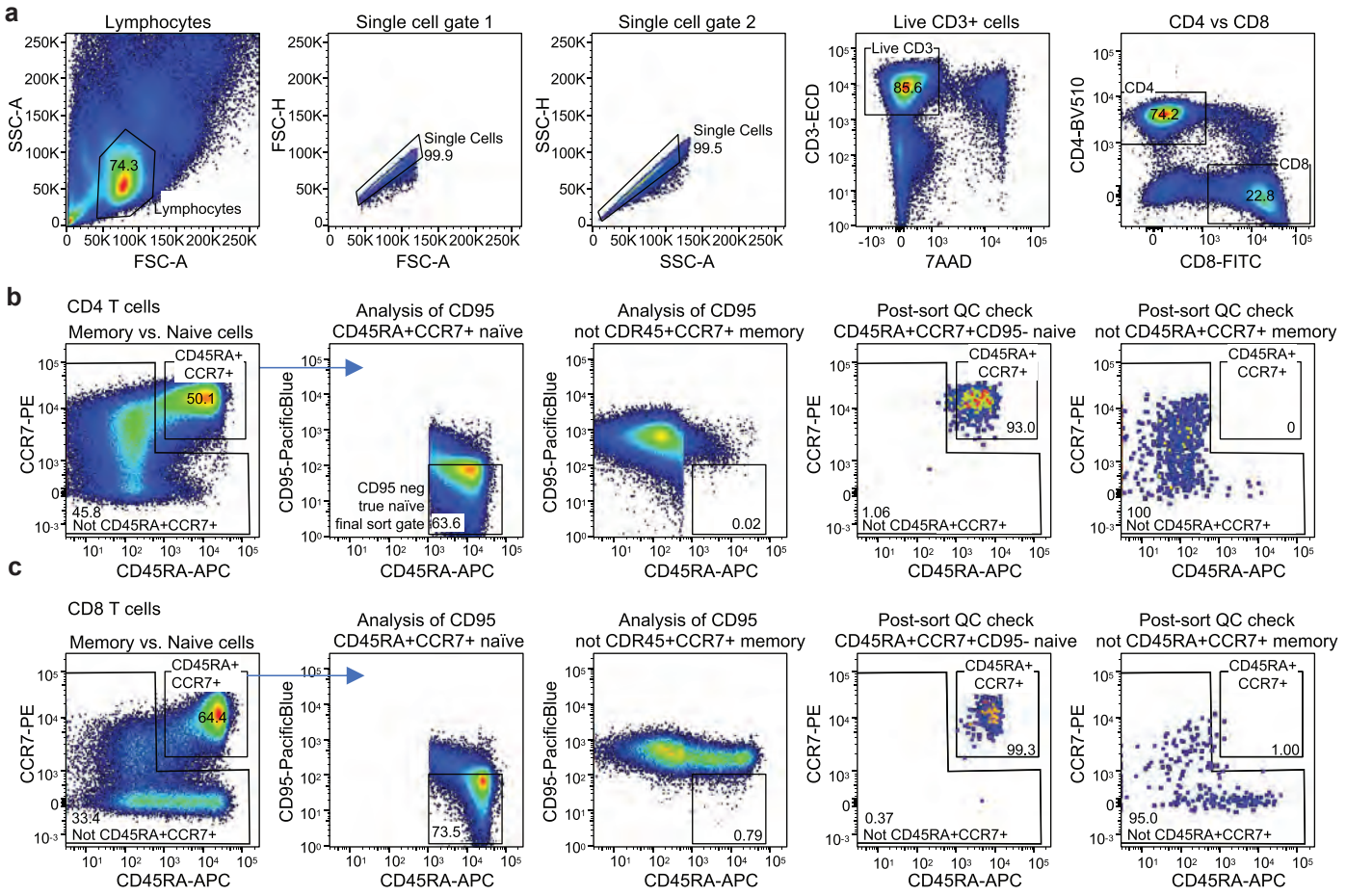
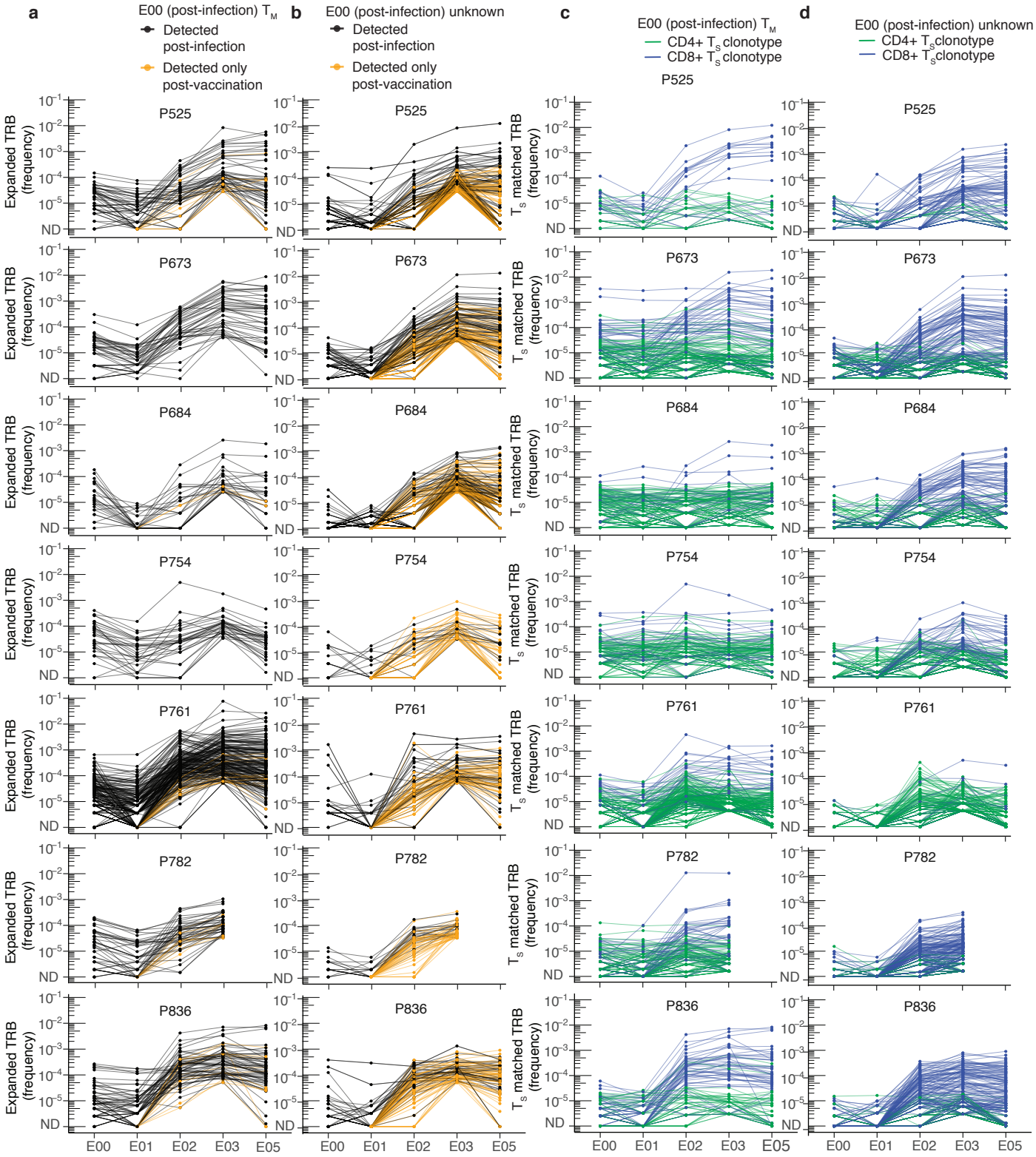
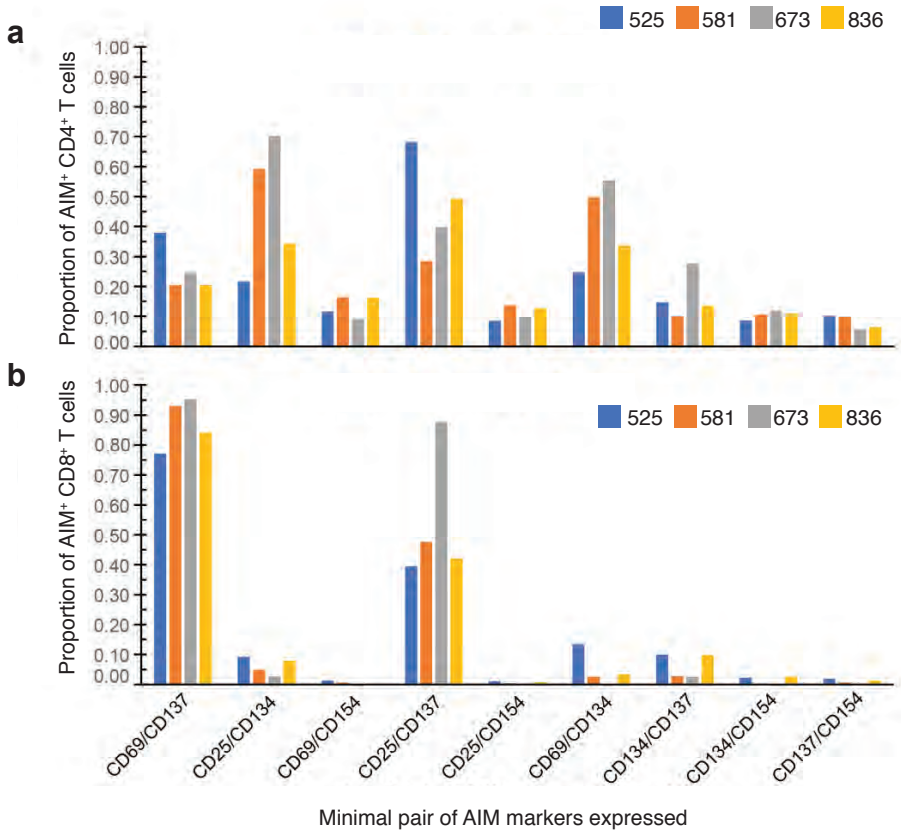


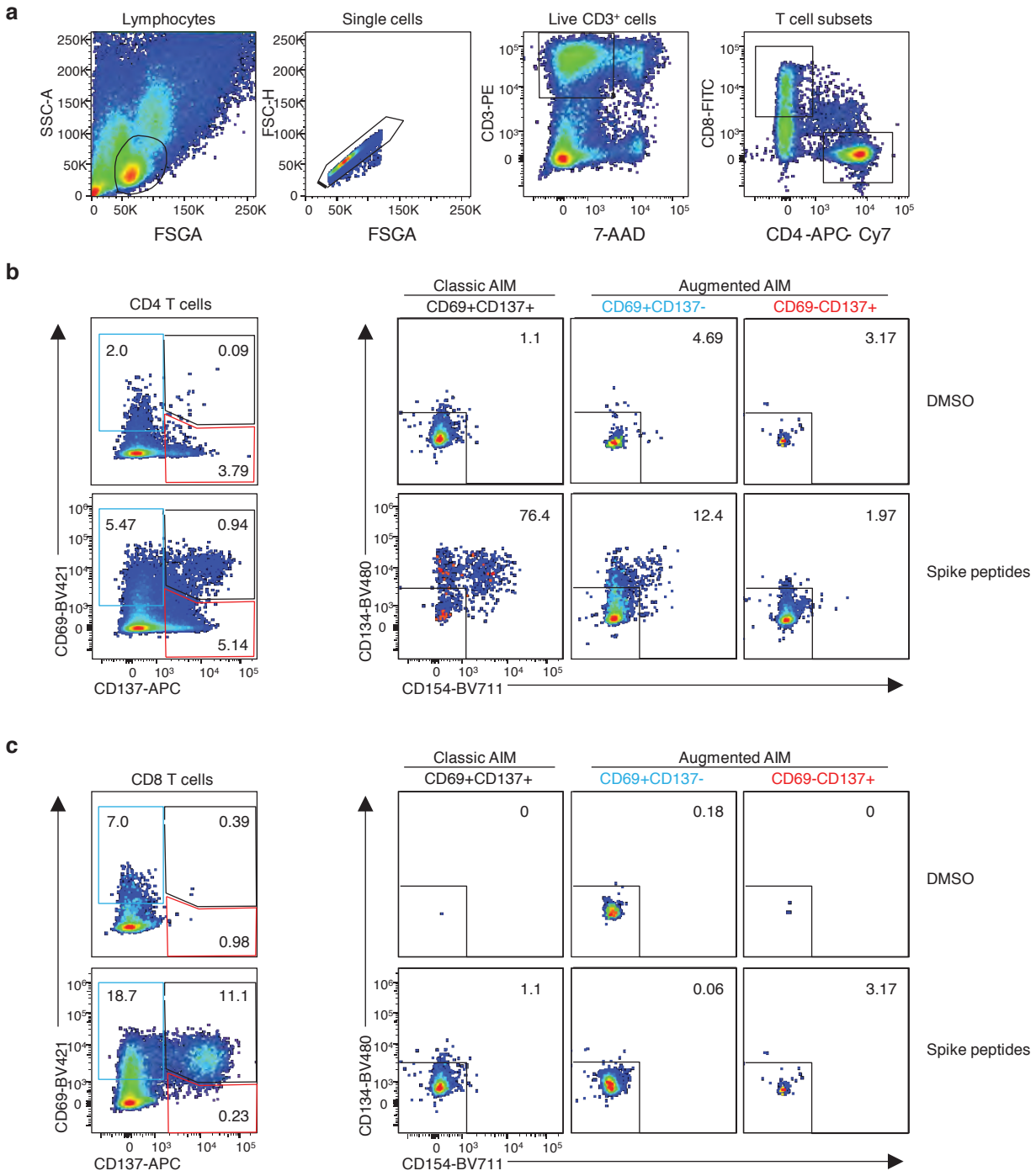
Supplementary Data Fig. 1. (**a**) Representative gating for evaluation of memory/naive phenotype at pre-vaccine timepoint E00. In the top row, PBMC were gated for single live $CD4^{+}CD8^{-}$ or $CD4^{-}CD8^{+}$ T cells. (**b**) Gated $CD4^{+}$ T cells were stained for CD45RA vs. CCR7 to preliminarily identify memory vs. naïve cells. The third dot plot illustrates CD95 expression for memory $CD4^{+}$ T cells. The same indicated gate for $CD95^{-}$ cells was applied to the second dot plot containing preliminarily identified naïve $CD4^{+}$ T cells. Cells within the square gate were retained as final naïve $CD4^{+}$ T cells. (**c**) A similar process was applied to $CD8^{+}$ T cells in the third row. The right two dot plots in (**b**) and (**c**) represent post-sort quality control checks of sorted $CD4^{+}$ and $CD8^{+}$ final naïve and memory cells. Numbers in boxed areas are percent of total cells in indicated gates.



