JEM Journal of Experimental Medicine

ARTICLE

Identification of conserved SARS-CoV-2 spike epitopes that expand public cTfh clonotypes in mild COVID-19 patients

Xiuyuan Lu[1], Yuki Hosono[1,2,3], Masamichi Nagae[1,2], Shigenari Ishizuka[1,2], Eri Ishikawa[1,2], Daisuke Motooka[4], Yuki Ozaki[4], Nicolas Sax[5], Yuichi Maeda[3,6], Yasuhiro Kato[3,7], Takayoshi Morita[3,7], Ryo Shinnakasu[8], Takeshi Inoue[8], Taishi Onodera[9], Takayuki Matsumura[9], Masaharu Shinkai[10], Takashi Sato[10], Shota Nakamura[4], Shunsuke Mori[11,12], Teru Kanda[13], Emi E. Nakayama[14], Tatsuo Shioda[14,15], Tomohiro Kurosaki[8,15,16], Kiyoshi Takeda[6,15,17], Atsushi Kumanogoh[3,7,15], Hisashi Arase[11,12,15], Hironori Nakagami[18], Kazuo Yamashita[5], Yoshimasa Takahashi[9], and Sho Yamasaki[1,2,15,19,20]

Adaptive immunity is a fundamental component in controlling COVID-19. In this process, follicular helper T (Tfh) cells are a subset of CD4⁺ T cells that mediate the production of protective antibodies; however, the SARS-CoV-2 epitopes activating Tfh cells are not well characterized. Here, we identified and crystallized TCRs of public circulating Tfh (cTfh) clonotypes that are expanded in patients who have recovered from mild symptoms. These public clonotypes recognized the SARS-CoV-2 spike (S) epitopes conserved across emerging variants. The epitope of the most prevalent cTfh clonotype, $S_{864-882}$, was presented by multiple HLAs and activated T cells in most healthy donors, suggesting that this S region is a universal T cell epitope useful for booster antigen. SARS-CoV-2–specific public cTfh clonotypes also cross-reacted with specific commensal bacteria. In this study, we identified conserved SARS-CoV-2 S epitopes that activate public cTfh clonotypes associated with mild symptoms.

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused a worldwide outbreak of coronavirus disease 2019 (COVID-19). B cells (antibodies) and cytotoxic (CD8⁺) and helper (CD4⁺) T cells are the fundamental components of adaptive immunity, the process by which the body attempts to control COVID-19 (Ni et al., 2020; Rydyznski Moderbacher et al., 2020). B cells produce neutralizing antibodies, and their critical epitopes are determined within the spike (S) protein (Liu et al., 2020); these epitopes have contributed to the development of current vaccines that employ S protein as an antigen (Baden et al., 2021; Walsh et al., 2020; Sadoff et al., 2021; Voysey et al., 2021; Shinde et al., 2021). In contrast, key T cell epitopes that contribute to protective immunity against COVID-19 have not been well characterized. In addition to the importance of CD8⁺ T cells (Sette and Crotty, 2021), rapid induction of CD4⁺ T cells is associated with mild COVID-19 symptoms (Peng et al., 2020; Tan et al., 2021; Sette and Crotty, 2021). Among CD4⁺ T cells, follicular helper T (Tfh) cells help B cells to produce protective antibodies (Liu et al., 2013; Vinuesa et al., 2016; Crotty, 2019), and inefficient induction of Tfh cells is correlated with more severe and fatal COVID-19 (Kaneko et al., 2020; Zhang et al., 2021; Gong et al., 2020). Although TCR clonotypes for SARS-CoV-2 are being extensively examined (Shomuradova et al., 2020; Dykema et al., 2021), public Tfh clonotypes

[1]Laboratory of Molecular Immunology, Immunology Frontier Research Center, Osaka University, Suita, Japan; [2]Department of Molecular Immunology, Research Institute for Microbial Diseases, Osaka University, Suita, Japan; [3]Department of Respiratory Medicine and Clinical Immunology, Graduate School of Medicine, Osaka University, Suita, Japan; [4]Department of Infection Metagenomics, Genome Information Research Center, Research Institute for Microbial Diseases, Osaka University, Suita, Japan; [5]KOTAI Biotechnologies, Inc., Suita, Japan; [6]Laboratory of Immune Regulation, Department of Microbiology and Immunology, Graduate School of Medicine, Osaka University, Suita, Japan; [7]Department of Immunopathology, Immunology Frontier Research Center, Osaka University, Suita, Japan; [8]Laboratory of Lymphocyte Differentiation, Immunology Frontier Research Center, Osaka University, Suita, Japan; [9]Research Center for Drug and Vaccine Development, National Institute of Infectious Diseases, Tokyo, Japan; [10]Tokyo Shinagawa Hospital, Tokyo, Japan; [11]Department of Immunochemistry, Research Institute for Microbial Diseases, Osaka University, Suita, Japan; [12]Laboratory of Immunochemistry, Immunology Frontier Research Center, Osaka University, Suita, Japan; [13]Division of Microbiology, Faculty of Medicine, Tohoku Medical and Pharmaceutical University, Sendai, Japan; [14]Department of Viral Infections, Research Institute for Microbial Diseases, Osaka University, Suita, Japan; [15]Center for Infectious Disease Education and Research, Osaka University, Suita, Japan; [16]Laboratory of Lymphocyte Differentiation, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan; [17]Department of Mucosal Immunology, Immunology Frontier Research Center, Osaka University, Suita, Japan; [18]Department of Health Development and Medicine, Graduate School of Medicine, Osaka University, Suita, Japan; [19]Division of Molecular Immunology, Medical Mycology Research Center, Chiba University, Chiba, Japan; [20]Division of Molecular Design, Research Center for Systems Immunology, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan.

Correspondence to Sho Yamasaki: yamasaki@biken.osaka-u.ac.jp.

associated with COVID-19 recovery have not yet been reported. The identification of such public clonotypes and their common epitopes would elucidate the S regions essential for protective T cell immunity and thereby provide invaluable information for future vaccine design against emerging variants. Furthermore, any cross-reactive antigens of such clonotypes may contribute to the severity of COVID-19.

In this study, using single-cell–based paired and bulk TCR sequencing (TCR-seq) of COVID-19 patient T cells and the worldwide TCR database, we identified public circulating Tfh (cTfh) clonotypes associated with mild symptoms and their epitopes in the S regions. These epitopes are conserved among the emerging SARS-CoV-2 variants.

Results

Identification of SARS-CoV-2 S protein–specific cTfh clones in convalescent COVID-19 patients

We first analyzed SARS-CoV-2–specific T cell subsets and their clonotypes using a single-cell–based RNA-seq platform. Peripheral blood mononuclear cells (PBMCs) isolated from healthy donors and convalescent COVID-19 patients (Table S1 and Table S2) were stimulated with antigens derived from SARS-CoV-2, including inactivated virus, recombinant S protein, overlapped peptide pools derived from S protein (S peptide pool) or membrane (M) and nucleocapsid (N) proteins (MN peptide pool). Activation marker–positive T cells were sorted and analyzed for their TCR sequences together with RNA expression by single-cell TCR-seq and RNA-seq analyses (Fig. S1 A). Uniform manifold approximation and projection (UMAP) embedding and clustering allowed identification of a cluster (#8) consisting of $CD4^+$ cells expressing Tfh-related genes, such as *CD200*, *PDCD1*, *ICOS*, *CXCL13*, *CD40LG*, and *CXCR5* (Fig. 1, A and B; Crotty, 2019; Schmitt et al., 2014), and activation signature genes (Fig. S1 B). Indeed, cluster #8 also exhibited a high Tfh score based on the reported gene sets (Meckiff et al., 2020), suggesting that it includes cTfh cells (Fig. 1, C and D). Within this cluster, we could identify 1,735 TCRs (Table S3). Among 147 TCRs from the first batch of analyses, which included five patients (Ts-002, -016, -017, -018, and -020) and one healthy donor (Oo-001), we detected S-reactive clones bearing TCRαβ pairs shared between patients Ts-017 and Ts-018, who exhibited anti-S and neutralizing antibodies (Table S1). Considering the cell number of single-cell TCR-seq, these are likely to be frequent TCRs, and we designated them as TCR-017 and TCR-018, respectively. TCR-017 and -018 had the same Vα, Jα, Vβ, and Jβ usages, and their complementarity-determining region 3 (CDR3) sequences were identical except for one amino acid in CDR3β (Fig. 1, E and F). TCR-018 was detected in patient Ts-018 as three different barcoded clones (Table S3).

TCR-017 and -018 recognize an identical S epitope restricted by the same HLAs

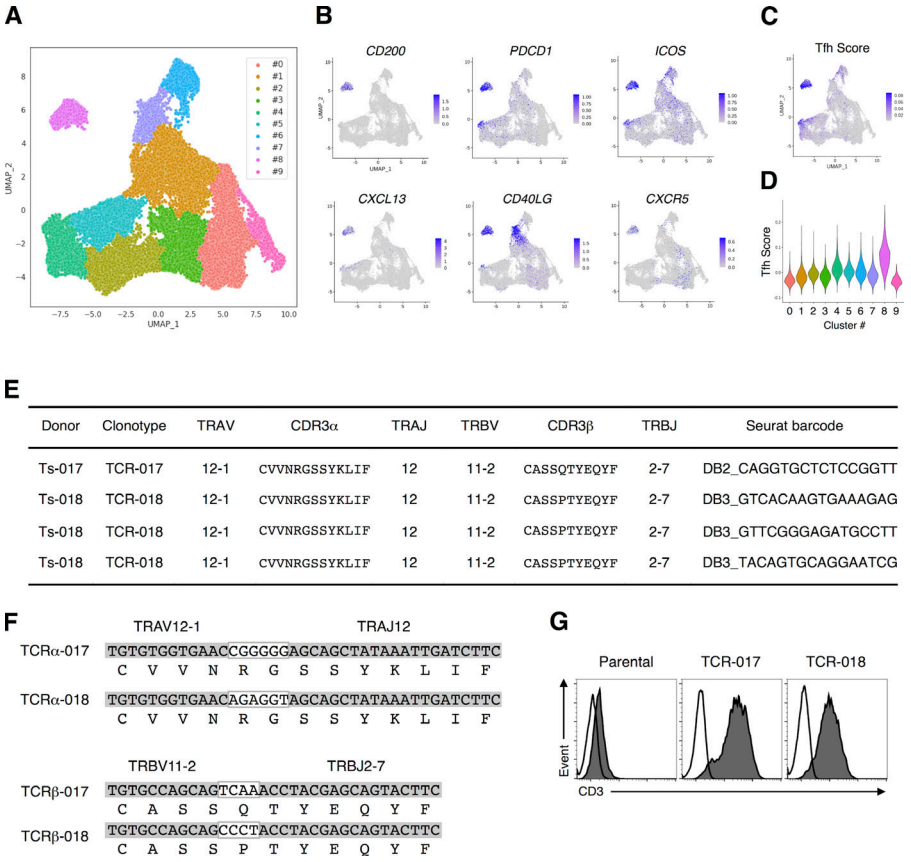
To examine the antigen specificity of TCR-017 and -018, the respective TCRα and β chains were reconstituted in a TCR-deficient T cell hybridoma bearing an NFAT-GFP reporter (Fig. 1 G). TCR transfectants were stimulated with SARS-CoV-2 antigens in the presence of transformed B cells derived from the same patients as autologous APCs. Cells expressing TCR-017 and -018 responded to recombinant S protein and the S peptide pool but not to the MN peptide pool (Fig. 2 A), which is consistent with the initial antigen specificity revealed in single-cell analysis (Table S3). Among two halves of pooled S peptides, #1 ($S_{1–643}$) and #2 ($S_{633–1273}$), only #2 pool reacted with both TCRs (Fig. 2 B). However, another SARS-CoV-2 S peptide pool ($S_{304–338,\ 421–475,\ 492–519,\ 683–707,\ 741–770,\ 785–802,\ 885–1273}$; termed selected S pool) did not activate either clone (Fig. S2 A and Fig. 2 B). These results suggest that the antigen epitopes for TCR-017 and -018 are located in regions 633–682, 708–740, 771–784, or 803–884 of the S protein (Fig. S2 A).

To further reduce the number of epitope candidates, we determined the HLA alleles that restrict TCR-017 and -018. Heterologous APCs from patient Ts-017 could activate TCR-018, and vice versa, demonstrating that these APCs are interchangeable (Fig. S2 B). Thus, the shared HLA class II alleles, DRA-DRB1*15:01, DQA1*01:02-DQB1*06:02, or DPA1*02:02-DPB1*05:01 (shown in red in Table S2), are considered to restrict this recognition. Since DRB1*15:01 and DQB1*06:02 are in linkage disequilibrium and would not be segregated from each other (Begovich et al., 1992), these HLA alleles were separately transfected into HEK293T cells to examine their ability to present S peptides to TCR-017 and -018. Only cells transfected with DRA and DRB1*15:01 could activate these TCRs in the presence of S peptides (Fig. 2 C). Using NetMHC Server software (Reynisson et al., 2020), two epitopes, $S_{828–846}$ and $S_{864–882}$, were predicted as strong binders to DRA-DRB1*15:01 within the candidate S regions described above (Fig. S2 A). Among them, $S_{864–882}$ was ultimately identified as the epitope for both TCR-017 and -018 (Fig. 2 D). Judging from their dose–response curves, the relative affinities of TCR-017 and -018 to this peptide appeared comparable (Fig. 2 E). Furthermore, serial overlapping peptides determined $S_{870–878}$ as a minimum epitope for both TCRs (Fig. 2 F). Given the same epitope on the same HLA with similar affinities, TCR-017 and -018 are considered as one clonotype, clonotype-017/018, shared by different patients. As would be expected from their similarity to DRB1*15:01, DRB1*15:02/*15:03/*15:04 also present this epitope and activate clonotype-017/018 (Fig. S2 C).

