Mutations in SARS-CoV-2 spike protein impair epitope-specific CD4+ T cell recognition

Emily X. C. Tye[1], Elizabeth Jinks[1], Tracey A. Haigh [1], Baksho Kaul[1], Prashant Patel [2], Helen M. Parry[1], Maddy L. Newby[3], Max Crispin [3], Nayandeep Kaur[1], Paul Moss [1], Samantha J. Drennan[1], Graham S. Taylor [1] & Heather M. Long [1] ✉

CD4+ T cells are essential for protection against viruses, including SARS-CoV-2. The sensitivity of CD4+ T cells to mutations in SARS-CoV-2 variants of concern (VOCs) is poorly understood. Here, we isolated 159 SARS-CoV-2-specific CD4+ T cell clones from healthcare workers previously infected with wild-type SARS-CoV-2 (D614G) and defined 21 epitopes in spike, membrane and nucleoprotein. Lack of CD4+ T cell cross-reactivity between SARS-CoV-2 and endemic beta-coronaviruses suggested these responses arose from naïve rather than pre-existing cross-reactive coronavirus-specific T cells. Of the 17 epitopes located in the spike protein, 10 were mutated in VOCs and CD4+ T cell clone recognition of 7 of them was impaired, including 3 of the 4 epitopes mutated in omicron. Our results indicated that broad targeting of epitopes by CD4+ T cells likely limits evasion by current VOCs. However, continued genomic surveillance is vital to identify new mutations able to evade CD4+ T cell immunity.

Coordinated adaptive immunity is essential for protection and clearance of viral infections, including SARS-CoV-2 (ref. [1]). Virus-specific neutralizing antibodies are considered the main correlate of protection against SARS-CoV-2 infection, but wane over time[2,3]. T cell responses are more durable[2,4,5] and increasing evidence supports their role in restricting SARS-CoV-2 infection and limiting the severity of COVID-19[6,7]. Worldwide efforts have rapidly delivered SARS-CoV-2 vaccines, mostly designed against the spike (S) protein, which mediates host cell entry. Studies enumerating the T cell response to whole SARS-CoV-2 S protein using pools of overlapping peptides covering the entire protein sequence (peptide mixes) show that memory T cell responses to S protein in previously infected or vaccinated individuals are dominated by CD4+ T cells[4,8–10].

SARS-CoV-2 CD4+ T cell epitopes have been identified, but mostly in assays that use high concentrations of stimulating peptides. In addition, their human leukocyte antigen (HLA) restriction has largely been inferred from in silico HLA-binding algorithms[11,12]. Detailed knowledge of the specificity of CD4+ T cell responses at the epitope level and their HLAII restriction is therefore currently lacking[13]. Furthermore, whether the CD4+ T cell epitopes are generated naturally through the HLA class II (HLAII) antigen processing pathway is currently unknown. Because of these limitations, the extent to which CD4+ T cells that recognize SARS-CoV-2 epitopes cross-react with other human β-coronaviruses (β-HCoVs) remains unknown. Low frequency CD4+ T cell reactivity to SARS-CoV-2 peptide mixes has been reported in some uninfected individuals[14–18] and has been suggested to originate from previous exposure to other HCoVs (HKU1, OC43, NL63, 229E, SARS or MERS). This raises the possibility that pre-existing HCoV immunity could potentially contribute to controlling SARS-CoV-2 infection.

The extent to which T cells induced by ancestral SARS-CoV-2 proteins can protect against VOCs is a critical question. In particular, the highly transmissible omicron VOC contains several mutations within