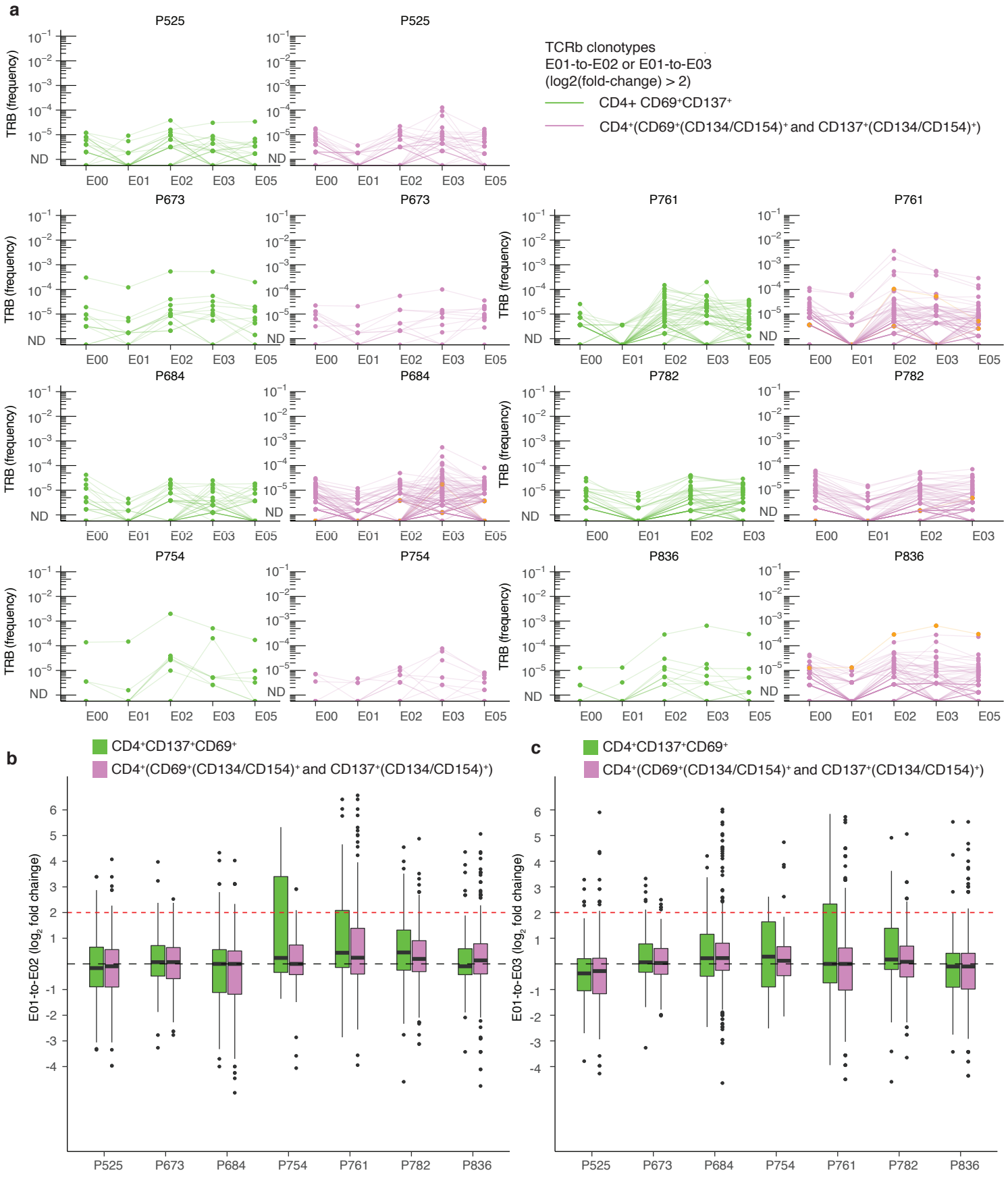
Supplementary Data Fig. 2. (**a**) Trajectories of clonotypes with confirmed pre-vaccine memory or unknown phenotypes among vaccine-expanded and AIM$^+$ S-reactive clonotypes. Deep TRB sequencing was performed after sorting PBMCs into memory and naïve subsets at time point E00 shortly after infection. Clonotypes with a confirmed memory phenotype prior to vaccination that were vaccine-expanded (log2(fold change) > 2 and Fisher's Exact Test q-value < 0.05 between pre-vaccine and post-vaccine repertoires). Black clonotypes were also detected in bulk PBMC repertoires prior to vaccination. Many clones decreased in the year after infection (E00 to E01). A few clonotypes (orange) were not detected prior to initiation of vaccination, but were detected in sorted phenotypic memory cells from E00. (**b**) Vaccine-expanded clones that were not detected amongst memory cells from E00. A few clonotypes were detectable during convalescence (black), but could not be confirmed as memory in (**a**). In general, there are far fewer of such clonotypes per subject than those confirmed as memory in (**a**). Orange clonotypes were not seen in pre-vaccine bulk PBMC, not detected as memory cells at E00, and were only noted after vaccination. (**c**) Spike-activated clonotypes (CD69$^+$CD137$^+$) with a documented memory phenotype at E00 shortly after infection. (**d**) Spike-activated clonotypes that were not detected amongst memory cells at E00. Comparing (**c**) and (**d**), confirmed CD8$^+$ T cell S-reactive clonotypes exhibited stronger vaccine-related expansion than did CD4$^+$ T cell clonotypes, regardless of evidence of memory status.