Public clonotypes expanded in recovered COVID-19 patients preferentially recognize conserved S epitopes

Consistent with the high frequency of DRB1*15, which is shared by 33% of the Japanese population (http://hla.or.jp), bulk TCR-seq of prepandemic healthy Japanese donors revealed that TCRα and β chains of clonotype-017/018 are prevalent (Fig. 2 G). As DRB1*15 is not a minor allele worldwide (Gonzalez-Galarza et al., 2020), we analyzed TCR databases of healthy and convalescent COVID-19 cohorts containing donors of various ethnicities (Emerson et al., 2017; Nolan et al., 2020 *Preprint*) and found that approximately one-fifth of donors possessed TCRβ-017/018 (Fig. 2 H). These results suggest that clonotype-017/018 is a public clonotype.

TCR-017/018 is not an exceptional public clonotype because 855 of 1,735 TCRs identified in our samples were also detected in these cohorts (Table S3). Among them, we identified 10 public cTfh clonotypes that were significantly expanded in recovered COVID-19 patients compared with the healthy cohort (Fig. 3 A,



E

Donor	Clonotype	TRAV	CDR3α	TRAJ	TRBV	CDR3β	TRBJ	Seurat barcode
Ts-017	TCR-017	12-1	CVVNRGSSYKLIF	12	11-2	CASSQTYEQYF	2-7	DB2_CAGGTGCTCTCCGGTT
Ts-018	TCR-018	12-1	CVVNRGSSYKLIF	12	11-2	CASSPTYEQYF	2-7	DB3_GTCACAAGTGAAAGAG
Ts-018	TCR-018	12-1	CVVNRGSSYKLIF	12	11-2	CASSPTYEQYF	2-7	DB3_GTTCGGGAGATGCCTT
Ts-018	TCR-018	12-1	CVVNRGSSYKLIF	12	11-2	CASSPTYEQYF	2-7	DB3_TACAGTGCAGGAATCG

Figure 1. **TCRs shared by SARS-CoV-2–reactive Tfh cells from different patients.** **(A)** UMAP projection of T cells in single-cell analysis of PBMCs identified a cluster representing Tfh cells (cluster #8). Each dot corresponds to a single cell and is colored according to the cluster. **(B)** The expression levels of the canonical Tfh cell markers *CD200*, *PDCD1*, *ICOS*, *CXCL13*, *CD40LG*, and *CXCR5* are shown as heatmaps in the UMAP plot. **(C)** Tfh scores of T cells in single-cell analysis of PBMCs are shown as heatmaps in the UMAP plot. **(D)** Tfh score of each cluster determined in A is shown in a violin plot. **(E)** Usages of V and J gene segments and CDR3 sequences of the α (TRAV, TRAJ, and CDR3α, respectively) and β chains of TCR-017 and TCR-018. For TCR-018, three different barcoded clones having the same nucleotide sequences detected in Ts-018 are shown. **(F)** Alignment of CDR3α and CDR3β nucleotide and amino acid sequences from TCR-017 and TCR-018. Nucleotide sequences within the white boxes indicate N-nucleotides. **(G)** Surface expression of paired TCRs on reporter cells are assessed by surface anti-CD3 staining. Filled histogram, anti-mouse CD3 mAb; open histogram, isotype control antibody. Data are representative of three independent experiments (G).

column J, asterisks). Of note, one-half of such clonotypes were indeed detected in multiple patients in our sample pool by single-cell TCR-seq (Fig. 3 A, column A, #), and TCR-017/018 was the fifth most expanded clonotype (Fig. 3 A, columns A and J).

To determine the epitopes of these clonotypes more efficiently, we established a rapid epitope determination platform (Fig. S3 A) and validated it using TCR-017/018 (Fig. 3 B, clonotype 5). Using this platform, we could determine the epitopes of additional public cTfh clonotypes expanded in patients (Fig. 3 B). The restricting HLAs of these epitopes were determined using transformed B cells and HLA transfectants (Table S2 and Fig. S3, B–G). All identified epitopes were located in two major regions within the conserved trimer-forming interface of the S protein and exhibited low mutation rates (Fig. 3 C; Singer et al., 2020 *Preprint*; Elbe and Buckland-Merrett, 2017). These results reveal that public cTfh clonotypes expanded during SARS-CoV-2 infection preferentially recognize epitopes located within the S region, and these regions are conserved across the emerging variants (Emerson et al., 2017; Nolan et al., 2020 *Preprint*). Of note, sera from donors who possessed these clonotypes were positive for anti–receptor binding domain (RBD) and neutralizing antibodies (Table S1).

Crystal structure of TCR-017 reveals sequence flexibility in CDR3 extending the publicity of SARS-CoV-2–specific clonotypes

The only difference between TCR-017 and -018 is the amino acid at position 94 (Q94 and P94, respectively). Since these two residues have distinct characteristics, we suspected that any amino acid at this position allows epitope recognition. Indeed, any of 17 amino acid substitutions, except for D94, retained reactivity to the $S_{864–882}$ epitope on DRA-DRB1*15:01 (Fig. 4 A). We thus solved the crystal structure of TCR-017 and found that the residue at position 94 in CDR3β is located away from CDR3α and thus unlikely to contribute to antigen recognition (Fig. 4, B and C). C90, A91, S92, and S93 were also distant from CDR3α (Fig. 4 C), suggesting that this TCRβ may mainly use a C-terminal half of CDR3β for recognition and allow diverse TCR sequence variants or even other Vβ genes.

Considering this flexibility within the CDR3β sequence, the occurrence of extended clonotype 5 was increased (Fig. 4 D), which gives us a more precise expansion ratio in recovered patients. Indeed, increase in the frequencies of this extended clonotype in individual COVID-19 patients compared with the healthy individuals in Fig. 3 A became more significant (Fig. 4 E). In the COVID-19 cohort, the frequency was also significantly higher in patients not admitted to the intensive care unit (non-ICU), but not in ICU patients, than that in the healthy control cohort. Furthermore, it was significantly higher in non-ICU than in ICU patients (Fig. 4 E). These results suggest that expansion of this clonotype during COVID-19 is associated with mild symptoms and, therefore, its epitope may serve as an antigen for inducing protective immunity.

Public cTfh clonotypes did not cross-react with HCoVs but with symbiotic bacteria

We next examined the cross-reactive antigens of these clonotypes. “Common cold” HCoVs have been reported as cross-reactive

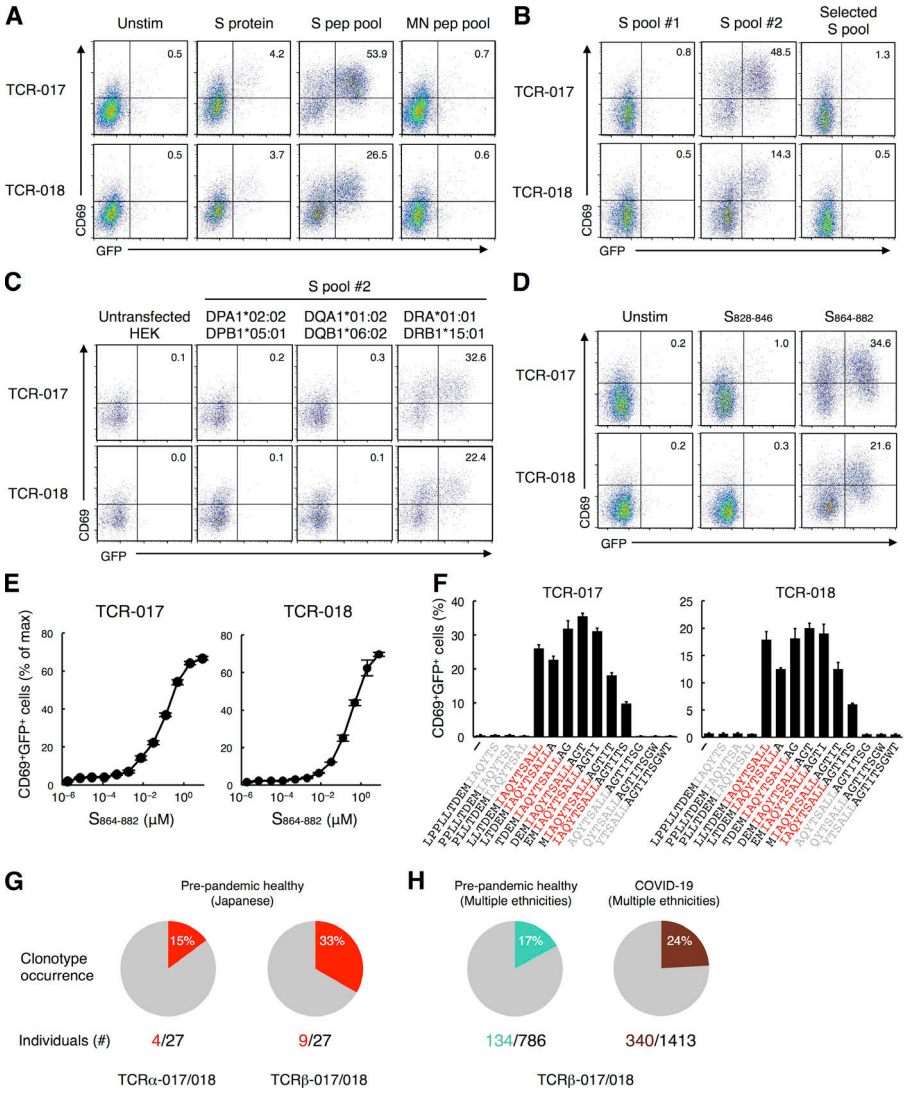


Figure 2. **TCR-017 and -018 recognize the same core epitope, $S_{870-878}$, presented on DRB1*15:01. (A)** Cells expressing TCR-017 or TCR-018 were left unstimulated or stimulated with 10 µg/ml recombinant S protein, 1 µg/ml S protein peptide pool, or 1 µg/ml MN peptide pool in the presence of autologous APCs and analyzed for GFP and CD69 expression. **(B–D)** Cells expressing TCR-017 or TCR-018 were stimulated with 0.3 µg/ml S peptide pool #1, #2, or selected S pool ($S_{304-338}$, $_{421-475}$, $_{492-519}$, $_{683-707}$, $_{741-770}$, $_{785-802}$, $_{885-1273}$; B), 1 µg/ml S peptide pool #2 in the presence of HEK293T cells expressing the indicated HLAs (C), or 1 µg/ml single peptides $S_{828-846}$ or $S_{864-882}$ (D). **(E)** Sensitivity of TCR-017 and TCR-018 to the $S_{864-882}$ epitope. TCR transfectants were stimulated with the indicated concentrations of $S_{864-882}$ peptide and APCs from donor Ts-018. Activities are shown as percentages of the maximum responses induced by plate-coated anti-CD3 mAb. Data are shown as mean ± SD of triplicates. **(F)** Determination of the core epitope recognized by TCR-017 and TCR-018 (shown in red). Cells were stimulated with 1 µg/ml of serial overlapping 15-mer peptides covering the $S_{861-887}$ region. Data are shown as mean ± SD of triplicates. **(G and H)** Occurrence of TCRα-017/018 (CDR3, CVVNRGSSYKLIF) and TCRβ-017/018 (CDR3, CASS[Q or P]TYEQYF) in a prepandemic healthy cohort from Japan (n = 27; G) and in cohorts of prepandemic healthy donors (n = 786) and convalescent COVID-19 patients (n = 1,413) from multiple ethnicities (H). Numbers of individuals possessing the TCRs among cohorts are indicated. Data are representative of two independent experiments (A–D and F), or four independent experiments (E). pep, peptide; Unstim, unstimulated.

examples for SARS-CoV-2 (Grifoni et al., 2020; Mateus et al., 2020; Braun et al., 2020). However, TCR-017/018 did not react with S proteins derived from HCoV-OC43 (Fig. S4, A and B). Furthermore, none of the SARS-CoV-2–specific public clonotypes characterized in Fig. 3 responded to HCoV peptides (Fig. 3 B; and Fig. S4, C and D). Thus, these clonotypes are unlikely to cross-react with reported HCoVs present before the outbreak of COVID-19. Instead, a homology search of core epitopes activating the most prevalent clonotype 5, $S_{870-878}$, revealed that oral commensal bacteria, *Selenomonas noxia*, possesses the same epitope in a protein called multidrug and toxic compound extrusion (MATE) family efflux transporter (Fig. 5 A). Indeed, the corresponding peptides ($MATE_{241-260}$ and $MATE_{242-256}$) that include this epitope were recognized by both TCR-017 and -018 on DRB1*15:01 (Fig. 5 B).

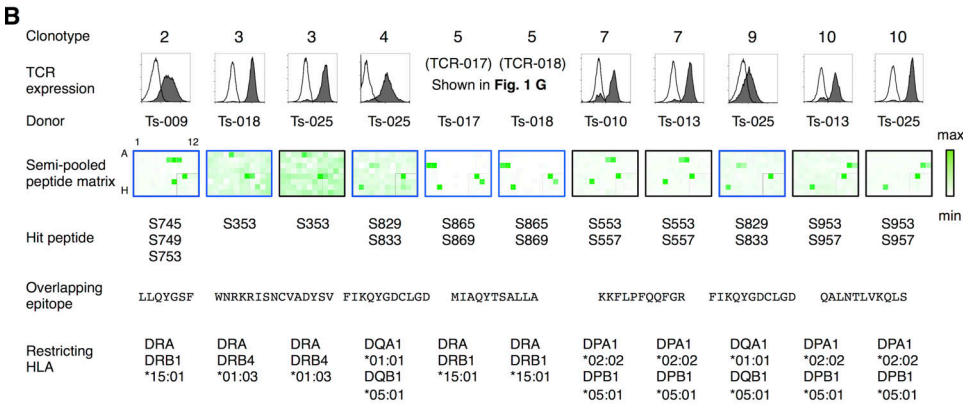
Among other epitopes, the epitope of the most expanded clonotype 2 (5.9-fold; Table S3), $S_{753-759}$, was contained in gut symbiotic microbes, such as Bacteroidales and *Klebsiella pneumoniae* (Fig. 5 A; Atarashi et al., 2017; Sefik et al., 2015). We confirmed that corresponding synthetic peptides activated T cells expressing the clonotype 2 TCR (Fig. 5 C). Furthermore, *Escherichia coli* transformed with expression vectors for these proteins could activate public clonotypes in the presence of HLA-matched dendritic cells (Fig. 5 D), confirming that cross-reactive antigenic epitopes are processed and presented. These observations imply that these bacteria might affect protective immunity against SARS-CoV-2 infection.