[1]Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham, UK. [2]Institute of Cancer and Genomics, University of Birmingham, Birmingham, UK. [3]School of Biological Sciences, University of Southampton, Southampton, UK. ✉e-mail: h.m.long@bham.ac.uk
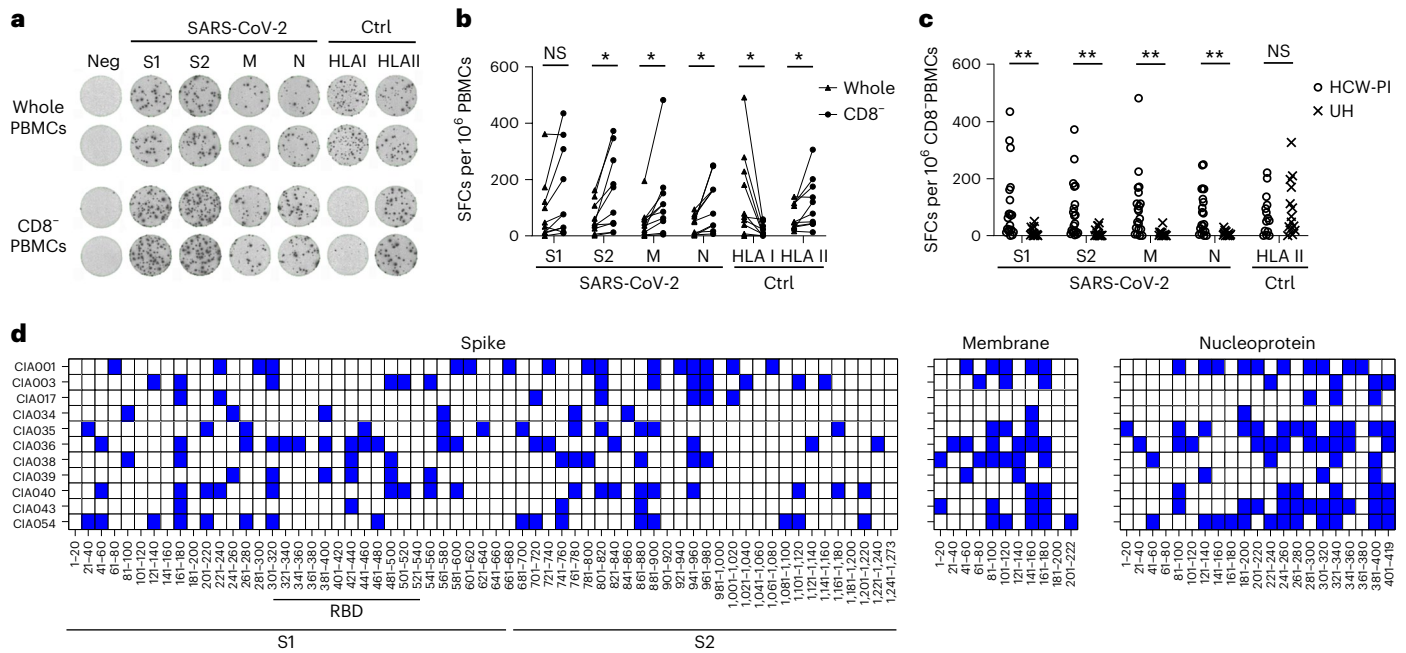
Fig. 1 | CD4+ T cell response to SARS-CoV-2. a, Representative Elispot assays for the production on IFN-γ in whole and CD8−PBMCs isolated from HCW-PI, plated at 2 × 10⁵ cells per well and incubated with mixes of 15 aa peptides (overlapping by 11 aa) from SARS-CoV-2 S1, S2, M and N proteins, control (Ctrl) HLAI and HLAII epitope mixes and DMSO solvent (Neg). **b**, Summary of Elispot assays for the production of IFN-γ in whole versus CD8−PBMCs using n = 10 HCW-PI incubated with peptide mixes as in **a**. Results are shown as mean SFC per 10⁶ PBMCs. Significance was determined by two-sided Wilcoxon test, *P < 0.05. **c**, Summary of

Elispot assays for the production of IFN-γ in CD8−PBMCs from HCW-PI (n = 20) and UH (n = 14) individuals incubated with peptide mixes as in **a**. Results are shown as mean SFC per 10⁶ CD8−PBMCs. Significance was determined by two-sided Mann–Whitney U-test, **P < 0.01. **d**, Summary of ELISAs measuring IFN-γ production by polyclonal CD4+ T cell lines generated by initial stimulation of CD8−PBMCs with peptide mixes as in **a**, then stimulated 7–14 days later with individual 20mer peptides (overlapping by 10 aa) spanning the relevant SARS-CoV-2 protein. Individual rows show the response from HCW-PI, n = 11. NS, not significant.

the receptor binding domain (RBD) of S protein—the main target of neutralizing antibodies[19,20]. These mutations reduce neutralization by the S-specific antibodies induced by ancestral SARS-CoV-2 variants or by the initial vaccines deployed in the pandemic, which were designed against the original S protein reference sequence[21–23]. Whilst neutralizing antibody titers can be partially restored by booster vaccination[24,25], continued virus evolution has caused a high prevalence of secondary and vaccine breakthrough infections[26]. Ex vivo studies of previously infected or vaccinated individuals using peptide mixes have shown minimal reduction in the overall frequency of CD4+ T cell responses against the alpha (B.1.1.7), beta (B1.351), gamma (P1) and delta (B.1.717.2)[23,27–29] or omicron (B.1.1.529) VOCs[22,30,31]. However, biologically relevant differences in epitope-specific recognition efficiency may have been missed[32] and little information exists to understand the extent of CD4+ T cell epitopes evaded by current VOCs or to predict CD4+ T cell epitope loss in future SARS-CoV-2 variants.