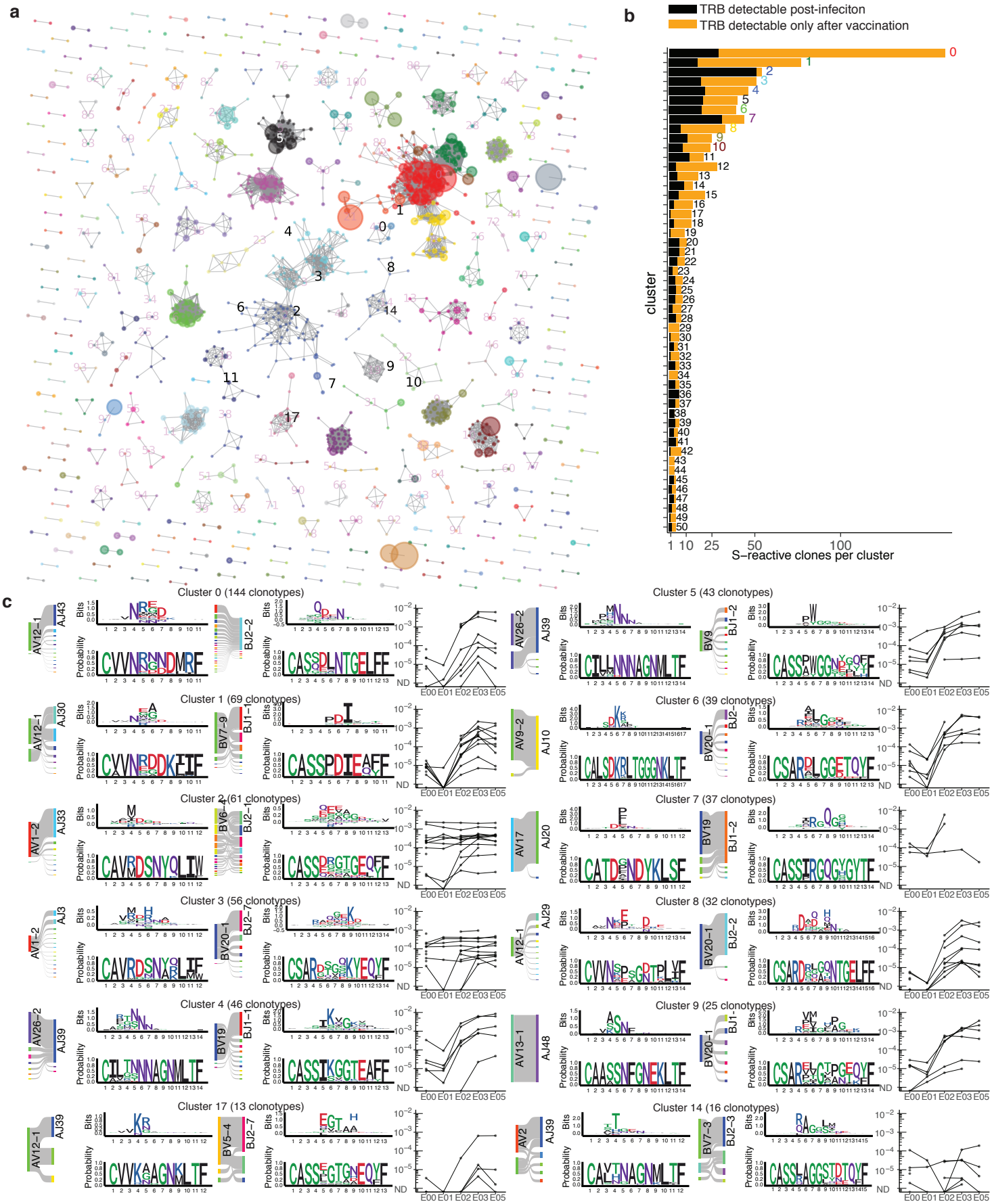
Supplementary Data Fig. 3. Sensitivity of different combinations of activation markers for detection of SARS-CoV-2 S peptide-reactive CD4 and CD8 T cells. E03 PBMC from 4 participants (P525, P581, P673, P836) were gated for **(a)** live, single $CD3^+CD4^+CD8^-$ T cells or **(b)** live, single $CD3^+CD4^-CD8^+$ lymphocytes. Bars are colored by participant. Net percentages of cells expressing each combination of activation induced markers (AIM) were calculated by subtracting values of DMSO-exposed samples from values of S peptides-exposed samples. Total AIM^+ $CD4^+$ and total AIM^+ $CD8^+$, calculated from net percentages, represents the sum of the cells expressing any combination of 2 or more AIM (except for $CD25^+CD69^+$) and is set to 100%. The y-axis shows the proportion of total $CD4^+$ or $CD8^+$ AIM^+ cells detected per AIM pair indicated on the x-axis. The number of $CD4^+$ T cells analyzed per condition ranged from 10,190 to 95,833 for DMSO and from 27,368 to 112,383 for S peptide pool. Bars represent results from single flow cytometry experiments.



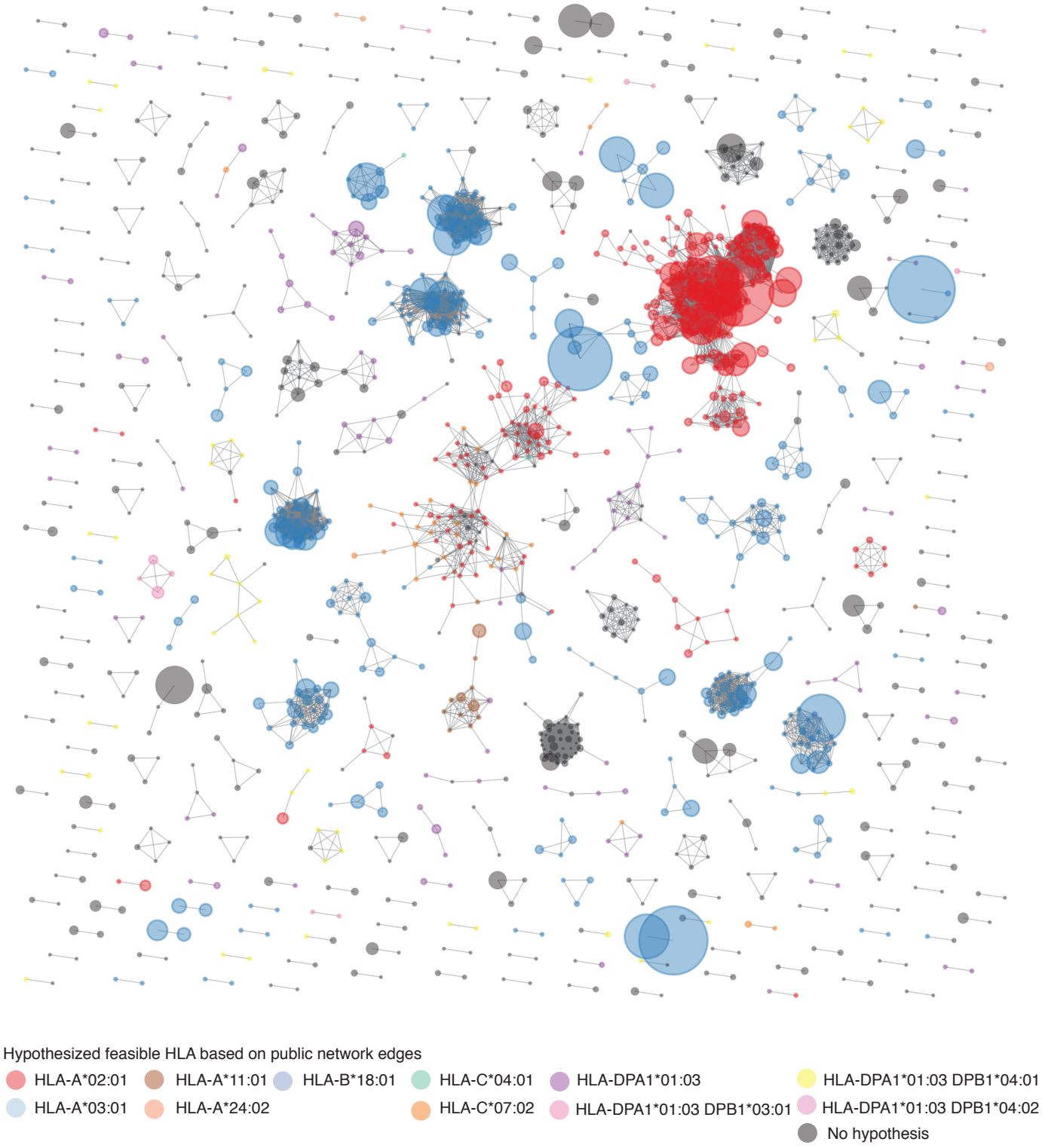
Supplementary Data Fig. 4. Representative gating scheme for sorting of CD69$^+$ CD137$^+$ and CD69$^+$ or CD137$^+$ plus CD134$^+$ and/or CD154$^+$ activation induced marker (AIM) cell fractions to investigate AIM identification of S-reactive clonotypes via bulk TCRβ-seq. PBMC from E03 were stimulated with S peptides or DMSO for 18 hours and stained for expression of the indicated markers. (**a**) Preliminary gating for live CD4$^+$ and CD8$^+$ T cells. (**b, c**) show the same scheme in CD4$^+$ **(b)** and CD8$^+$ (**c**) T cells, respectively. Left dot plots show expression of CD69$^+$ and CD137$^+$ CD4$^+$ and CD8$^+$ T cells in the CD69$^+$ CD137$^+$ double positive (black) gate, were collected for CD4$^+$ TCRβ-seq as 'classic' AIM (black gate). The right three