$S_{864-882}$ could activate multiple public cTfh clonotypes

$S_{864-882}$ is an epitope activating the most prevalent clonotypes (Fig. 4 D). To examine the ability of this epitope to activate multiple cTfh clonotypes, T cells from DRB1*15 donors were subjected to single-cell TCR- and RNA-seq (Fig. 6 A and Table S4). We identified 104 public cTfh clonotypes that responded to $S_{864-882}$ after filtering by two criteria: (1) expanded during stimulation based on single-cell sequencing (>0.1% of total $CD4^+$ T cells) and (2) detected in cohort databases (Table S4, columns P–S). Bulk TCR-seq confirmed that both α and β chains of all 104 clonotypes were indeed expanded after $S_{864-882}$ stimulation (Table S4, columns J and K). These results indicate that the $S_{864-882}$ peptide is capable of expanding multiple public cTfh clones in DRB1*15 donors. Consistent with this, we indeed

A

Clonotype	Donor	TRBV	CDR3β	TRBJ	TRAV	CDR3α	TRAJ	β occurrence (%) Healthy	β occurrence (%) COVID	β expansion p-value	αβ occurrence (%) Healthy (Japan) α	αβ occurrence (%) Healthy (Japan) β
1#	Ts-018	7-3	CASSLAGGSTDTQYF	2-3	8-2	CVVSRTGGGNKLTF	10	14.5	21.9	2.4e-8*	14.8	25.9
	Ts-009	7-3	CASSQAGGSTDTQYF	2-3	2	CAALNAGNMLTF	39					
2	Ts-009	6-1	CASSEGASNQPQHF	1-5	12-1	CVVNKGSSASKIIF	3	3.1	13.8	5.0e-7*	11.1	3.7
3#	Ts-018	20-1	CSARAGVQETQYF	2-5	12-3	CAMTYSGGYQKVTF	13	6.0	15.1	0.0001*	22.2	18.5
	Ts-025	20-1	CSARAGVQETQYF	2-5	12-3	CAMTPSGGYQKVTF	13					
4	Ts-025	20-1	CSARDWREETQYF	2-5	8-4	CAVSDQGRDDKIIF	30	2.0	16.6	0.0003*	3.7	0.0
5# (TCR-017/018)	Ts-017	11-2	CASSQTYEQYF	2-7	12-1	CVVNRGSSYKLIF	12	17.1	24.1	0.003*	14.8	33.3
	Ts-018	11-2	CASSPTYEQYF	2-7	12-1	CVVNRGSSYKLIF	12					
6	Ts-002	20-1	CSAKTSGRGETQYF	2-5	13-2	CAENSGGSNYKLTF	53	1.5	8.8	0.006*	96.3	0.0
7#	Ts-010	7-9	CASSSREDTQYF	2-3	23	CAATLYSGAGSYQLTF	28	1.7	3.1	0.01*	7.4	3.7
	Ts-013	7-9	CASSSREDTQYF	2-3	23	CAATSYSGAGSYQLTF	28					
8	Ts-013	29-1	CSVYRGLDEKLFF	1-4	23	CAASPPSYGGSQGNLIF	42	0.3	0.1	0.02*	0.0	0.0
9	Ts-025	20-1	CSARDWRGETQYF	2-5	9-2	CALSGGNRDDKIIF	30	1.9	10.3	0.02*	7.4	3.7
10#	Ts-013	11-2	CASSHGTGGRYEQYF	2-7	13-1	CAAQYALNF	41	1.5	2.6	0.03*	0.0	3.7
	Ts-025	11-2	CASSLGTSGRYEQYF	2-7	13-1	CAAGYALNF	41					



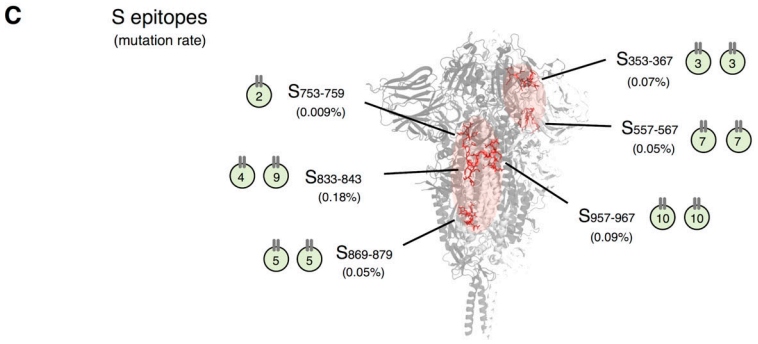


Figure 3. **Determination of epitopes of public Tfh clonotypes expanded in COVID-19 patients within conserved S regions exhibiting low mutation rates. (A)** Characterization of public clonotypes. Clonotypes detected in multiple donors from our sample pool are indicated by # (column A). Clonotype occurrences were calculated based on a prepandemic healthy cohort from Japan (n = 27, column K) and cohorts of prepandemic healthy donors (n = 786) and convalescent COVID-19 patients (n = 1,413) from multiple ethnicities (column I). For clonotypes that contain different TCRα or TCRβ sequences, occurrences were calculated using pooled sequences (columns K and I). Clonotypes are listed following the order of significance of TCRβ expansion. Clonotypes expanded significantly in recovered patients are indicated by * ($P < 0.05$, column J). **(B)** The TCR pair of each clonotype listed in A was reconstituted into reporter cells, and their epitopes and restricting HLAs were identified. TCR expression levels of paired TCRs on reporter cells are indicated by surface CD3 expression as shown in the histograms. Filled histogram, TCR-reconstituted reporter cells; open histogram, parental cells. Semi-pooled peptide matrix described in Fig. S3 A was used for epitope determination, in which reporter cells were stimulated in the presence of HEK293T expressing shared HLAs or autologous APCs and GFP reporter activities are shown as heatmaps. Unless noted otherwise, a peptide name that starts with an S followed by a number in normal size refers to a 15-mer S peptide, and the number indicates the position of its N-terminal amino acid in S protein. Restricting HLAs were determined as described in Fig. S3, B–G; Table S2, and Materials and methods. Clonotypes 1, 6, and 8 did not respond to S peptide pools. **(C)** Locations of epitopes described in B are highlighted in red in the three-dimensional structure of the SARS-CoV-2 S protein (Protein Data Bank under accession no. 6XR8). Frequencies of mutation (mutation rate) defined relative to Wuhan-Hu-1 were calculated based on data from CoV-GLUE and GISAID. Data are representative of two independent experiments (B).

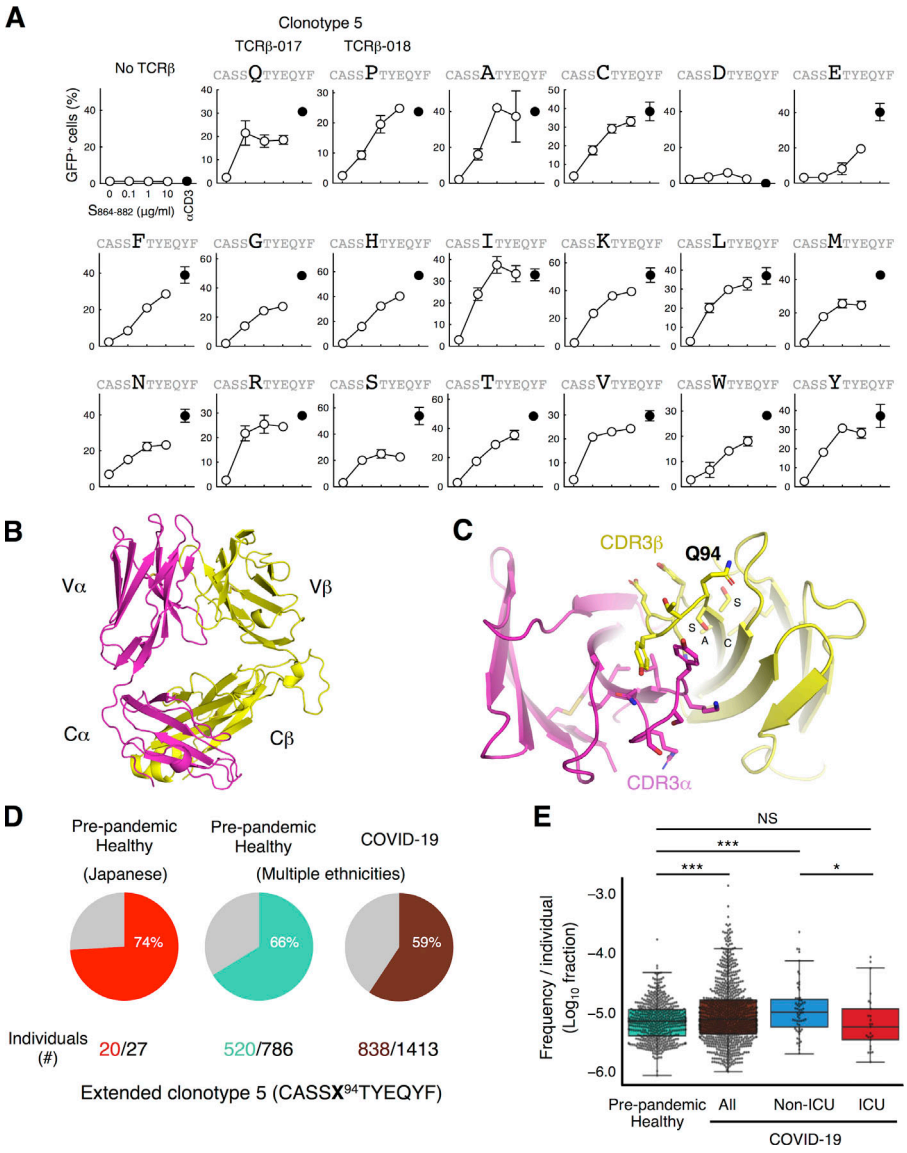


Figure 4. **Antigen recognition and clonal expansion of public clonotype 5 (TCR-017/018).** **(A)** TCRβ-017, -018, or mutants at position 94 were coexpressed with TCRα-017/018 and human CD4 in reporter cells. Cells were kept unstimulated or stimulated with peptide $S_{864–882}$ or anti-CD3 antibody in the presence of APCs from donor Ts-018. Percentages of GFP$^+$ cells in the CD3$^+$ populations are shown. Data are shown as mean ± SD of duplicates. **(B and C)** Crystal structure of TCRαβ heterodimer of clonotype 5 (TCR-017). The entire view (B) and a close-up of the CDR3 domains (C) are shown. Q94 and adjacent amino acids, C90, A91, S92, and S93 within CDR3β are indicated. **(D)** Occurrences of extended antigen-reactive TCRβ chains harboring the 17 acceptable substitutions at position 94 (shown as X; except for D94 of TCRβ-017/018) as determined in A within individuals in a prepandemic healthy cohort from Japan (n = 27) and within cohorts of prepandemic healthy donors (n = 786) and convalescent COVID-19 patients (n = 1,413) from multiple ethnicities. Numbers of individuals possessing this TCRβ among cohorts are indicated. **(E)** Frequencies of extended antigen-reactive TCRβ chains harboring the 17 acceptable substitutions at position 94 (except for D94 of TCRβ-017/018) as determined in A within the prepandemic healthy donors, all convalescent COVID-19 patients, and patients admitted to the ICU and non-ICU setting from multiple ethnicities. Horizontal line, median; bottom/top, first/third quartile; whiskers, 1.5 times the interquartile range. Data are representative of two independent experiments (A). *, P < 0.05; ***, P < 0.00001.

identified another clonotype within the Tfh cluster (Table S3) that recognizes the identical epitope $S_{864–882}$ on the same HLA through TCR distinct from clonotype 5 (Fig. 6, B–D); this clonotype was also public (Table S3, Seurat barcode, DB2_CAAGAAAGTGCGGTAA).

To examine the presence and longevity of these $S_{864–882}$-reactive clones in convalescent patients, we assembled all 104 clonotypes with bulk TCR-seq of unstimulated peripheral T cells (Fig. 6 A, day 0). 19 of these 104 clonotypes were present in the periphery of convalescent patients (Fig. 6 E). These data suggest that a substantial frequency of $S_{864–882}$-reactive public cTfh clones was maintained as a memory pool for at least 3 mo after infection (Table S1).

$S_{864–882}$ is a promiscuous epitope presented by multiple HLA alleles

DRB1*15 alleles are frequent, being shared by 25% of the population worldwide (Gonzalez-Galarza et al., 2020). However, it is also important to examine whether $S_{864–882}$ can be presented by individuals lacking DRB1*15. In silico analysis predicted 77 major HLA alleles as potential binders to $S_{864–882}$, implying that most of the world population is capable of presenting peptides that include this epitope (Table S5; Vita et al., 2019). In line with this, most randomly collected healthy donors possessed some of these predicted binding alleles (Table S6); this was confirmed by assessing the direct binding of biotinylated epitope (Fig. S5). Indeed, peripheral T cells from these donors responded to $S_{864–882}$ to express activation markers (Fig. 6 F). The number of activated CD4$^+$CXCR5$^+$ cells was also increased upon epitope stimulation (Fig. 6 G). These results suggest that $S_{864–882}$ is a universal epitope activating multiple CD4$^+$ T cells upon presentation on multiple HLAs and, therefore, is a critical region of the S protein serving as a T cell antigen in SARS-CoV-2 infection.