Here, we performed a detailed analysis of CD4+ T cell immunity against SARS-CoV-2 in healthcare workers (HCW) infected in the first wave of the pandemic. We examined 159 CD4+ T cell clones and identified and characterized 21 HLAII-restricted T cell epitopes. Responses to most epitopes located in the S protein were also present in vaccinated individuals of appropriate HLAII genotype. Cross-reactivity of SARS-CoV-2 S-specific T cell clones was observed only against the closely related S2 region of the SARS virus, with no cross-reactivity observed for any other β-HCoV. Mutations were present in 10 of the 17 S protein epitopes within one or more SARS-CoV-2 VOC. Minor amino acid (aa) changes in seven epitope sequences, including those within the RBD region of omicron, were sufficient to reduce or evade recognition by S-specific CD4+ T cells. However, the breadth of responses to several CD4+ T cell epitopes seen in each individual suggested that current VOC mutations confer only limited evasion from CD4+ T cell surveillance.

Results

SARS-CoV-2 infection induces broad CD4+ T cell immunity

Peripheral blood mononuclear cells (PBMCs) were collected from June to September 2020 from 20 HCW 3–6 months postinfection (PI) during the first wave of SARS-CoV-2 wild-type (WT, D614G) infection in the United Kingdom and 14 uninfected healthy (UH) volunteers. HCW-PI had detectable antibodies to S protein and nucleoprotein (N) at this timepoint, while UH volunteers had no detectable S- or N-specific antibody responses (Extended Data Fig. 1). To characterize T cell immunity against whole antigens, PBMCs and CD8-depleted PBMCs (hereafter CD8−PBMCs) were tested against SARS-CoV-2 peptide mixes comprising 15mer peptides overlapping by 11 aa and spanning the entire open reading frame of S, membrane (M) and N proteins in ex vivo interferon-γ (IFN-γ) Elispot assays. Compared with control whole PBMCs, as expected, the response to a peptide mix of HLA class I (HLAI)-restricted epitopes was significantly lower in CD8−PBMCs (Fig. 1a,b). In contrast, responses to S protein (tested as two pools, S1 and S2), M and N peptide mixes were increased in CD8−PBMCs (Fig. 1a,b), confirming CD4+ T cell memory responses were predominant[4,8]. As previously reported[14–18], we detected weak responses to individual SARS-CoV-2 protein pepmixes in CD8−PBMCs from 8 of 14 UH volunteers (Fig. 1c). The magnitude of responses in UH volunteers was significantly lower than in HCW-PI (P < 0.01, two-sided Mann–Whitney U-test) (Fig. 1c) and may represent cross-reactive T cells primed by previous exposure to β-HCoVs.