columns of dot plots show expression of additional activation markers for cells in the indicated quadrants in the CD69/CD137 plots. CD4$^+$ T cells expressing either CD134 and/or CD154 from the CD69$^+$ CD137$^-$ (blue) and CD69$^-$ CD137$^+$ (red) gates were pooled for CD4$^+$ TCRβ-seq as 'augmented' AIM$^+$ cells. In this representative person, 76.4% of CD69$^+$ CD137$^+$ CD4+ T cells were positive for CD134 and/or CD154. However, 12.4% of CD69$^+$ CD137$^-$ CD4$^+$ T cells and 2.0% of CD69$^-$ CD137$^+$ CD4 T cells were positive for these additional markers. Few CD8$^+$ T cells from the CD69$^+$ CD137$^-$ and CD69$^-$ CD137$^+$ gates expressed either CD134 and/or CD154 and so these were not sequenced by TCRβ-seq.



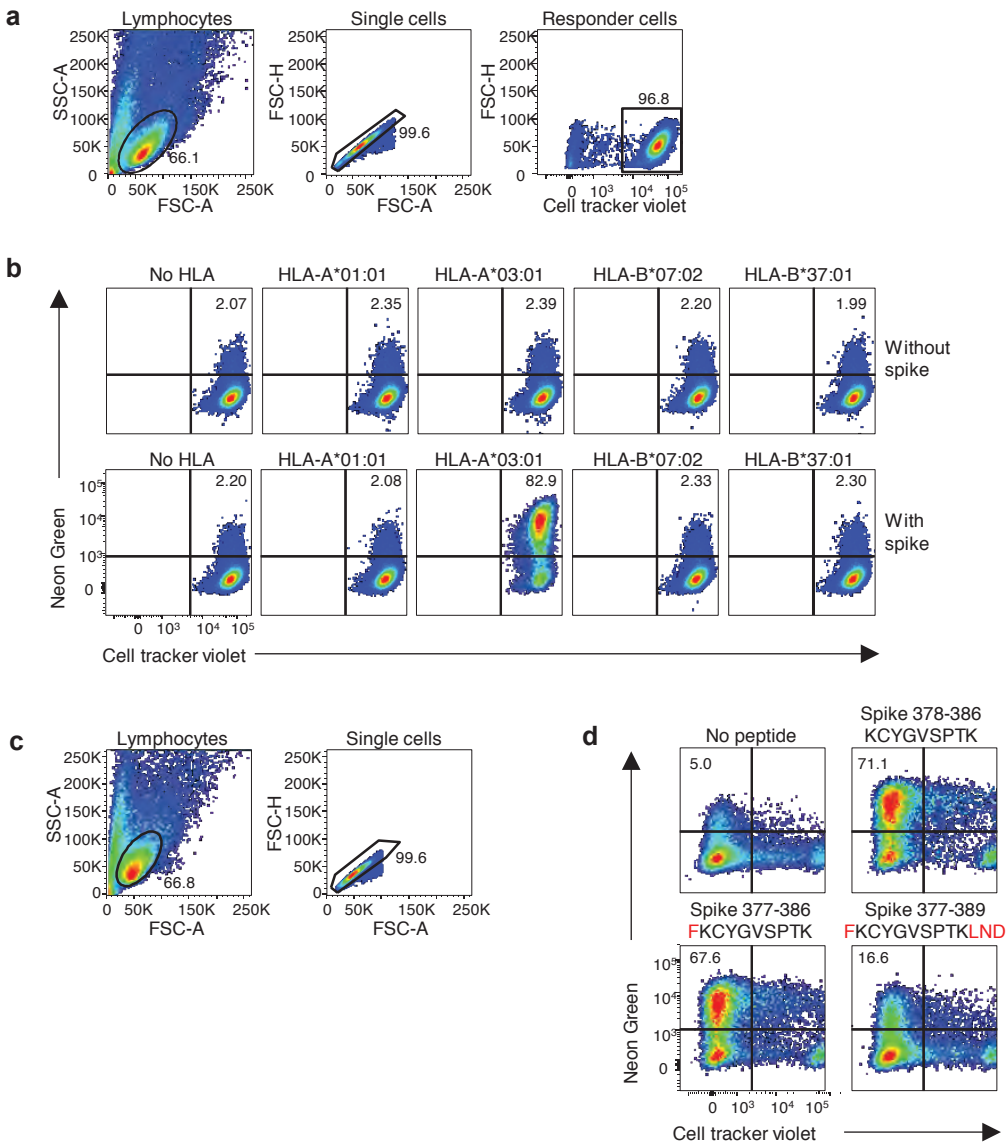
Supplementary Data Fig. 5. Longitudinal kinetics and fold change distribution of AIM-scTCRαβ-seq selected T cells comparing CD69, CD137, and CD134/ CD154 marker sets in 7 persons. (**a**) The frequency of each expanding (log_2(FC) > 2) TRB clonotype is tracked from E00 to E05. Clonotypes shown were those enriched in the total CD4+ sort and enriched in subsequent split sample sorts for $CD69^+CD137^+$ (green) or augmented CD69/137 single positive $CD134^+$ and/or $CD154^+$ (pink, 'augmented'), or both (orange). Among all TS cells identified using either marker set, 12-49% (median 27%) TRB clonotypes were $CD69^+CD137^+$ double positive and 48-86% (median 69%) were CD69/CD137 single positive plus $CD134^+$ and/or $CD154^+$. A small fraction of TRB sequences, 2-12% (median 4%) were identified in both AIM+ populations. **(b, c)** Distribution of fold change of all AIM^+ $CD4^+$ TRB clonotypes between pre-vaccine E01 and E02 (b) and E01 and E03 (c). Median, box (IQR), and whiskers (1.5 * IQR) indicate distribution of fold changes. Clonotypes falling above the red line in either (**b**) or (**c**) are those with expansion of log_2(fold-change) > 2 relative to E01.