Discussion

This study reports the identification of cTfh epitopes within the SARS-CoV-2 S protein that may contribute to the protective

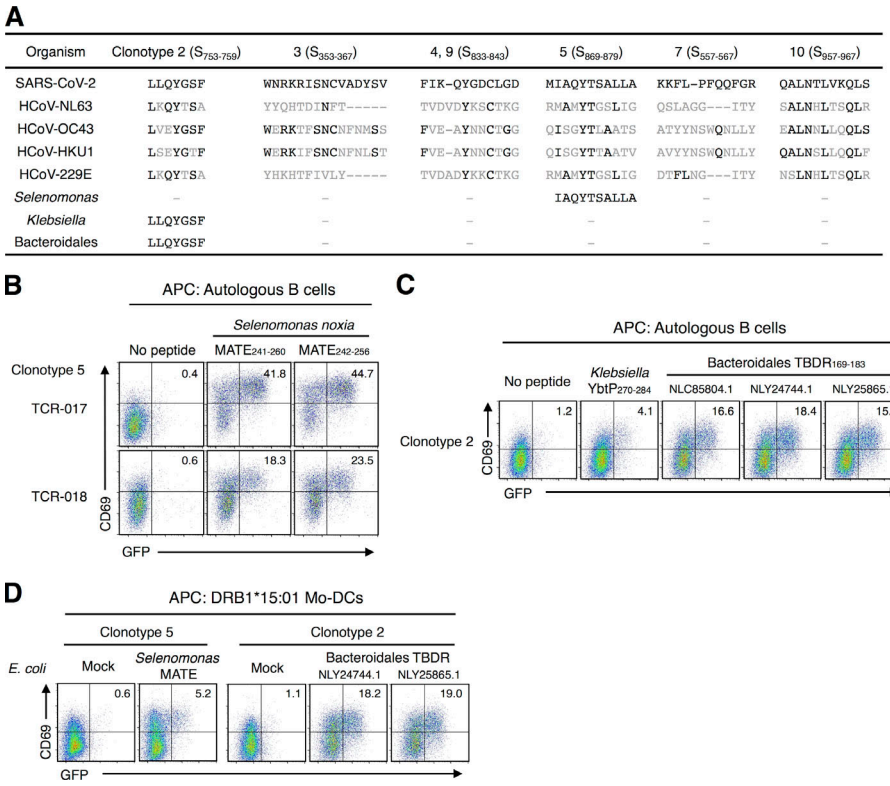


Figure 5. **Cross-reactivity of SARS-CoV-2–specific public Tfh clonotypes. (A)** Alignment of corresponding peptides. Amino acids identical to SARS-CoV-2 epitopes are shown in black. **(B)** Cells expressing clonotype 5 (TCR-017 or TCR-018) were stimulated with 1 µg/ml of peptides derived from the MATE family efflux transporter of *S. noxia*. **(C)** Cells expressing the TCR from clonotype 2 were stimulated with 1 µg/ml of peptides derived from *K. pneumoniae* yersiniabactin ABC transporter ATP-binding/permease protein (YbtP) or Bacteroidales bacterium TonB-dependent receptor (TBDR). Accession numbers of TBDRs from different Bacteroidales bacteria in GenBank are shown. **(D)** Cells expressing TCRs from clonotype 5 (TCR-017) or clonotype 2 were stimulated with heat-killed *E. coli* transformed with expression vectors for the indicated microbial proteins in the presence of monocyte-derived dendritic cells (Mo-DCs) expressing DRB1*15:01. Data are representative of two independent experiments (B–D).

T cell responses against COVID-19. In the wake of the COVID-19 outbreak, some public TCRs recognizing SARS-CoV-2 have recently been reported (Shomuradova et al., 2020; Dykema et al., 2021). In the present study, we first identified and determined the crystal structure of SARS-CoV-2–specific public TCRs of $CD4^+$ T cells associated with mild COVID-19 symptoms. Individuals possessing such public cTfh clonotypes are expected to recognize S protein and to induce the production of antibodies against this cognate antigen for immune protection (Kaneko et al., 2020; Zhang et al., 2021; Gong et al., 2020). Consistent with this view, convalescent patients who expressed the public cTfh clonotypes in our sample pool had anti–S-neutralizing antibodies and recovered from COVID-19 (Table S1).

As these public clonotypes were widely detected in multiple ethnicities, precise monitoring of the frequencies of public Tfh clonotypes by developing specific probes could become a novel option for prognosis prediction. Given that their restricting HLAs cover a large global population (Gonzalez-Galarza et al., 2020) and, in particular, $S_{864–882}$ is a universal epitope presented by multiple HLAs, such common S epitopes are promising antigens for promoting protective Tfh responses against SARS-CoV-2 worldwide.

Since most of the identified epitopes are located in the non-RBD region and included in the current full-length S vaccines, this study provides a molecular affirmation of and additional rationale for the vaccines being administered worldwide (Baden et al., 2021; Walsh et al., 2020; Sadoff et al., 2021; Voysey et al., 2021; Shinde et al., 2021). As these epitopes are conserved among SARS-CoV-2 variants (Fig. 3 C; Singer et al., 2020 *Preprint*; Elbe and Buckland-Merrett, 2017), these peptides are expected to generate T cell memory against Alpha: B.1.1.7 (UK), Beta: B.1.351 (South Africa), Gamma: P.1 (Brazil), Delta: B.1.617.2 (India), and future mutants that escape from neutralizing antibodies (Wang et al., 2021; Collier et al., 2021; Cele et al., 2021). The universal T epitope $S_{864–882}$ could be a candidate for peptide vaccines upon coupling with appropriate linear B cell epitopes (Sauer et al., 2021); in fact such a combined peptide may provide an "adjustable" booster for variants in the postvaccine era. Furthermore, identification of additional public clonotypes and their epitopes using the platform established in this study would be an intriguing next step in extending the number of candidate T cell antigens. In addition, cross-reacting symbiotic bacteria may contribute to the priming of SARS-CoV-2–specific T cells, thus providing a novel perspective on prognosis and prevention. The search for antigenic cross-reactive epitopes having acceptable substitutions will expand the pool of environmental "priming" antigens. Further global metagenomic analysis might clarify the relationship between the presence of cross-reactive commensals and resistance to COVID-19, offering a potential explanation of local/ethnic variation in disease severity and another route to fighting this virus.

The current study has a limitation in the definition of Tfh cells, as we collected T cells from PBMCs. Although circulating Tfh cells in the periphery reflect germinal center Tfh cells (Locci et al., 2013; Hill et al., 2019; Heit et al., 2017; Schmitt et al., 2014; Crotty, 2019), the analysis of tissue-resident Tfh cells using biopsy samples is ideal. Different TCR-seq methods also have advantages and limitations. Single-cell TCR-seq can determine αβ pairs, but throughput is limited. Bulk sequencing can detect more diverse clonotypes, but pair information cannot be obtained. Taken together, the current study shows that a combination of single-cell and bulk TCR-seq, global TCR databases, and

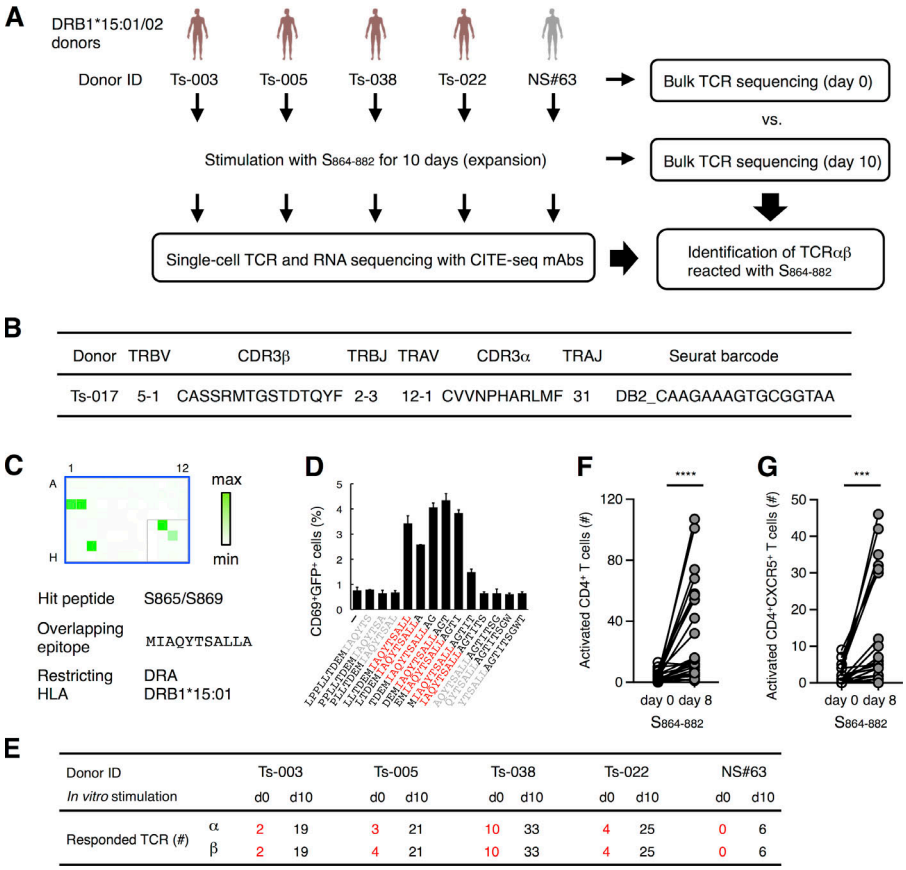


Figure 6. **Characterization of $S_{864–882}$-reactive T cells. (A)** Workflow of the analysis for the $S_{864–882}$-induced T cell expansion. PBMCs from convalescent COVID-19 patients (red) and healthy (gray) donors possessing DRB1*15:01 or *15:02 were analyzed by bulk TCR-seq or stimulated with 1 µg/ml $S_{864–882}$ peptide for 10 d before $CD4^+$ T cells were sorted and analyzed by single-cell TCR- and RNA-seq as well as bulk TCR-seq. **(B)** V and J usages and CDR3 sequences of the newly identified clonotype that recognizes $S_{870–878}$ epitope. **(C)** Reporter cells expressing TCRαβ in B were stimulated with pooled peptide matrix described in Fig. S3 A in the presence of DRA-DRB1*15:01–expressing APCs. Wells containing S865 or S869 (C1, C2, and G3), S pool #2 (E10), and $S_{864–882}$ (F11) were positive. **(D)** 1 µg/ml of serial overlapped 15-mer peptides covering the $S_{861–887}$ region were tested for reactivity to the clonotype in B to determine the minimum epitope (shown in red). Data are shown as mean ± SD of triplicates. **(E)** Numbers of TCRα and β of expanded clonotypes (proportion >0.1% in single-cell TCR-seq) detected in bulk TCR-seq described in Table S4 are shown. **(F and G)** 10^5 of PBMCs from healthy donors were left unstimulated (day 0) or stimulated with $S_{864–882}$ for 8 d (day 8). Numbers of $CD4^+CD69^+CD137^+$ T cells (F) and $CD4^+CD69^+CD137^+CXCR5^+$ Tfh cells (G) are shown. Data are representative of two independent experiments (C and D). ***, $P < 0.005$; ****, $P < 0.001$. ID, identifier.

the TCR reconstitution/epitope determination platform is an efficient workflow to identify beneficial public helper/cytotoxic clonotypes and epitopes. This methodology could provide a powerful tool for future outbreaks of other infectious diseases.

Materials and methods

Human subjects

The institutional review boards of Osaka University (approval number 898-4) and National Institute of Infectious Diseases (approval number 1237) approved blood draw protocols for convalescent COVID-19 and healthy individuals. The institutional review board of the nonprofit organization MINS (approval number 190210) approved the analyses in KOTAI Biotechnologies using blood samples from healthy individuals. The research was performed in accordance with all relevant guidelines and regulations. Written informed consent was obtained from all participants or designated health care surrogates if participants were unable to provide informed consent. Study enrollment criteria included subjects >20 yr old, regardless of disease severity and genders (Table S1). Prior to enrollment in this study, all COVID-19 donors were confirmed to be positive for SARS-CoV-2 by PCR using nasopharyngeal swab specimens. Blood from convalescent COVID-19 donors and healthy donors was obtained at Tokyo Shinagawa Hospital and Osaka University. Samples were deidentified before analyses. For PBMC preparation, whole blood was collected in heparin-coated tubes and centrifuged to separate the cellular fraction and plasma followed by density-gradient sedimentation. For neutralizing antibody assay, plasma samples from patients were heat inactivated for 30 min at 56°C. Plasma diluted at 1:5 followed by twofold serial dilutions was incubated with equal volume of solution containing 100 median tissue culture infectious dose of SARS-CoV-2 for 1 h at 37°C and added to VeroE6/TMPRSS2 cells. After incubation for 5 d, the highest plasma dilution that protected 100% of cells from cytopathic effect was taken as the neutralization titer. Disease severity was defined as mild, moderate I, moderate II, or severe in accordance with the Japanese COVID-19 Clinical Practice Guideline Version 2.2 (Ministry of Health, Labour and Welfare, 2020). The study using samples from healthy donors was approved by the ethical committee of Osaka University Hospital (approval number 20483). Comprehensive informed consent has been previously obtained and approved by the institutional review board of Osaka Universal Hospital (approval number 209008-B). Waiver of informed consent was approved for projects involving the secondary analysis of the residual samples, and an opt-out method was applied to obtain consent in this study by using the announcement online.

SARS-CoV-2

For the preparation of inactivated SARS-CoV-2, SARS-CoV-2 (KNG19-020) was kindly supplied by Dr. Tomohiko Takasaki (Kanagawa Prefectural Institute of Public Health, Kanagawa, Japan). The virus was propagated in VeroE6/TMPRSS2 cells (JCRB1819) and purified by sucrose gradient centrifugation

(Dent and Neuman, 2015). Concentrated virus was then inactivated by ultraviolet light (0.6 J/cm^2).