We next examined the CD4+ T cell response to S, M and N proteins following SARS-CoV-2 infection at the epitope level. Polyclonal CD4+ T cell lines were initially established from HCW-PI by stimulating PBMCs with S, M and N peptide mixes, ensuring complete coverage of the proteins. These lines were then tested with individual 20mer peptides overlapping by 10 aa, or dimethylsulfoxide (DMSO) solvent as negative control, to determine the regions of reactivity against each
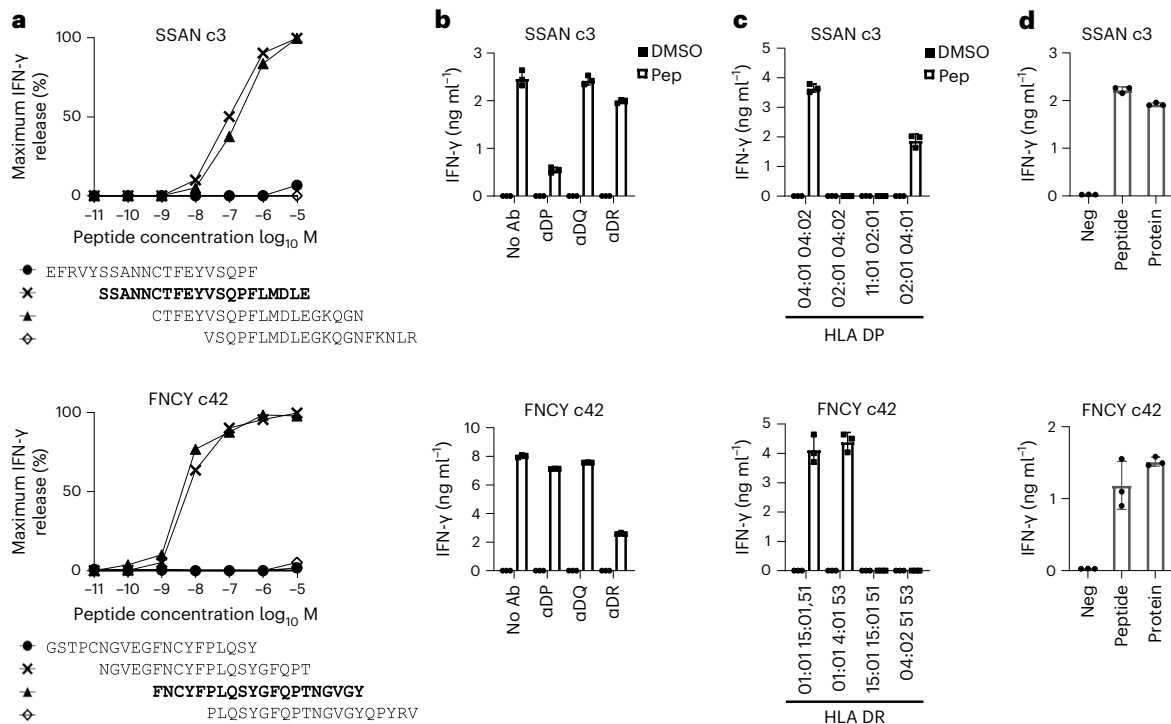
Fig. 2 | Characterization of new spike CD4⁺ T cell epitopes. a, ELISA assays for the production of IFN-γ from CD4⁺ T cell clones cocultured in overnight assays with autologous LCL loaded with individual 20mer peptides overlapping by 15 aa (10^{-5} to 10^{-11}M). **b–d**, ELISA assays for the production of IFN-γ from CD4⁺ T cell clones cocultured in overnight assays with autologous LCL prepulsed with epitope peptide or DMSO solvent and either tested alone (no antibody; No Ab), or in the presence of blocking antibodies against HLA-DP (αDP), HLA-DQ (αDQ) or HLA-DR (αDR) (**b**), autologous LCL and allogeneic LCLs with HLAII types partially matched to the autologous LCL, either prepulsed with 5 μM 20mer epitope peptide or DMSO solvent (Neg) (**c**) or autologous LCL either prepulsed with epitope DMSO solvent (Neg), peptide or 1 ng ml⁻¹S tetrameric protein (**d**). (**a–d**) Results show mean IFNγ release ±1 s.d. and are representative of three experiments.

protein. CD4⁺ T cells can be expanded using shorter 15mer peptides; however, the optimal response is often to a longer peptide[33,34]. For each individual, the polyclonal CD4⁺ T cell lines contained several responses against peptides distributed throughout the S protein, including the RBD, S1 and S2 regions, M and N (Fig. 1d). As noted by others, these data show that SARS-CoV-2 induced a broad CD4⁺ T cell response[11,14].

HLAII type determines SARS-CoV-2 CD4⁺ T cell epitope responses

We next performed limiting dilution cloning from five individuals with different HLAII types and isolated CD4⁺ T cell clones specific for 21 epitopes (17 S protein, 2 M protein and 2 N protein epitopes). To identify the optimal peptide recognized by each T cell clone, we titrated four individual 20mer peptides (overlapping by 15 aa) covering the regions where T cell reactivity was detected in the screening assays. This defined peptides SSAN (aa 161–180) and FNCY (aa 486–505) as the epitopes for the CD4⁺ T cell clones c3 and c42, respectively (Fig. 2a). Peptides RGHL and RNSS were defined using the same approach in M and N respectively (Extended Data Fig. 2). The avidities of the S, M and N specific T cell clones were comparable with CD4⁺ T cell clones against other viruses previously measured by peptide titration (Fig. 2 and Extended Data Fig. 2)[35–37].

