shows flow cytometry analysis of germinal-center B cells (GL7+/CD38-). Bottom panels show the percentage of CD45.1 donor mouse B cells that recognized the HIV-1 Env trimer [SOSIP(+)] within the germinal center. (**C**) The diameter of LNP particles was monitored by dynamic light scattering to quality control individual LNP preparations. The distribution profile shows mRNA-LNP expressing the indicated SOSIP-TM variants. (**D**) The ability of sera from these mice to bind to 16055-ConM-v8.1ds SOSIP protein was determined by ELISA. Colored bars indicate pooled sera from 7 immunized mice engrafted with HCDR3-edited B cells, and gray bars indicated pooled sera from 3 mice immunized without engraftment with edited B cells. Each sample was assayed with three technical replicates. Neutralization studies of the same sera are shown in **Fig. 6E**. (**E**) Competition ELISA assays were performed to evaluate the specificity of antibody responses. Specifically, the indicated sera analyzed in panel **D** was used to prevent binding of VRC26.25 or a germline form of VRC26.25 to 16055-ConM SOSIP. The percentage of competition was calculated based on the binding of VRC26.25 in the presence of naïve mouse serum. Grey indicates vaccinated mice without engraftment with edited B cells. Significant differences in panels **D-E** were determined by two-way ANOVA (ns, not significant; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001).

Figure S7. A proposed approach for developing an apex-focused HIV-1 vaccine, related to Figure 7 (**A**) An alignment of the HCDR3 sequences of PG9, PG16, VRC26.25 and the VRC26-UCA, each deriving from IGHD3-3, and the open reading frames of four

D3 diversity segments encoding YYDX. Homology with the YYDF motif present in D3-3 are highlighted in green. Other homologies with D3-3 are indicated in grey. Tyrosines shown to be sulfated are indicate in red. (**B**) Previously reported structures of Env trimers complexed with the bnAbs VRC26.25 (PDB: 6VTT) or PG9 (PDB: 5VJ6, (Wang et al., 2017))are presented with only YYDF region of the HCDR3 of these bnAbs shown. Individual Env protomers are indicated in shades of grey, YYDF contact residues on Env are indicated in light blue. Green, red, blue, and yellow indicate YYDF carbon, oxygen, nitrogen, and sulfur atoms, respectively. (**C**) Previously described BCR repertoires from 10 HIV-1 uninfected human donors were analyzed. The percentage of HCDR3 deriving from the indicated D segment is subtracted from the percentage of the same D segment in the population of 'apex-like' HCDR3. A HCDR3 is defined as apex-like if its length is 24 or greater, and it has a tyrosine-sulfation motif 7 or more residues from the HCDR3 amino-terminus and 10 or more residues from its carboxy-terminus. A tyrosine sulfation motif is described as a tyrosine adjacent to an acidic residue without a proximal positive residue. Note that four D3 segments are enriched in the apex-like population. Each of these bears a YYDX motif. (**D**) The frequency of long, sulfated 'apex-like' HCDR3 from each donor is shown, along with subsets with bearing the indicated motifs, and subsets deriving from the indicated D3 gene segments.