Antigens

Inactivated virus was used at 2.9 × 10^6 viral particles/ml (~1 µg S protein/ml) for stimulation. For the preparation of recombinant SARS-CoV-2 S protein, the ectodomain of S protein was cloned into a mammalian expression vector pCMV, with a foldon sequence followed by 9× His-tag and Strep-tag at the C-terminus. Polybasic cleavage site (RRAR) of S protein was replaced by a single alanine, and K986P and V987P mutations were introduced to stabilize the conformation as previously described (Amanat et al., 2021). After transient transfection into Expi293F cells (Gibco), secreted protein was purified from the culture supernatant 4 d after transfection using TALON Metal Affinity Resin (Clontech) and an Amicon Ultra 10K Centrifugal Filter Device (Millipore). For the expression of S proteins of SARS-CoV-2 and HCoV-OC43 for antigen presentation, pME18S expression vectors for each protein were transfected into HEK293T cells. For the expression of bacterial proteins in *E. coli*, cDNA for MATE family efflux transporter, TonB-dependent receptor, and TonB-dependent receptor plug domain–containing protein was cloned into pCold expression vector and transformed into *E. coli* BL21 (DE3) competent cells, Champion21 (SMBIO Technology). The protein expressions were induced by the addition of 0.5 or 1.0 mM isopropyl-β-D-thiogalactopyranoside at 37°C. Bacterial pellets were washed and resuspended with PBS and incubated at 95°C for 5 min for inactivation. The expressions of the proteins were confirmed by SDS-PAGE and Coomassie brilliant blue staining. PepMix SARS-CoV-2 (Spike Glycoprotein), which contains pool #1 and #2, was purchased from JPT Peptide Technologies. PepTivator SARS-CoV-2 for S, M, and N protein were purchased from Miltenyi Biotec. Domains of SARS-CoV-2 S protein shown in Fig. S2 A were previously described (Wrapp et al., 2020). All of other peptides were from GenScript. Three-dimensional structure of SARS-CoV-2 S protein (Protein Data Bank accession no. 6XR8) was depicted with program PyMOL (The PyMOL Molecular Graphics System, Version 2.0; Schrödinger, LLC).

In vitro stimulation of PBMCs

Cryopreserved PBMCs were thawed and washed with RPMI 1640 medium (Sigma) supplemented with 5% human AB serum (Gemini Bio), penicillin (Sigma), streptomycin (MP Biomedicals), and 2-mercaptoethanol (Nacalai Tesque). 5 × 10^5 PBMCs were stimulated in the same medium with inactivated SARS-CoV-2, 1 or 10 µg/ml of recombinant S protein, 1 µg/ml of S peptide pool, or 1 µg/ml of MN peptide pool for 20 h, followed by staining with anti-human CD3, CD69, CD137, CD154, and TotalSeq-C Hashtag antibodies. CD69$^+$, CD137$^+$, or CD154$^+$ cells within a CD3$^+$-gated population were sorted by SH800S Cell Sorter (Sony Biotechnology) and used for single-cell TCR- and RNA-seq analyses. For epitope-specific clonal expansion, 1–5 × 10^5 PBMCs were stimulated with 1 µg/ml of $S_{864–882}$ for 10 d, and recombinant human IL-2 (1 ng/ml, PeproTech) was added at day 4 and day 7. CD4$^+$ T cells were sorted and analyzed by single-cell TCR- and RNA-seq.

Single-cell–based transcriptome and TCR repertoire analysis

Single-cell capturing and library preparation were performed using the following reagents: Chromium Next GEM Single Cell 5′ Library & Gel Bead Kit v1.1, 16 rxns, PN-1000165; Chromium Next GEM Chip G Single Cell Kit, 48 rxns, PN-1000120; Chromium Single Cell V(D)J Enrichment Kit, Human T Cell, 96 rxns, PN-1000005; Single Index Kit T Set A, 96 rxns, PN-1000213; Chromium Single Cell 5′ Feature Barcode Library Kit, 16 rxns, PN-1000080; and Single Index Kit N Set A, 96 rxns, PN-1000212. Single-cell suspension containing ~2 × 10^4 cells were loaded into Chromium microfluidic chips to generate single-cell gel bead-in-emulsion using Chromium Controller (10x Genomics) according to the manufacturer's instructions. RNA from the barcoded cells for each sample was subsequently reverse-transcribed inside gel bead-in-emulsion using a Veriti Thermal Cycler (Thermo Fisher Scientific), and all subsequent steps to generate single-cell libraries were performed according to the manufacturer's protocol, with 14 cycles used for cDNA amplification. Then, ~50 ng of cDNA was used for gene expression library amplification by 14 cycles in parallel with cDNA enrichment and library construction for TCR libraries. Fragment size of the libraries was confirmed with an Agilent 2100 Bioanalyzer. Libraries were sequenced on Illumina NovaSeq 6000 as paired-end mode (read 1, 28 bp; read 2, 91 bp). The raw reads were processed by Cell Ranger 3.1.0 (10x Genomics). Gene expression–based clustering was performed using the Seurat R package (v3.1; Hafemeister and Satija, 2019). Briefly, cells with a mitochondrial content >10% and cells with <200 or >4,000 genes detected were considered outliers (dying cells and empty droplets and doublets, respectively) and filtered out. The Seurat SCTransform function was used for normalization, and data were integrated without performing batch-effect correction as all samples were processed simultaneously. Hashtag oligo demultiplexing was performed on centered log ratio–normalized hashtag unique molecular identifier counts, and clonotypes were matched to the gene expression data through their droplet barcodes, using Python scripts. Only cells assigned a single hashtag and a β-chain clonotype were retained for downstream analyses. Tfh scores and activation scores were generated using a published list of Tfh-enriched transcripts (Meckiff et al., 2020) and four well-known activation genes (*CD69*, *TNFRSF9*, *TNFRSF4*, and *CD40LG*), respectively. Calculations were performed in R with the AddmoduleScore function of Seurat, using the default parameters. Briefly, the average expression levels of all genes of the gene list were computed and subtracted by the average expression of control genes randomly selected from similar aggregate expression level bins.

Bulk TCR-seq and analysis

1–3 × 10^5 PBMCs were lysed in QIAzol (QIAGEN). Full-length cDNA was then synthesized using SMARTer technology (Takara Bio), and the variable regions of TCRα and β genes were amplified using TRAC/TRBC-specific primers. After sequencing of the variable region amplicons, each pair of reads was assigned a clonotype (defined as TR(A/B)V and TR(A/B)J genes and CDR3) using MiXCR software (Bolotin et al., 2015). For each α/β clonotype, expansion was defined as the fraction of reads for that

clone divided by the total number of reads for the α/β chain, respectively, and fractions were converted to $\log_{10}$ for plotting and statistical analyses.

HLA class II typing

For HLA class II typing, genomic DNA samples were isolated from PBMCs using QIA DNA Mini Kit (QIAGEN). AllType FASTplex NGS 11 Loci Flex Kit (One Lambda) was used to prepare DNA sequence libraries (DPA1, DPB1, DQA1, DQB1, DRB1, and DRB3/4/5) according to the manufacturer's protocol. Sequencing was performed on a MiSeq System (Illumina). TypeStream Visual version 2.0 (One Lambda) was used to analyze the DNA sequences. For DRB1*15:01/02 typing, genomic DNA was amplified with a DRB1*15/16-specific forward primer (5′-CGTTTCCTG TGGCAGCCTAAGAGG-3′) and a DRB1*15-specific reverse primer (5′-CCGCGCCTGCTCCAGGAT-3′) followed by DNA sequencing.

APCs

Recombinant EBV was produced as previously reported (Kanda et al., 2015). Viral stocks were obtained by concentrating virus-containing culture supernatants by ultracentrifugation at 32,000 rpm for 1 h. 3×10^5 PBMCs were incubated with an aliquot of the viral stock for 1 h at 37°C. The infected cells were cultured with RPMI 1640 medium supplemented with 20% FBS (Capricorn Scientific GmbH) containing 0.1 µg/ml cyclosporine A (Cayman Chemical). EBV-immortalized B lymphoblastoid cell lines were obtained after 3-wk culture and used as APCs. To generate APCs expressing specific HLA, HEK293T cells were transfected with plasmids encoding HLA class II alleles as previously described (Jiang et al., 2013). EBV experiments were approved by the Ministry of Education, Culture, Sports, Science and Technology (approval number 539) and the institutional review board of Osaka University (approval number 04658). To generate monocyte-derived dendritic cells, $CD14^+$ cells were isolated from PBMCs using CD14 MicroBeads, human (Miltenyi Biotec), and cultured in RPMI 1640 medium supplemented with 10% FBS, 0.1 mM Non-Essential Amino Acids Solution (Gibco), 1 mM sodium pyruvate (Gibco), 10 ng/ml human GM-CSF (PeproTech), and IL-4 (PeproTech) for 6 d.

TCR reconstitution and stimulation

TCRα and β chain cDNA sequences were synthesized and cloned into retroviral vectors pMX-IRES-rat CD2 and pMX-IRES-human CD8, respectively. Two vectors containing paired TCRα and β chains were transfected together into Phoenix packaging cells using PEI MAX (Polysciences). Supernatant containing retroviruses was used for infection into mouse T cell hybridoma with an NFAT-GFP reporter gene (Matsumoto et al., 2021) to reconstitute TCRαβ pairs. TCRβ mutants were constructed by site-directed mutagenesis. For antigen stimulation, TCR-reconstituted cells were cocultured with stimulants in the presence of immortalized autologous B cells unless indicated otherwise. After 20 h, T cell activation was assessed by GFP or CD69 expression.

Rapid epitope determination platform

For the preparation of pooled peptide matrix, 15-mer peptides with 11-amino acid overlap that cover the full length of S protein of SARS-CoV-2 were individually synthesized (GenScript). Peptides were dissolved in DMSO at 12 mg each peptide/ml, and 2–12 peptides were mixed to create 75 different semipools, as indicated in Fig. S3 A, so that the responsible epitopes could be determined from the cell reactivities to horizontal and vertical pools. Pooled peptides were adjusted to 1 mg each peptide/ml in DMSO, followed by 10× dilution with water, and 1 µl of the solution was added into each corresponding well. Commercial S peptide pools and $S_{864–882}$ peptide were also loaded as controls. S peptide pools of four common cold HCoVs (229E, NL63, OC43, and HKU1) were loaded to assess cross-reactivities. For reporter cell assay, 100 µl of media containing T cells and APCs were added to the wells so that T cells were stimulated with each peptide at 1 µg/ml. After 20 h, GFP reporter induction was assessed in a T cell–gated population using an Attune NxT flow cytometer (Thermo Fisher Scientific).

Restricting HLA determination

For the clonotypes detected in more than one donor from our sample pool, reporter cells were stimulated in the presence of HEK293T cells expressing individual pairs of alleles shared by the donors. Otherwise, reporter cells were stimulated in the presence of transformed B cells from various donors that have partial overlapped HLAs with the original donor of the clonotype.

Database analyses

Public repertoire datasets of prepandemic healthy donors (n = 786, Adaptive ImmuneACCESS; Emerson et al., 2017) or COVID-19 convalescent donors (n = 1,413, Adaptive ImmuneRACE; Nolan et al., 2020 *Preprint*) of various ethnicities from the US, Italy, and Spain were downloaded from the Adaptive Biotechnologies website, and V-GENE names were renamed to match the IMGT nomenclature. Finally, clonotypes and their expansions were defined in the same way as in-house bulk TCR-seq. TCR occurrence in a dataset was defined as the fraction of patients whose repertoire contained the given TCR, regardless of its expansion. A two-sided Student's t test was used to compare expansion values. Considering the extremely high diversity of observed TCR clonotype sequences, conventional P value adjustment for multiple testing has not been performed in the field (Emerson et al., 2017), and P values are therefore nonadjusted unless stated otherwise.

Expression and purification of soluble TCRαβ heterodimer

Expression constructs encoding the extracellular domains of TCRα-017 (Q1-D207 in mature protein) and TCRβ-017 (G1-G241) subunits were incorporated into pCold vector including 6× His-tag and a tobacco etch virus protease cleavage site. For crystallization, the point mutations (T159C in α and S166C/C184A in β) were introduced to form an artificial disulfide bond as described previously (Boulter et al., 2003). The plasmids were transformed into *E. coli* BL21 competent cells Champion21 and Rosetta2 (DE3; Novagen), respectively. The protein expression was induced by the addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside at 18°C. Cells were suspended with 500 mM NaCl-containing Tris-HCl buffer (pH 8.0) and disrupted with sonication. The inclusion bodies, including target proteins, were collected by centrifugation.

The inclusion body was then solubilized by 50 mM Tris-HCl buffer (pH 8.0) containing 6 M guanidine HCl, 10 mM EDTA, and 2 mM dithiothreitol at room temperature. The equal amount of solubilized TCRα and TCRβ (35 mg each) was mixed and rapidly diluted with 1 liter of 100 mM Tris-HCl buffer (pH 8.0) containing 5 M urea, 0.4 M L-arginine, 5 mM reduced glutathione, and 0.5 mM oxidized glutathione at 4°C. The diluted solution was further dialyzed against 10 mM Tris-HCl buffer (pH 8.0) at 4°C for 2 d. The dialyzed solution was applied onto 5 ml nickel-nitrilotriacetic acid agarose (FUJIFILM Wako), and His-tagged TCRαβ were eluted with elution buffer (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, and 250 mM imidazole). After removal of His-tag by tobacco etch virus protease, the eluted protein was concentrated and further applied to Superdex 75 (GE Healthcare) equilibrated with 20 mM Tris-HCl (pH 8.0) buffer containing 100 mM NaCl. The TCRαβ heterodimer fractions were concentrated up to 1.2 mg/ml by Amicon Ultra (molecular weight cutoff, 10 kD). The purity of the proteins was assessed by SDS-PAGE and Coomassie brilliant blue staining. The numbering of amino acids in CDR3 sequences of TCRs was based on mature protein.

Crystallization, data collection, and structure determination of TCR-017 ectodomain

All crystallization trials were performed by sitting drop vapor diffusion method. Initial crystallization conditions were screened using Index (Hampton Research), SG1 Screen, and SG2 Screen (Molecular Dimensions). The best diffracted crystal was obtained under the condition of 0.1 M Hepes-Na (pH 7.0) and 1.1 M sodium malonate at 20°C. Prior to x-ray diffraction experiments, crystals were soaked in the reservoir containing 20% ethylene glycol and flash cooled in liquid nitrogen. X-ray diffraction datasets were collected at the beamline BL-1A in the Photon Factory. Diffraction data were integrated with program XDS (Kabsch, 2010) and scaled with program SCALA (Evans, 2006). The phases of datasets were determined by molecular replacement method using program MOLREP (Vagin and Teplyakov, 2010) with the coordinate of 1E6 TCR (Protein Data Bank accession no. 5C0B; Cole et al., 2016). After initial phase determination, the model buildings were manually performed using program COOT (Emsley et al., 2010). Refinement was performed using program REFMAC5 (Vagin et al., 2004) at the initial step and Phenix.refine for the final model (Afonine et al., 2012). The stereochemical quality of the final model was assessed by program MolProbity (Williams et al., 2018). Data collection and refinement statistics are summarized in Table S7. Structural factors and the atomic coordinates of TCRαβ-017 ectodomain have been deposited in Protein Data Bank under accession no. 7EA6. All figures of 3D structure were depicted with program PyMOL (The PyMOL Molecular Graphics System, Version 2.0; Schrödinger, LLC).