Next, to identify the HLAII allele restricting each epitope, we tested each clone against a peptide-loaded autologous lymphoblastoid cell line (LCL) in the presence of blocking antibodies against HLA-DR, HLA-DP and HLA-DQ, and after that against peptide-loaded allogeneic LCLs with partially matched HLA-DR, -DP or -DQ types. This approach indicated that HLA-DPB1*04:01 restricted the presentation of peptide SSAN to clone c3, while HLA-DRB1*01:01 presented peptide FNCY to clone c42 (Fig. 2b,c). We then defined the optimal peptide and presenting HLAII allele for all 21 S, M and N epitopes (Table 1), using the clones

isolated against each epitope (Supplementary Table 1). Individual T cell clones specific for the same epitope recognized the same optimal peptide and HLAII combination (Extended Data Fig. 3). The T cell clones specific for all 17 S protein epitopes recognized autologous LCL pre-exposed to low concentrations of purified S protein (1 ng ml⁻¹) (Fig. 2d and Extended Data Fig. 4), indicating that all S protein epitopes were efficiently generated through the exogenous HLAII processing pathway. All 21 S, M and N epitopes were presented by HLA-DR or HLA-DP alleles, with no HLA-DQ-restricted T cells found (Table 1).

To address whether natural infection and vaccination elicited similar CD4⁺ T cell immunity, we repeated the ex vivo Elispots assays for the production of IFN-γ by CD8⁻PBMCs from each HCW-PI using the defined epitope peptides appropriate to each individual's HLAII type. Of the 11 epitopes presented by HLAII alleles present in several donors, responses to 9 were present in more than one donor (Table 1). Next, we examined the relative immunogenicity of SARS-CoV-2 vaccination by testing blood samples collected 1–5 months postvaccine from 9 of the 14 donors originally used as UH. All nine donors had S-specific antibodies, but undetectable N-specific antibodies (Extended Data Fig. 5), indicating they had responded to vaccination and had no history of natural infection. Based on the HLAII genotypes of these donors, we tested 14 S epitopes and detected responses to 13 in ex vivo IFNγ Elispot assays (Table 1). Collectively, these results indicated that SARS-CoV-2 infection and vaccination induced broad CD4⁺ T cell responses to shared epitopes and that, in both contexts, HLAII genotype was a key determinant of the SARS-CoV-2 S-specific CD4⁺ T cell response.

Spike-specific CD4⁺ T cell clones do not cross-recognize β-HCoVs

Next we asked whether spike-specific CD4⁺ T cells elicited by SARS-COV-2 infection cross-reacted with closely related β-HCoVs

Table 1 | SARS-CoV-2 CD4⁺ T cell epitopes and responses

Protein	aa coordinates	Epitope	HLAII restriction	HCW-PI	UH-PV
Spike	71–90	SGTNGTKRFDNPVLPFNDGV	DPB1*05:01	2/2	NT
Spike	161–180	SSANNCTFEYVSQPFLMDLE	DPB1*04:01	2/2	1/2
Spike	226–245	LVDLPIGINITRFQTLLALH	DRB1*04:01, *04:03, *04:04	1/1	1/2
Spike	236–255	TRFQTLLALHRSYLTPGDSS	DRB1*04:02	1/1	0/1
Spike	296–315	LSETKCTLKSFTVEKGIYQT	DRB1*04:01, *04:04	1/2	2/2
Spike	446–465	GGNYNYLYRLFRKSNLKPFE	DPB1*02:01	3/3	1/2
Spike	486–505	FNCYFPLQSYGFQPTNGVGY	DRB1*01:01	2/4	1/2
Spike	511–530	VVLSFELLHAPATVCGPKKS	DRB1*01:01	3/4	2/3
Spike	746–765	STECSNLLLQYGSFCTQLNR	DRB1*15:01	2/2	5/5
Spike	801–820	NFSQILPDPSKPSKRSFIED	DRB1*04:01, *04:03, *04:04	2/2	2/2
Spike	816–835	SFIEDLLFNKVTLADAGFIK	DRB1*08:01	1/1	2/2
Spike	891–910	GAALQIPFAMQMAYRFNGIG	DRB1*07:01	3/5	2/3
Spike	956–975	AQALNTLVKQLSSNFGAISS	DRB1*01:01	2/4	2/3
Spike	961–980	TLVKQLSSNFGAISSVLNDI	DRB1*04:01, *04:04	1/1	2/2
Spike	1,006–1,025	TYVTQQLIRAAEIRASANLA	DRB1*04:02	1/1	NT
Spike	1,011–1,030	QLIRAAEIRASANLAATKMS	DRB1*04:01, *04:04	1/2	1/2
Spike	1,061–1,080	VFLHVTYVPAQEKNFTTAPA	DRB1*04:02	1/1	NT
Membrane	146–165	RGHLRIAGHHLGRCDIKDLP	DRB4*01:03	NT	N/A
Membrane	161–180	IKDLPKEITVATSRTLSYYK	DRB1*04:02	NT	N/A
Nucleoprotein	196–215	RNSSRNSTPGSSRGTSPARM	DPB1*09:01	NT	N/A
Nucleoprotein	281–300	QTQGNFGDQELIRQGTDYKH	DRB4*01:03	NT	N/A