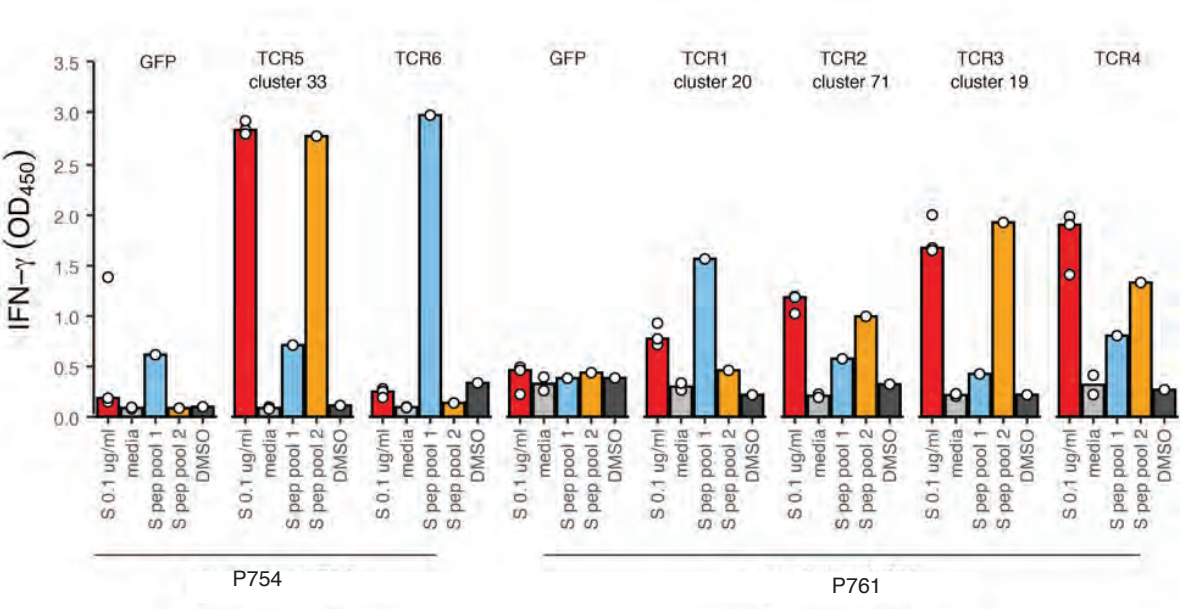
Supplementary Data Fig. 6. Selected TS clonotype clusters by CD69+CD137+ AIM-scTCRαβ-seq. (**a**) Sequence similarity graph with 1,448 clones and 248 convergent clusters of two or more TS clonotypes recovered from 17 individuals at E03. Circle size represents the relative frequency of each TCRαβ clonotype. Equivalent TCRαβ AA sequences may be included more than once if found in multiple participants. Color is by cluster number. (**b**) The number of clones comprising each cluster that had a corresponding TRB detectable post-infection (black) or only post-vaccination (orange) for the top 50 clusters. (**c**) Logo plots for representative clusters indicated by integers in (**a**) with inferred restricting HLA class I alleles (Methods), if able to be defined. For each CDR3 motif, the lower sequence logo shows the probability of each amino acid residue at each CDR3 position and the upper sequence logo depicts the information content in bits comparing the residue usage to a set of randomly selected CDR3 with the same V and J gene usage as the sequence cluster (Methods). Residue color indicates chemistry: acidic (red), basic (blue), hydrophobic (black), neutral (purple), polar (green). Line graphs show the sums of the abundances of the TRB sequences in each cluster in longitudinal PBMC repertoires for each participant contributing sequences to the cluster. Clusters 2 and 3 have many members with characteristic MAIT cell V-J gene usage (*TRAV1-2* with *TRAJ33*, *TRAJ20*, or *TRAJ12*).