Antibodies

Anti-human CD3 (HIT3a), anti-human CD8 (RPA-T8), anti-human CD69 (FN50), anti-human CD137 (4B4-1), anti-human CD154 (24-31), anti-human CD4 (OKT4), anti-human CXCR5 (J252D4), TotalSeq-C Hashtags (LNH-94; 2M2), anti-mouse CD3 (17A2), anti-mouse CD69 (H1.2F3), and anti-rat CD2 (OX-34) antibodies were purchased from BioLegend. Rat IgG2b κ Isotype Control (eB149/10H5) was purchased from eBioscience.

Statistical analysis

Statistical analysis in Fig. 3 A and Fig. 4 E was performed with a two-sided Student's *t* test to compare expansion values, and P values are indicated. Analyses in Fig. 6, F and G, were performed with paired *t* test, and P values are indicated.

Online supplemental material

Fig. S1 shows our workflow for single-cell–based analyses of SARS-CoV-2–responsive T cells. Fig. S2 shows epitope peptides and restricting HLAs of clonotype-017/018. Fig. S3 shows identification of the epitopes and restricting HLAs of public clonotypes. Fig. S4 shows no cross-reactivity of clonotype-017/018 to other HCoVs. Fig. S5 shows presentation of $S_{868–880}$ epitope on multiple HLA alleles. Table S1 shows donor characteristics. Table S2 shows HLA class II types of involved blood donors. Table S3 shows cTfh clonotypes identified in single-cell analysis. Table S4 shows $S_{864–882}$-reactive public cTfh clonotypes. Table S5 shows frequencies of HLAs in the Japanese population that are predicted to present epitopes of public clonotypes. Table S6 shows predicted binding alleles to $S_{864–882}$ in healthy donors. Table S7 shows data collection and refinement statistics of the crystallographic analysis of TCRαβ-017.

Data availability

Recombinant SARS-CoV-2 S protein, TCR-reconstituted cells, and soluble TCRαβ heterodimers are available from the corresponding author on request under a standard material transfer agreement. Single-cell–based transcriptome data have been deposited in Gene Expression Omnibus datasets under accession no. GSE184806). Structural factors and the atomic coordinates of the TCRαβ-017 ectodomain have been deposited in Protein Data Bank under accession no. 7EA6). Other data needed to support the study conclusions are included in the main text and online supplemental material.

Acknowledgments

We thank T. Ito, K. Toyonaga, Y. Adachi, S. Moriyama, Y. Harima, D. Azizah, H. Hayashi, and J. Sun for experimental support and C. Schutt, W. Ise, J. B. Wing, D. Okuzaki, D. Standley, S. Teraguchi, S. Futami, T. Sasazuki, and T. Kobayashi for discussion.

This research was supported by Japan Agency for Medical Research and Development (20fk0108542, 20fk0108403 [H. Arase], 20fk0108265, 20nk0101602, 20fk0108454, 21nf0101623, 21gm0910010, 21ak0101070, 20fk0108075, 20fk0108104, 21fk0108608, and 21fk0108534 [S. Yamasaki]), Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research JP18H05279 (H. Arase) and JP20H00505 (S. Yamasaki), the Kansai Economic Federation, and the Mitsubishi Foundation (S. Yamasaki). The Department of Health Development and Medicine is an endowed department supported by AnGes, Daicel, and FunPep.

Author contributions: S. Yamasaki conceptualized research; X. Lu, Y. Hosono, S. Ishizuka, E. Ishikawa, M. Nagae, and

Y. Ozaki did investigation; R. Shinnakasu, T. Inoue, T. Onodera, T. Matsumura, M. Shinkai, T. Sato, S. Mori, T. Kanda, E.E. Nakayama, T. Shioda, T. Kurosaki, H. Arase, and Y. Takahashi provided resources; D. Motooka, N. Sax, Y. Maeda, Y. Kato, T. Morita, S. Nakamura, M. Nagae, K. Takeda, A. Kumanogoh, H. Nakagami, and K. Yamashita did data curation; S. Yamasaki supervised the research; and X. Lu, Y. Hosono, and S. Yamasaki wrote the manuscript.

Disclosures: N. Sax reported personal fees from KOTAI Biotechnologies, Inc. outside the submitted work; in addition, N. Sax had a patent to "follicular helper T cells specific for SARS-CoV-2 virus" pending and a patent to "novel medical technology using follicular T cells" pending. H. Nakagami reported that the Department of Health Development and Medicine is an endowed department supported by Anges, Daicel, and Funpep. K. Yamashita reported personal fees from KOTAI Biotechnologies, Inc. outside the submitted work; in addition, K. Yamashita had a patent to "follicular helper T cells specific for SARS-CoV-2 virus" pending and a patent to "novel medical technology using follicular T cells" pending. No other disclosures were reported.

Submitted: 18 June 2021
Revised: 21 August 2021
Accepted: 28 September 2021

References

Afonine, P.V., R.W. Grosse-Kunstleve, N. Echols, J.J. Headd, N.W. Moriarty, M. Mustyakimov, T.C. Terwilliger, A. Urzhumtsev, P.H. Zwart, and P.D. Adams. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* 68:352–367. https://doi.org/10.1107/S0907444912001308

Amanat, F., S. Strohmeier, R. Rathnasinghe, M. Schotsaert, L. Coughlan, A. García-Sastre, and F. Krammer. 2021. Introduction of two prolines and removal of the polybasic cleavage site lead to higher efficacy of a recombinant spike-based sars-cov-2 vaccine in the mouse model. *MBio.* 12:1–10. https://doi.org/10.1128/mBio.02648-20

Atarashi, K., W. Suda, C. Luo, T. Kawaguchi, I. Motoo, S. Narushima, Y. Kiguchi, K. Yasuma, E. Watanabe, T. Tanoue, et al. 2017. Ectopic colonization of oral bacteria in the intestine drives T_H1 cell induction and inflammation. *Science.* 358:359–365. https://doi.org/10.1126/science.aan4526

Baden, L.R., H.M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S.A. Spector, N. Rouphael, C.B. Creech, et al. COVE Study Group. 2021. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N. Engl. J. Med.* 384:403–416. https://doi.org/10.1056/NEJMoa2035389

Begovich, A.B., G.R. McClure, V.C. Suraj, R.C. Helmuth, N. Fildes, T.L. Bugawan, H.A. Erlich, and W. Klitz. 1992. Polymorphism, recombination, and linkage disequilibrium within the HLA class II region. *J. Immunol.* 148:249–258.

Bolotin, D.A., S. Poslavsky, I. Mitrophanov, M. Shugay, I.Z. Mamedov, E.V. Putintseva, and D.M. Chudakov. 2015. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods.* 12:380–381. https://doi.org/10.1038/nmeth.3364

Boulter, J.M., M. Glick, P.T. Todorov, E. Baston, M. Sami, P. Rizkallah, and B.K. Jakobsen. 2003. Stable, soluble T-cell receptor molecules for crystallization and therapeutics. *Protein Eng.* 16:707–711. https://doi.org/10.1093/protein/gzg087

Braun, J., L. Loyal, M. Frentsch, D. Wendisch, P. Georg, F. Kurth, S. Hippenstiel, M. Dingeldey, B. Kruse, F. Fauchere, et al. 2020. SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature.* 587:270–274. https://doi.org/10.1038/s41586-020-2598-9

Cele, S., I. Gazy, L. Jackson, S.-H. Hwa, H. Tegally, G. Lustig, J. Giandhari, S. Pillay, E. Wilkinson, Y. Naidoo, et al. COMMIT-KZN Team. 2021. Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. *Nature.* 593:142–146. https://doi.org/10.1038/s41586-021-03471-w

Cole, D.K., A.M. Bulek, G. Dolton, A.J. Schauenberg, B. Szomolay, W. Rittase, A. Trimby, P. Jothikumar, A. Fuller, A. Skowera, et al. 2016. Hotspot autoimmune T cell receptor binding underlies pathogen and insulin peptide cross-reactivity. *J. Clin. Invest.* 126:2191–2204. https://doi.org/10.1172/JCI85679

Collier, D.A., A. De Marco, I.A.T.M. Ferreira, B. Meng, R.P. Datir, A.C. Walls, S.A. Kemp, J. Bassi, D. Pinto, C. Silacci-Fregni, et al. COVID-19 Genomics UK (COG-UK) Consortium. 2021. Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies. *Nature.* 593:136–141. https://doi.org/10.1038/s41586-021-03412-7

Crotty, S. 2019. T follicular helper cell biology: a decade of discovery and diseases. *Immunity.* 50:1132–1148. https://doi.org/10.1016/j.immuni.2019.04.011

Dent, S., and B.W. Neuman. 2015. Purification of coronavirus virions for Cryo-EM and proteomic analysis. *Methods Mol. Biol.* 1282:99–108. https://doi.org/10.1007/978-1-4939-2438-7_10

Dykema, A.G., B. Zhang, B.A. Woldemeskel, C.C. Garliss, L.S. Cheung, D. Choudhury, J. Zhang, L. Aparicio, S. Bom, R. Rashid, et al. 2021. Functional characterization of CD4+ T cell receptors crossreactive for SARS-CoV-2 and endemic coronaviruses. *J. Clin. Invest.* 131:e146922. https://doi.org/10.1172/JCI146922

Elbe, S., and G. Buckland-Merrett. 2017. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Chall.* 1:33–46. https://doi.org/10.1002/gch2.1018

Emerson, R.O., W.S. DeWitt, M. Vignali, J. Gravley, J.K. Hu, E.J. Osborne, C. Desmarais, M. Klinger, C.S. Carlson, J.A. Hansen, et al. 2017. Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire. *Nat. Genet.* 49:659–665. https://doi.org/10.1038/ng.3822

Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan. 2010. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66:486–501. https://doi.org/10.1107/S0907444910007493

Evans, P. 2006. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* 62:72–82. https://doi.org/10.1107/S0907444905036693

Gong, F., Y. Dai, T. Zheng, L. Cheng, D. Zhao, H. Wang, M. Liu, H. Pei, T. Jin, D. Yu, and P. Zhou. 2020. Peripheral CD4+ T cell subsets and antibody response in COVID-19 convalescent individuals. *J. Clin. Invest.* 130:6588–6599. https://doi.org/10.1172/JCI141054

Gonzalez-Galarza, F.F., A. McCabe, E.J.M.D. Santos, J. Jones, L. Takeshita, N.D. Ortega-Rivera, G.M.D. Cid-Pavon, K. Ramsbottom, G. Ghattaoraya, A. Alfirevic, et al. 2020. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res.* 48(D1):D783–D788. https://doi.org/10.1093/nar/gkz1029

Grifoni, A., D. Weiskopf, S.I. Ramirez, J. Mateus, J.M. Dan, C.R. Moderbacher, S.A. Rawlings, A. Sutherland, L. Premkumar, R.S. Jadi, et al. 2020. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell.* 181:1489–1501.e15. https://doi.org/10.1016/j.cell.2020.05.015

Hafemeister, C., and R. Satija. 2019. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* 20:296. https://doi.org/10.1186/s13059-019-1874-1

Heit, A., F. Schmitz, S. Gerdts, B. Flach, M.S. Moore, J.A. Perkins, H.S. Robins, A. Aderem, P. Spearman, G.D. Tomaras, et al. 2017. Vaccination establishes clonal relatives of germinal center T cells in the blood of humans. *J. Exp. Med.* 214:2139–2152. https://doi.org/10.1084/jem.20161794

Hill, D.L., W. Pierson, D.J. Bolland, C. Mkindi, E.J. Carr, J. Wang, S. Houard, S.W. Wingett, R. Audran, E.F. Wallin, et al. 2019. The adjuvant GLA-SE promotes human Tfh cell expansion and emergence of public TCRβ clonotypes. *J. Exp. Med.* 216:1857–1873. https://doi.org/10.1084/jem.20190301

Jiang, Y., N. Arase, M. Kohyama, K. Hirayasu, T. Suenaga, H. Jin, M. Matsumoto, K. Shida, L.L. Lanier, T. Saito, and H. Arase. 2013. Transport of misfolded endoplasmic reticulum proteins to the cell surface by MHC class II molecules. *Int. Immunol.* 25:235–246. https://doi.org/10.1093/intimm/dxs155

Kabsch, W. 2010. XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66:125–132. https://doi.org/10.1107/S0907444909047337

Kanda, T., M. Miyata, M. Kano, S. Kondo, T. Yoshizaki, and H. Iizasa. 2015. Clustered microRNAs of the Epstein-Barr virus cooperatively downregulate an epithelial cell-specific metastasis suppressor. *J. Virol.* 89:2684–2697. https://doi.org/10.1128/JVI.03189-14

Kaneko, N., H.H. Kuo, J. Boucau, J.R. Farmer, H. Allard-Chamard, V.S. Mahajan, A. Piechocka-Trocha, K. Lefteri, M. Osborn, J. Bals, et al. Massachusetts Consortium on Pathogen Readiness Specimen Working Group. 2020. Loss of Bcl-6-expressing T follicular helper cells and germinal centers in COVID-19. *Cell.* 183:143–157.e13. https://doi.org/10.1016/j.cell.2020.08.025

Liu, X., R.I. Nurieva, and C. Dong. 2013. Transcriptional regulation of follicular T-helper (Tfh) cells. *Immunol. Rev.* 252:139–145. https://doi.org/10.1111/imr.12040

Liu, L., P. Wang, M.S. Nair, J. Yu, M. Rapp, Q. Wang, Y. Luo, J.F.W. Chan, V. Sahi, A. Figueroa, et al. 2020. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature.* 584:450–456. https://doi.org/10.1038/s41586-020-2571-7