HCW-PI, healthcare workers postinfection; UH-PV, uninfected healthy individuals postvaccine; NT, not tested; N/A, not applicable

known to infect humans (SARS, MERS, HKU1 and OC43), in which the S proteins share 34–76% aa similarity to SARS-CoV-2 S protein (Fig. 3a). Within the N-terminal S1 region, the highest similarity (64.7%) is between SARS and SARS-CoV-2, whilst the similarity of all other β-HCoV S1 regions with SARS-CoV-2 S1 is low (<32%) (Fig. 3b). The C-terminal S2 regions exhibit greater overall similarity, with SARS-CoV-2 and SARS having 90.0% similarity, and the similarity between SARS-CoV-2 and the other β-HCoV is up to 45%[38] (Fig. 3b).

CD4⁺ T cell clones specific for six SARS-CoV-2 S epitopes were tested against peptide mixes comprising 15mer peptides overlapping by 11 aa from each β-HCoV or DMSO solvent as a negative control. As expected, all six S protein-specific clones showed similar recognition against S1 or S2 SARS-CoV-2 peptide mix and the respective cognate epitope peptide of that clone (Fig. 3c,d). The SSAN c3, VVLS c21 and FNCY c42 T cell clones specific for epitopes located within the more divergent S1 region of SARS-CoV-2 did not cross-recognize any other β-HCoV peptide mix (Fig. 3c). These SARS-CoV-2 20mer epitopes had between 2 and 10 aa differences compared with the corresponding regions of SARS (Fig. 3c)—the virus with greatest overall sequence similarity (Fig. 3a); the corresponding epitope sequences within the other β-HCoVs were even more divergent (Fig. 3c). Of the CD4⁺ T cell clones specific for three epitopes within the S2 region of SARS-CoV-2, which has greater aa similarity with other β-HCoVs (Fig. 3b), NFSQ c117 did not cross-recognize any other β-HCoV (Fig. 3d). STEC c41 and SFIE c55 both cross-recognized the SARS peptide mix, but not the MERS, HKU1 or OC43 peptide mixes (Fig. 3d); both epitopes had only a single aa difference between SARS-CoV-2 and SARS (Fig. 3d). We extended the work to an additional 11 clones: 5 specific for epitopes within S1 and 6 within S2. The epitopes within the S1 region of SARS-CoV-2 differed from SARS by 5–16 aa, whereas the epitopes within the S2 region differed by 0–2 aa (Extended Data Fig. 6). None of the five S1-specific

clones, but all six S2-specific clones, cross-recognized the SARS peptide mix (Table 2); the peptide mixes from the other β-HCoVs, with lower aa sequence similarity to SARS-CoV-2, were never recognized (Table 2). These data indicated that S epitope-specific CD4⁺ T cells isolated following SARS-CoV-2 infection could recognize highly homologous epitopes within the S2 region of SARS, but did not cross-react with MERS or the extant β-HCoVs HKU1 and OC43, consistent with the S protein-specific CD4⁺ T cell clones described here being primed by SARS-CoV-2 infection.