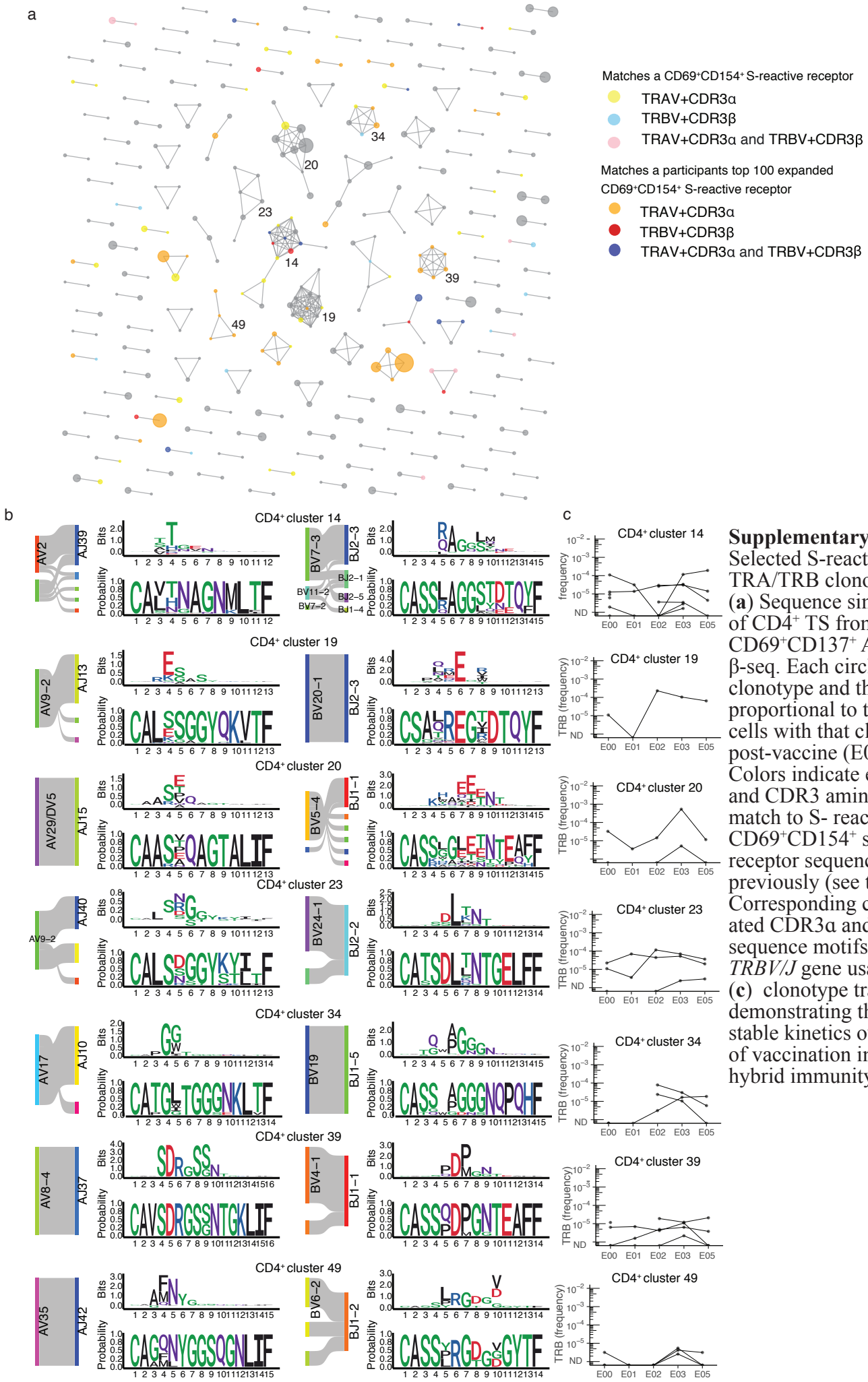
Supplementary Data Fig. 7. Feasible cognate HLA alleles (Feasible indicating clusters formed from groups of persons expressing a shared HLA class I or class II allele, suggesting common specificity for a peptide ligand restricted by this allele.) T cells co-expressing CD69 and CD137 after 18 hour stimulation with S peptides we e sorted and TRA/TRB sequences recovered by AIM-scTCRαβ-seq (see Methods). Sequence similarity graph of paired TRA/TRB sequences with 1,448clones and 248 convergent clusters of two or more S-reactive clones recovered from 17 individuals following second immunization. HLA feasibility analysis (see Methods) was used to assess feasible shared HLA allele among participants contributing TCRs to convergent sequence clusters. U indicates the feasible allele could not be narrowed down to two or fewer candidates (see Methods).



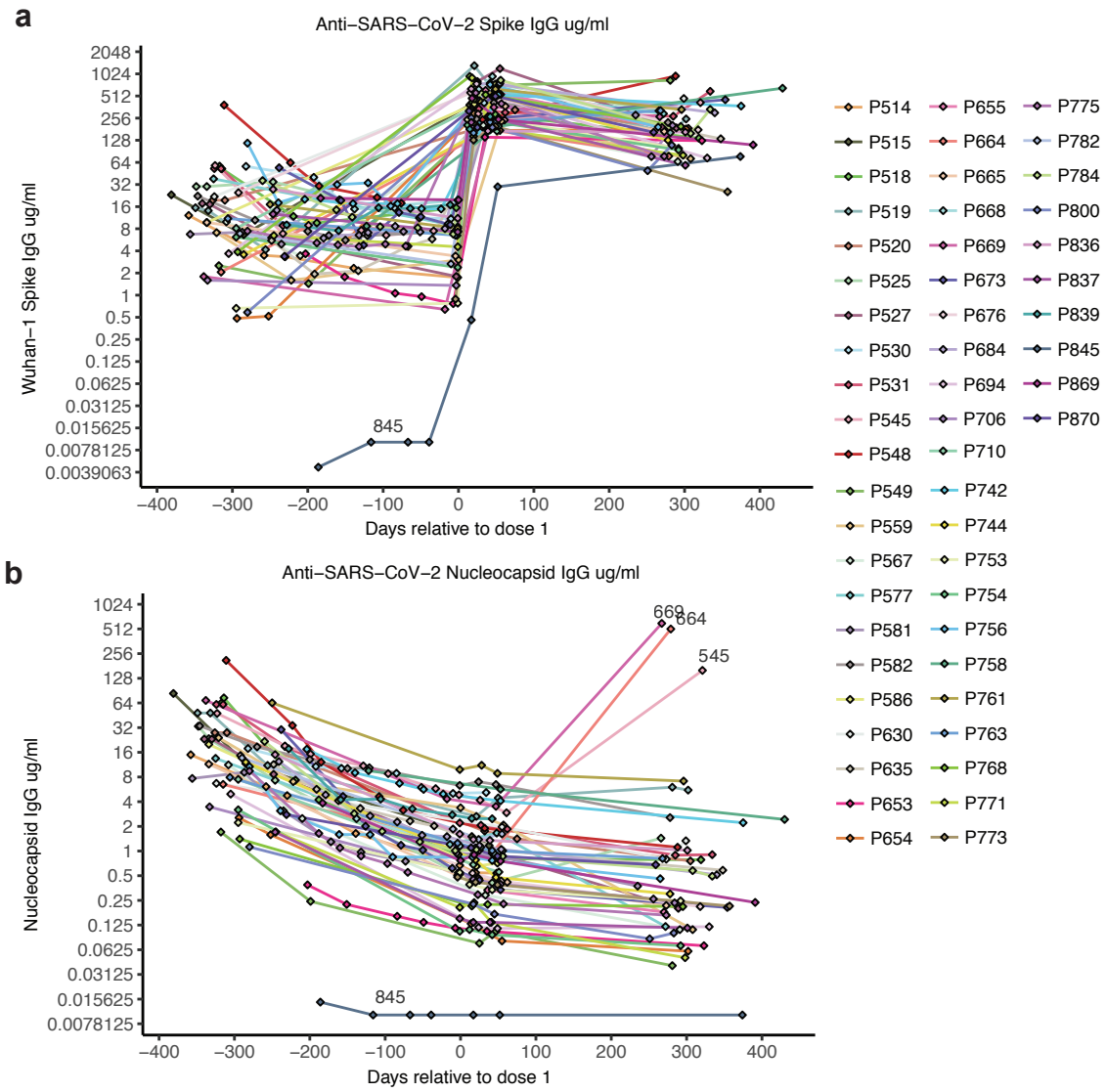
Supplementary Data Fig. 8. Evaluation of specificity of cluster-identified TCR for SARS-CoV-2 S. Representa-tive example shows evaluation of Spike-specificity of TCR3 from subject P673. (**a**) Representative gating strategy shows selection of CTV- labeled TCR3-transduced reporter cells after incubation with artificial antigen presenting cells expressing SARS-CoV-2 spike and each subject-specific HLA-A or HLA-B. (**b**) T cell activation was determined by measuring the percentage of mNeonGreen expression on TCR3-transduced reporter cells for each of the indicated HLA allelic variants. (**c**) Gating scheme for peptide specificity workup with EBV-LCL used as antigen presenting cells. (**d**) Cell tracker violet stained EBV-LCL were de-gated, and percent mNeonGreen expression determined on TCR3-transduced reporter cells. Amino acids lateral to the core 9 amino acid antigenic peptide recognized by TCR3 are shown in red. Numbers are percent of cells in indicated gates.