Locci, M., C. Havenar-Daughton, E. Landais, J. Wu, M.A. Kroenke, C.L. Arlehamn, L.F. Su, R. Cubas, M.M. Davis, A. Sette, et al. International AIDS Vaccine Initiative Protocol C Principal Investigators. 2013. Human circulating PD-1+CXCR3-CXCR5+ memory Tfh cells are highly functional and correlate with broadly neutralizing HIV antibody responses. *Immunity.* 39:758–769. https://doi.org/10.1016/j.immuni.2013.08.031

Mateus, J., A. Grifoni, A. Tarke, J. Sidney, S.I. Ramirez, J.M. Dan, Z.C. Burger, S.A. Rawlings, D.M. Smith, E. Phillips, et al. 2020. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science.* 370:89–94. https://doi.org/10.1126/science.abd3871

Matsumoto, Y., K. Kishida, M. Matsumoto, S. Matsuoka, M. Kohyama, T. Suenaga, and H. Arase. 2021. A TCR-like antibody against a proinsulin-containing fusion peptide ameliorates type 1 diabetes in NOD mice. *Biochem. Biophys. Res. Commun.* 534:680–686. https://doi.org/10.1016/j.bbrc.2020.11.019

Meckiff, B.J., C. Ramírez-Suástegui, V. Fajardo, S.J. Chee, A. Kusnadi, H. Simon, S. Eschweiler, A. Grifoni, E. Pelosi, D. Weiskopf, et al. 2020. Imbalance of regulatory and cytotoxic SARS-CoV-2-reactive CD4$^+$ T cells in COVID-19. *Cell.* 183:1340–1353.e16. https://doi.org/10.1016/j.cell.2020.10.001

Ministry of Health, Labour and Welfare. 2020. Japanese COVID-19 Clinical Practice Guideline Version 2. https://www.mhlw.go.jp/content/000650160.pdf (accessed October 7, 2021)

Ni, L., F. Ye, M.L. Cheng, Y. Feng, Y.Q. Deng, H. Zhao, P. Wei, J. Ge, M. Gou, X. Li, et al. 2020. Detection of SARS-CoV-2-Specific Humoral and Cellular Immunity in COVID-19 Convalescent Individuals. *Immunity.* 52:971–977.e3. https://doi.org/10.1016/j.immuni.2020.04.023

Nolan, S., M. Vignali, M. Klinger, J.N. Dines, I.M. Kaplan, T. Craft, K. Boland, M. Mazza, and K. Dobbs. 2020. A large-scale database of T-cell receptor beta (TCR β) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2. Res. Sq. https://doi.org/10.21203/rs.3.rs-51964/v1 (Preprint posted August 4, 2020)

Peng, Y., A.J. Mentzer, G. Liu, X. Yao, Z. Yin, D. Dong, W. Dejnirattisai, T. Rostron, P. Supasa, C. Liu, et al. ISARIC4C Investigators. 2020. Broad and strong memory CD4$^+$ and CD8$^+$ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat. Immunol.* 21:1336–1345. https://doi.org/10.1038/s41590-020-0782-6

Reynisson, B., C. Barra, S. Kaabinejadian, W.H. Hildebrand, B. Peters, and M. Nielsen. 2020. Improved prediction of MHC II antigen presentation through integration and motif deconvolution of mass spectrometry MHC eluted ligand data. *J. Proteome Res.* 19:2304–2315. https://doi.org/10.1021/acs.jproteome.9b00874

Rydyznski Moderbacher, C., S.I. Ramirez, J.M. Dan, A. Grifoni, K.M. Hastie, D. Weiskopf, S. Belanger, R.K. Abbott, C. Kim, J. Choi, et al. 2020. Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell.* 183:996–1012.e19. https://doi.org/10.1016/j.cell.2020.09.038

Sadoff, J., G. Gray, A. Vandebosch, V. Cárdenas, G. Shukarev, B. Grinsztejn, P.A. Goepfert, C. Truyers, H. Fennema, B. Spiessens, et al. ENSEMBLE Study Group. 2021. Safety and efficacy of single-dose Ad26.COV2.S vaccine against Covid-19. *N. Engl. J. Med.* 384:2187–2201. https://doi.org/10.1056/NEJMoa2101544

Sauer, M.M., M.A. Tortorici, Y.J. Park, A.C. Walls, L. Homad, O.J. Acton, J.E. Bowen, C. Wang, X. Xiong, W. de van der Schueren, et al. 2021. Structural basis for broad coronavirus neutralization. *Nat. Struct. Mol. Biol.* 28:478–486. https://doi.org/10.1038/s41594-021-00596-4

Schmitt, N., S.E. Bentebibel, and H. Ueno. 2014. Phenotype and functions of memory Tfh cells in human blood. *Trends Immunol.* 35:436–442. https://doi.org/10.1016/j.it.2014.06.002

Sefik, E., N. Geva-Zatorsky, S. Oh, L. Konnikova, D. Zemmour, A.M. McGuire, D. Burzyn, A. Ortiz-Lopez, M. Lobera, J. Yang, et al. 2015. MUCOSAL IMMUNOLOGY. Individual intestinal symbionts induce a distinct population of RORγ$^+$ regulatory T cells. *Science.* 349:993–997. https://doi.org/10.1126/science.aaa9420

Sette, A., and S. Crotty. 2021. Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell.* 184:861–880. https://doi.org/10.1016/j.cell.2021.01.007

Shinde, V., S. Bhikha, Z. Hoosain, M. Archary, Q. Bhorat, L. Fairlie, U. Lalloo, M.S.L. Masilela, D. Moodley, S. Hanley, et al. 2019nCoV-501 Study Group. 2021. Efficacy of NVX-CoV2373 Covid-19 vaccine against the B.1.351 variant. *N. Engl. J. Med.* 384:1899–1909. https://doi.org/10.1056/NEJMoa2103055

Shomuradova, A.S., M.S. Vagida, S.A. Sheetikov, K.V. Zornikova, D. Kiryukhin, A. Titov, I.O. Peshkova, A. Khmelevskaya, D.V. Dianov, M. Malasheva, et al. 2020. SARS-CoV-2 epitopes are recognized by a public and diverse repertoire of human T cell receptors. *Immunity.* 53:1245–1257.e5. https://doi.org/10.1016/j.immuni.2020.11.004

Singer, J.B., R.J. Gifford, M. Cotten, and D.L. Robertson. 2020. CoV-GLUE: a web application for tracking SARS-CoV-2 genomic variation. *Preprints.* https://doi.org/10.20944/preprints202006.0225.v1 (Preprint posted June 16, 2020)

Tan, A.T., M. Linster, C.W. Tan, N. Le Bert, W.N. Chia, K. Kunasegaran, Y. Zhuang, C.Y.L. Tham, A. Chia, G.J.D. Smith, et al. 2021. Early induction of functional SARS-CoV-2-specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. *Cell Rep.* 34:108728. https://doi.org/10.1016/j.celrep.2021.108728

Vagin, A., and A. Teplyakov. 2010. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* 66:22–25. https://doi.org/10.1107/S0907444909042589

Vagin, A.A., R.A. Steiner, A.A. Lebedev, L. Potterton, S. McNicholas, F. Long, and G.N. Murshudov. 2004. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* 60:2184–2195. https://doi.org/10.1107/S0907444904023510

Vinuesa, C.G., M.A. Linterman, D. Yu, and I.C.M. MacLennan. 2016. Follicular helper T cells. *Annu. Rev. Immunol.* 34:335–368. https://doi.org/10.1146/annurev-immunol-041015-055605

Vita, R., S. Mahajan, J.A. Overton, S.K. Dhanda, S. Martini, J.R. Cantrell, D.K. Wheeler, A. Sette, and B. Peters. 2019. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* 47(D1):D339–D343. https://doi.org/10.1093/nar/gky1006

Voysey, M., S.A. Costa Clemens, S.A. Madhi, L.Y. Weckx, P.M. Folegatti, P.K. Aley, B. Angus, V.L. Baillie, S.L. Barnabas, Q.E. Bhorat, et al. Oxford COVID Vaccine Trial Group. 2021. Single-dose administration and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: a pooled analysis of four randomised trials. *Lancet.* 397:881–891. https://doi.org/10.1016/S0140-6736(21)00432-3

Walsh, E.E., R.W. Frenck Jr., A.R. Falsey, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, K. Neuzil, M.J. Mulligan, R. Bailey, et al. 2020. Safety and immunogenicity of two RNA-based Covid-19 vaccine candidates. *N. Engl. J. Med.* 383:2439–2450. https://doi.org/10.1056/NEJMoa2027906

Wang, P., M.S. Nair, L. Liu, S. Iketani, Y. Luo, Y. Guo, M. Wang, J. Yu, B. Zhang, P.D. Kwong, et al. 2021. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature.* 593:130–135. https://doi.org/10.1038/s41586-021-03398-2

Williams, C.J., J.J. Headd, N.W. Moriarty, M.G. Prisant, L.L. Videau, L.N. Deis, V. Verma, D.A. Keedy, B.J. Hintze, V.B. Chen, et al. 2018. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* 27:293–315. https://doi.org/10.1002/pro.3330

Wrapp, D., N. Wang, K.S. Corbett, J.A. Goldsmith, C.L. Hsieh, O. Abiona, B.S. Graham, and J.S. McLellan. 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science.* 367:1260–1263. https://doi.org/10.1126/science.abb2507

Zhang, J., Q. Wu, Z. Liu, Q. Wang, J. Wu, Y. Hu, T. Bai, T. Xie, M. Huang, T. Wu, et al. 2021. Spike-specific circulating T follicular helper cell and cross-neutralizing antibody responses in COVID-19-convalescent individuals. *Nat. Microbiol.* 6:51–58. https://doi.org/10.1038/s41564-020-00824-5

Supplemental material

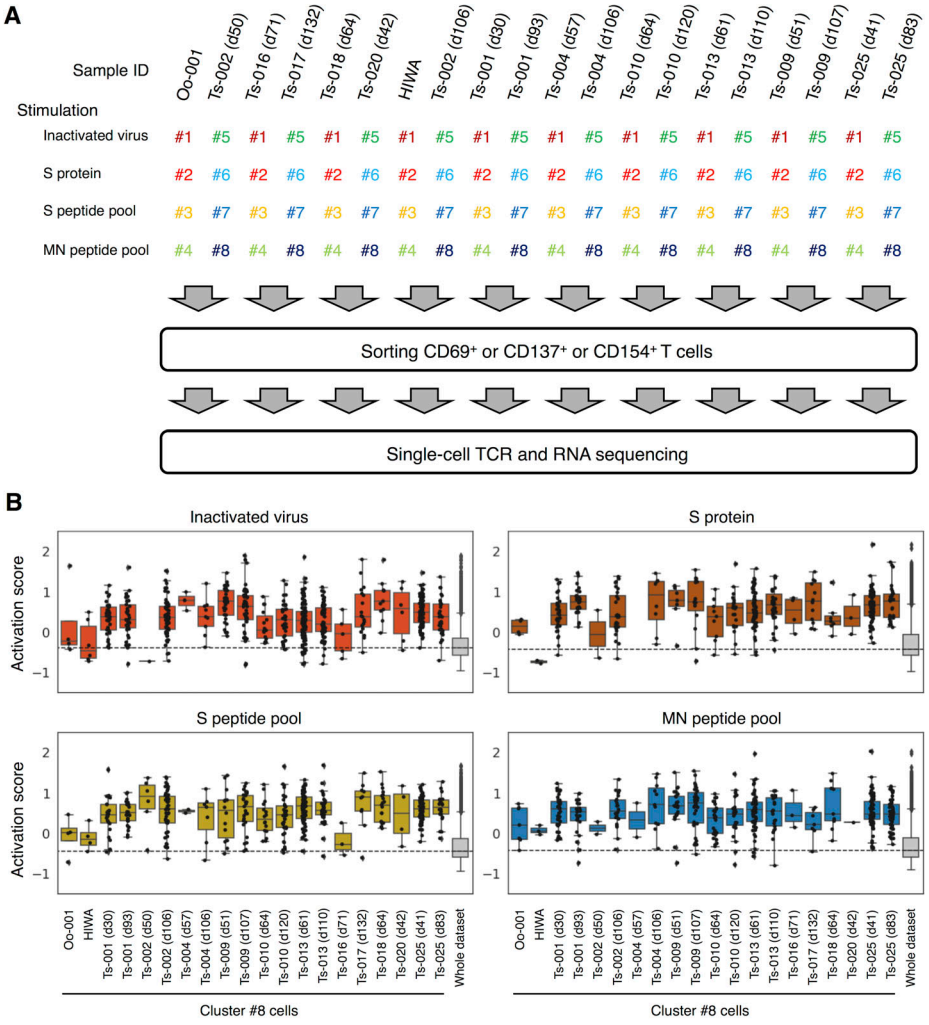


Figure S1. **Single-cell–based analyses of SARS-CoV-2–responsive T cells.** PBMCs from healthy donors and convalescent COVID-19 patients were stimulated with SARS-CoV-2–derived antigens as indicated for 20 h before antigen-responding T cells were sorted and analyzed by single-cell TCR- and RNA-seq. **(A)** A workflow of single-cell–based analyses. Antigen-stimulated PBMCs were stained with eight different hashtag antibodies (#1–8), anti-CD3, and the indicated activation marker antibodies. Hashtagged cells from two patients were pooled. Antigen-responding ($CD69^+$ or $CD137^+$ or $CD154^+$) $CD3^+$ T cells were sorted and immediately subjected to single-cell TCR- and RNA-seq. Each collection day after diagnosis is shown after the donor identifier (ID). **(B)** Activation score of each single cell was calculated based on the expression of activation genes (*CD69*, *TNFRSF9*, *TNFRSF4*, and *CD40LG*).