Mutations in variants of concern impair CD4⁺ T cell recognition

Next, we examined the recognition of previous and current SARS-CoV-2 VOCs, including alpha (B.1.1.7), beta (B.1.351), gamma (P.1), delta (B.1.617.2), zeta (P.2), theta (P.3) and omicron BA.1 and BA.2 (B.1.1.529), by S protein-specific CD4⁺ T cell clones. Amino acid substitutions or deletions were present in 10 of the 17 S epitope sequences within VOCs (Table 3), with some common to several VOCs, such as the N501Y mutation in the FNCY epitope identified in alpha, beta and gamma variants (Table 3), and others unique to particular viral isolates, such as the N764K mutation within the STEC epitope in omicron (Table 3). To study the impact of these mutations on CD4⁺ T cell recognition, we tested the S-specific CD4⁺ T cell clones from the HCW-PI donors, who had been infected during the WT (D614G) wave of SARS-CoV-2, against the optimal WT epitope peptide and the corresponding mutated peptides from the VOCs; 20mer peptides were employed to encompass the peptide flanking regions, as mutations within the MHCII-binding core and proximal flanking regions can interfere with epitope binding to MHCII (ref. [39]). Peptides were tested at concentrations of 10^{-5} to 10^{-11} M to detect effects of these mutations at low concentrations that may not be evident at higher concentrations. Single central aa substitutions in epitope
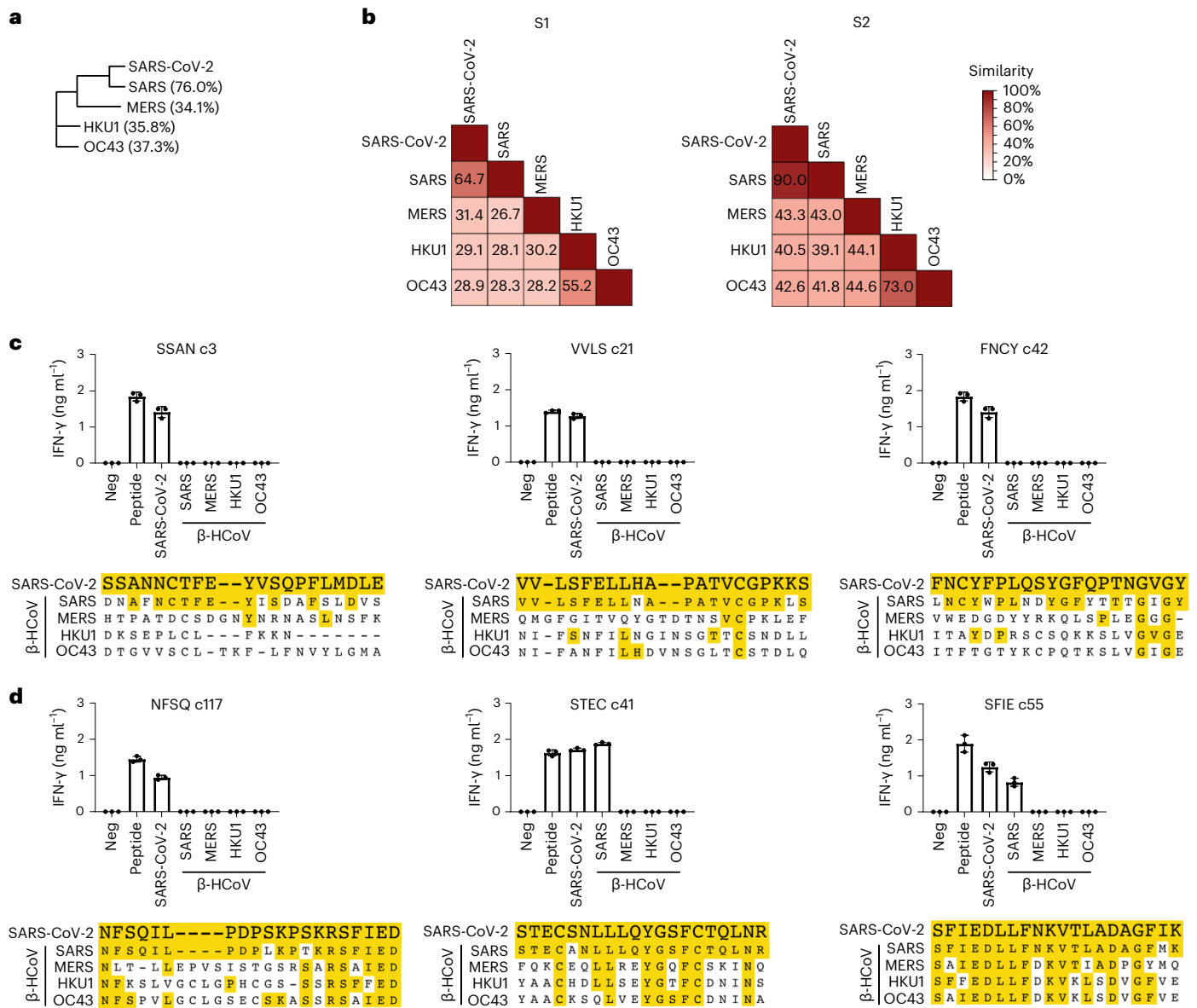
Fig. 3 | Cross-recognition of SARS-CoV-2 spike-specific CD4⁺ T cell clones with β-HCoV. a, Dendrogram showing the evolutionary relationship and percentage aa similarity of SARS-CoV-2 S protein with S proteins from all known β-HCoVs. **b**, Correlogram of the percentage aa similarities between the S1 and S2 regions of the β-HCoVs. **c,d**, ELISA assays for the production of IFN-γ from CD4⁺ T cell clones cocultured with autologous LCL pre-exposed to DMSO solvent (Neg) or epitope peptide or S1 (**c**), or S2 (**d**) peptide mixes from β-HCoVs (15 aa overlapping by 11 aa) in overnight assays. Results show mean IFN-γ release ±1 s.d. and are representative of three independent experiments. Alignments show the aa sequences of SARS-CoV-2 S CD4⁺ T cell epitopes and the corresponding β-HCoV sequences. Amino acids highlighted in yellow are conserved with SARS-CoV-2.