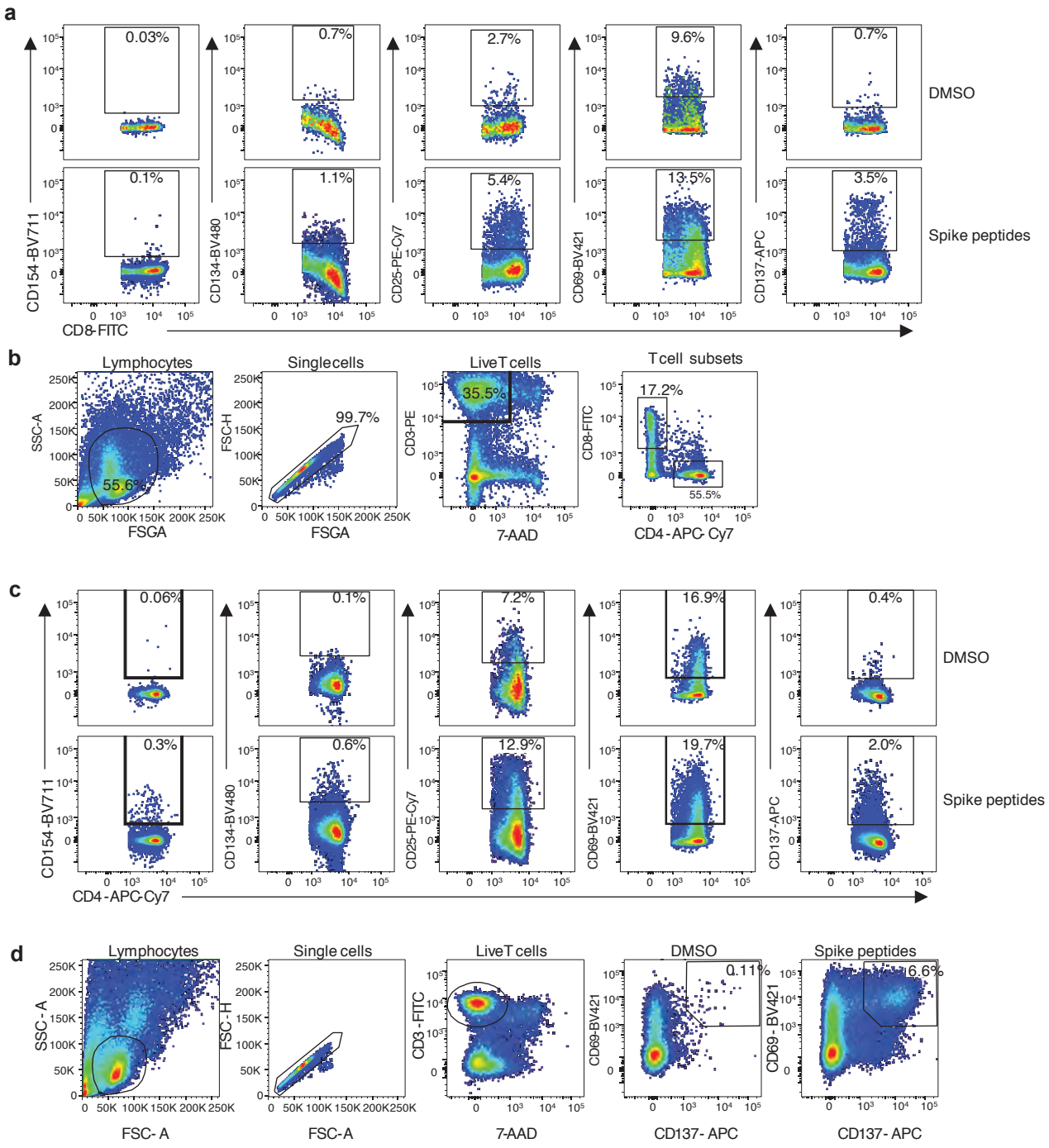
Supplementary Data Fig. 9. Confirmation of S-specificity of CD4+ TCRs identified by AIM-scTCRα β-seq. IFNγ response of reporter T cells with engineered TCRs selected from public clusters after stimulation with 0.1μg/ml whole S protein (red) or overlapping 15-mer peptides (1μg/ml each) spanning S protein amino terminal domain S1 (light blue) and carboxy terminal domain S2 (orange). The height of each bar represents the mean of three replicates (individual dots are shown). The negative control level of activation of $CD4^+$ T cells lentivirus-transduced with the control GFP gene is also shown.



Supplementary Data Fig. 10. Selected S-reactive CD4$^+$ TRA/TRB clonotype clusters. (**a**) Sequence similarity graph of CD4$^+$ TS from CD69$^+$CD137$^+$ AIM-scTCRαβ-seq. Each circle is a unique clonotype and the size is proportional to the number of cells with that clonotype at the post-vaccine (E03) time point. Colors indicate exact V-gene and CDR3 amino acid-level match to S- reactive CD69$^+$CD154$^+$ single-chain receptor sequences identified previously (see text). (**b**) Corresponding cluster-associated CDR3α and CDR3β sequence motifs, *TRAV/J* and *TRBV/J* gene usage plots, and (**c**) clonotype trajectories demonstrating the observed stable kinetics over the course of vaccination in persons with hybrid immunity.



Supplemental Data Fig. 11. Serial anti-SARS-CoV-2 plasma IgG levels over time from convalescence to vaccine booster. (**a**) Anti-spike and (**b**) anti-nucleocapsid antibodies. Day 0 on X axis is date of first mRNA vaccination. Participant P845, who was seronegative for both anti-nucleocapsid and anti-S at study entry and seroconverted (anti-S only) with vaccination, is shown. Participants with breakthrough infections (P545, P664, P669) are indicated by increase in anti-nucleocapsid antibodies between E03 and E05. Participant numbers and color codes for both graphs are provided.



Supplementary Data Fig. 12 | Gating scheme for comparison of AIM candidates. (**a**) Representative results for person with hybrid immunity after stimulation with DMSO negative control or SARS-CoV-2 spike peptide pool for $CD3^+$ T cells expressing activation markers CD69 and CD137. (**b**) Gating on single live $CD4^+$ or $CD8^+$ cells. (**c, d**) Representative results of gated $CD4^+$ T cells (**c**) or $CD8^+$ T cells (**d**) after stimulation with DMSO negative control (top) or S peptides (bottom) for five indicated candidate activation markers. Numbers are percent of $CD4^+$ or $CD8^+$ T cells in indicated gates.