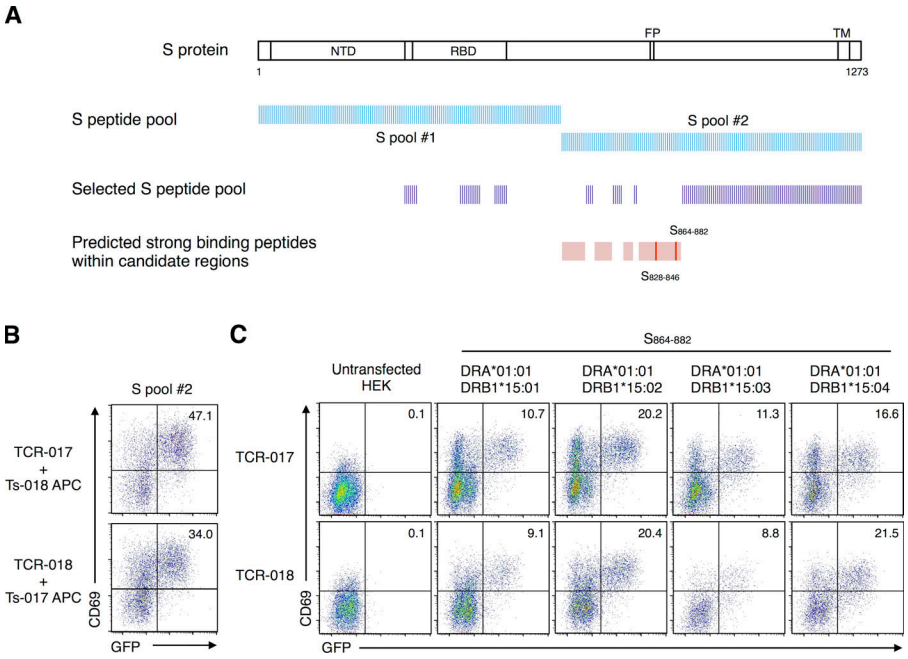


Figure S2. **Epitope peptide and restricting HLAs of clonotype-017/018. (A)** Relative positions within the S protein that are covered by each S peptide pool are shown. Strong binding peptides (%RANK ≤2) presented by DRA-DRB1*15:01 in regions 633–682, 708–740, 771–784, and 803–884 (pink regions) were predicted using NetMHC 4.0 Server. 19-mer peptides (red bars) that cover the strong binding peptides were synthesized and used for further epitope identification. **(B)** Cells expressing TCR-017 or -018 were stimulated with 1 µg/ml of S peptide pool #2 in the presence of heterologous APCs from exchanged donors (Ts-018 and Ts-017, respectively). **(C)** Cells expressing TCR-017 or -018 were stimulated with 1 µg/ml of peptide $S_{864–882}$ in the presence of HEK293T cells expressing the indicated HLAs. Data are representative of two independent experiments (B and C). NTD, N-terminal domain; FP, fusion peptide; TM, transmembrane domain.

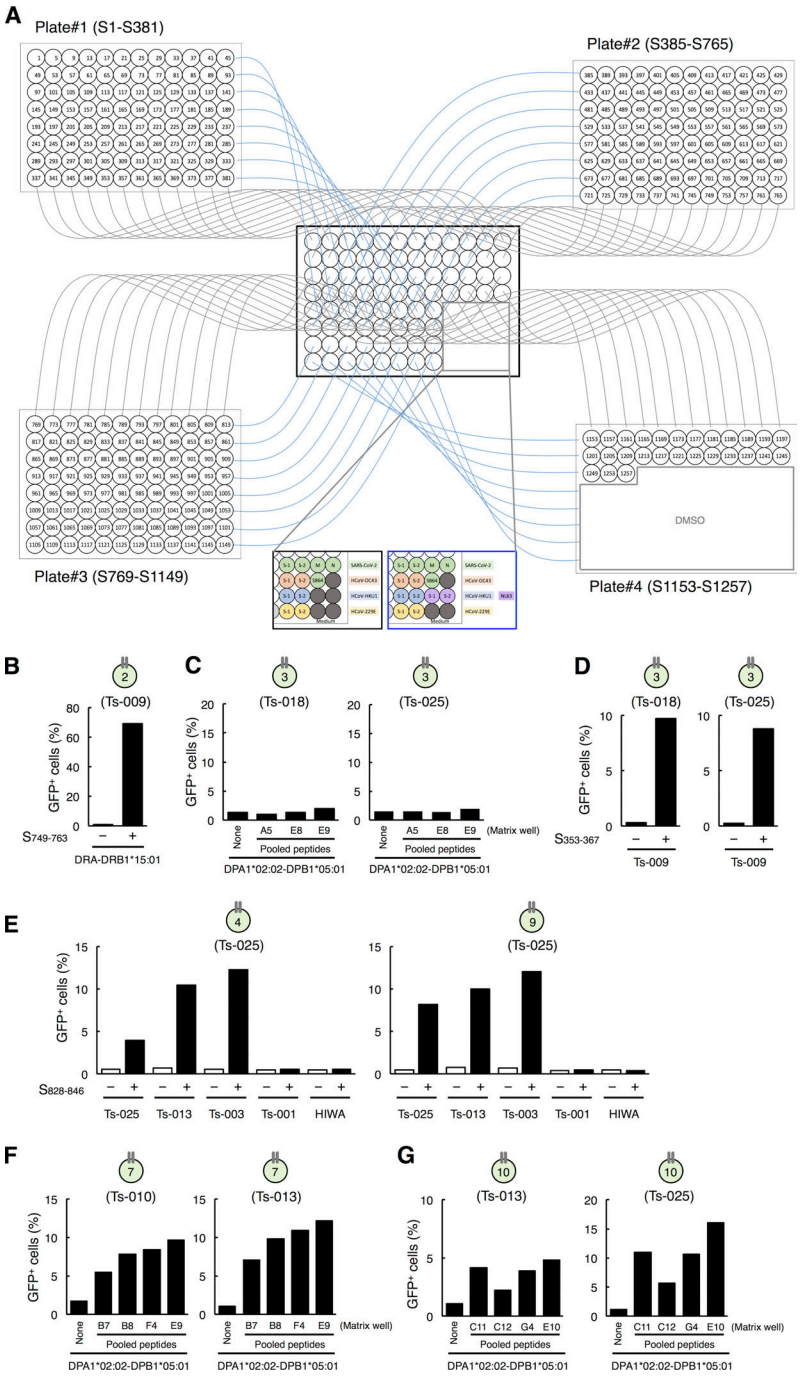


Figure S3. **Identification of the epitopes of public clonotypes. (A)** A scheme of the rapid epitope determination platform used for identification of the epitopes for T cell clonotypes. 15-mer peptides with 11-amino acid overlap that cover the full length of the S protein of SARS-CoV-2 were synthesized and pooled as described in Materials and methods. Each peptide is present in a unique combination of two different pools; thus, the epitope can be revealed by the common peptides in the pools that activate the clonotype. Peptide pools of S, M, and N proteins of SARS-CoV-2 or HCoV S proteins were also added to the matrix as controls. S864 indicates $S_{864-882}$ peptide. **(B)** Identification of DRB1*15:01 as a restricting HLA of $S_{753-759}$ for clonotype 2. Reporter cells expressing clonotype 2 were activated by $S_{749-763}$ in the presence of HEK293T cells transfected with DRA*01:01-DRB1*15:01. Data are shown as mean ± SD of triplicates. **(C)** Exclusion of DPA1*02:02-DPB1*05:01 as a restricting HLA of $S_{353-367}$ for clonotype 3. Reporter cells expressing clonotype 3 were not activated by S353-including peptide pools, such as A5, E8, and E9 (S pool #1) in A, in the presence of HEK293T cells transfected with DPA1*02:02-DPB1*05:01. **(D)** DRB4*01:03 is a restricting HLA of $S_{353-367}$ for clonotype 3. Reporter cells expressing clonotype 3 were activated by S353 in the presence of B cells from Ts-009. **(E)** DQA1*01:01-DQB1*05:01 is a restricting HLA of $S_{833-843}$ for clonotypes 4 and 9. Reporter cells expressing clonotype 4 or 9 were activated by $S_{828-846}$ in the presence of B cells from donors Ts-025 (autologous), Ts-013, and Ts-003 but not activated in the presence of B cells from donors Ts-001 and HIWA. **(F)** DPA1*02:02-DPB1*05:01 is a restricting HLA of $S_{557-567}$ for clonotype 7. Reporter cells expressing clonotype 7 were activated by S553- or S557-including peptide pools, such as B7, B8, F4, and E9 (S pool #1) in A, in the presence of HEK293T cells transfected with DPA1*02:02-DPB1*05:01. **(G)** DPA1*02:02-DPB1*05:01 is a restricting HLA of $S_{957-967}$ for clonotype 10. Reporter cells expressing clonotype 10 were activated by S953- or S957-including peptide pools, such as C11, C12, G4, and E10 (S pool #2) in A, in the presence of HEK293T cells transfected with DPA1*02:02-DPB1*05:01. Data are representative of two independent experiments (B–G).

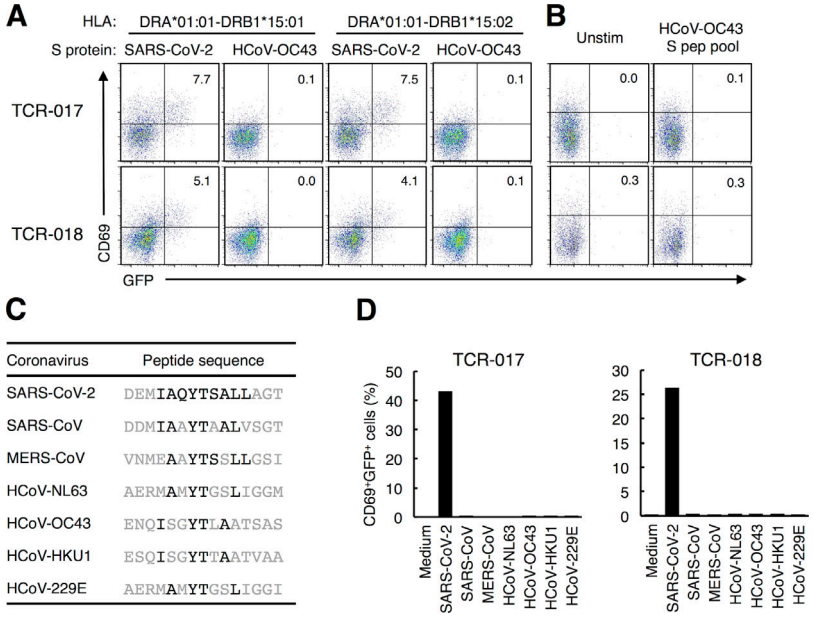


Figure S4. **Cross-reactivity of public Tfh clonotypes. (A)** Cells expressing TCR-017 or -018 were cocultured with HEK293T cells expressing the indicated HLAs and S proteins derived from the indicated CoVs. **(B)** Cells expressing TCR-017 or -018 were stimulated with 1 µg/ml of peptide pool derived from HCoV-OC43 S protein. **(C)** Sequences of peptides from HCoVs corresponding to $S_{867–881}$ of SARS-CoV-2. Amino acids that are the same as those in SARS-CoV-2 in the core epitope are shown in black. **(D)** Cells expressing TCR-017 or -018 were stimulated with 1 µg/ml of the peptides shown in C. Data are representative of two independent experiments (A, B, and D). pep, peptide; Unstim, unstimulated.

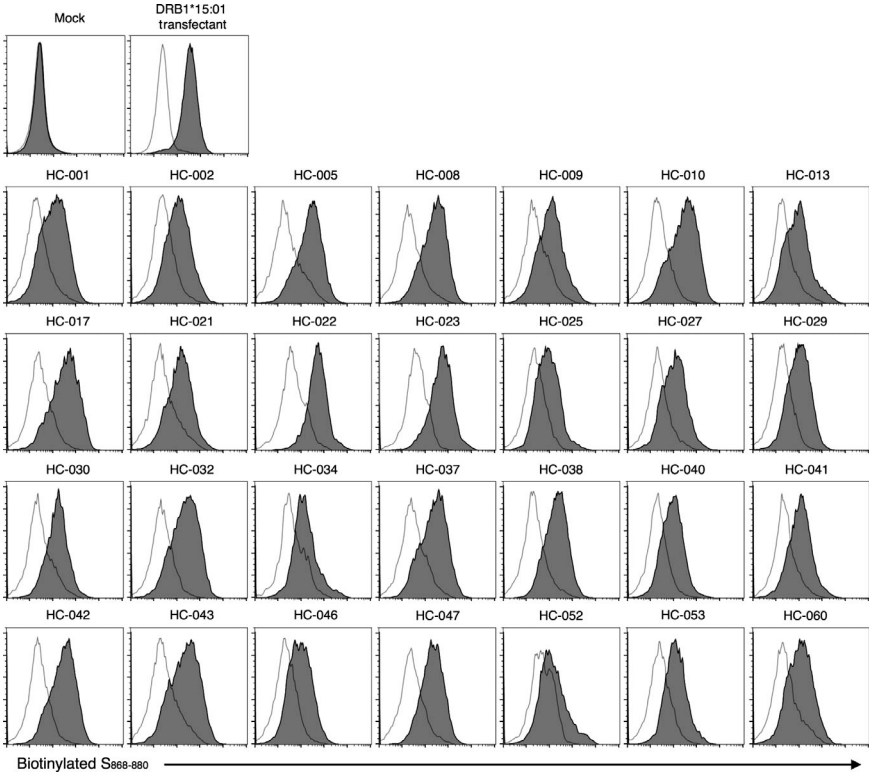


Figure S5. **Presentation of $S_{868–880}$ epitope on multiple HLA alleles.** Transformed B cells from healthy donors described in Fig. 6, F and G, were stained with (filled histogram) or without (open histogram) 10 µg/ml biotinylated $S_{868–880}$ peptide with a GSGSGS linker at the N-terminus at 37°C for 12 h followed by adding streptavidin-phycoerythrin. DRA-DRB1*15:01 transfectant was used as a control. Data are representative of two independent experiments.

Table S1, Table S2, Table S5, Table S6, and Table S7 are provided online as separate Word files. Table S3 and Table S4 are provided online as separate Excel files. Table S1 shows donor characteristics. Table S2 shows HLA class II types of involved blood donors. Table S3 shows cTfh clonotypes identified in single-cell analysis. Table S4 shows $S_{864–882}$-reactive public cTfh clonotypes. Table S5 shows frequencies of HLAs in the Japanese population that are predicted to present epitopes of public clonotypes. Table S6 shows predicted binding alleles to $S_{864–882}$ in healthy donors. Table S7 shows data collection and refinement statistics of the crystallographic analysis of TCRαβ-017.