TLVK (N969K) in omicron and SGTN (D80Y) in zeta eliminated CD4⁺ T cell recognition, except at supraphysiological peptide concentrations for the latter (Fig. 4). Two separate point mutations in epitope QLIR, A1022S in beta and T1027I in gamma impaired CD4⁺ T cell recognition at equivalent peptide concentrations (Fig. 4). However, several epitope peptides containing point mutations, such as LSET (T307A) in theta, STEC (N764K) in omicron and TYVT (A1022S) in beta were recognized equally to the WT epitope (Fig. 4). A double aa deletion (Δ43–44) in theta had no effect on the recognition of LVDL and TRFQ epitopes (Fig. 4); however, a triple aa deletion (Δ42–44) in beta reduced CD4⁺ T cell stimulation compared with the WT peptide (Fig. 4).

The epitopes with most mutations were GGNY and FNCY, both present within the RBD region—a frequent target of mutation in VOCs[19,20]. Although single point mutations arising in earlier VOCs in epitopes GGNY (L452R in delta and B.1.324) and FNCY (N501Y in alpha, beta, gamma and theta) did not affect recognition, the numerous mutations accumulated in omicron in GGNY (G446S, L452R, R457N) and FNCY (Q493K, G496S, Q498R, N501Y, Y505H) eliminated recognition by the CD4⁺ T cells specific for these epitopes (Fig. 4). Overall, recognition of 7 of the 17 epitopes in S protein by the CD4⁺ T cells specific for these epitopes was affected by mutations present in one or more VOCs (Fig. 4 and Table 3). The affected epitopes were restricted through a range of HLAII alleles (Table 1) and, in every case where recognition of an epitope was lost, the same person possessed CD4⁺ T cell responses against other epitopes that were not impacted by mutation (Fig. 1d).

T cell clones allow only a small number of TCRs to be studied. In vivo, epitope-specific CD4⁺ T cell responses comprise a multitude of different TCRs[40]. To test whether other TCRs specific for the same

Table 2 | CD4⁺ T cell clone recognition of β-HCoV peptide mixes

Spike region	aa coordinates	Epitope	SARS-CoV-2	SARS	MERS	HKU1	OC43
S1	71–90	SGTN	+	–	–	–	–
	161–180	SSAN	+	–	–	–	–
	226–245	LVDL	+	–	–	–	–
	236–255	TRFQ	+	–	–	–	–
	296–315	LSET	+	–	–	–	–
	446–465	GGNY	+	–	–	–	–
	486–505	FNCY	+	–	–	–	–
	511–530	VVLS	+	–	–	–	–
S2	746–765	STEC	+	+	–	–	–
	801–820	NFSQ	+	–	–	–	–
	816–835	SFIE	+	+	–	–	–
	891–910	GAAL	+	+	–	–	–
	956–975	AQAL	+	+	–	–	–
	961–980	TLVK	+	+	–	–	–
	1,006–1,025	TYVT	+	+	–	–	–
	1,011–1,030	QLIR	+	+	–	–	–
	1,061–1,080	VFLH	+	+	–	–	–

Table 3 | S epitope sequences in SARS-CoV-2 WT and VOCs

aa coordinates	Epitope sequence	Pango lineage	T cell recognition
71–90 SGTN	SGTNGTKRFDNPVLPFNDGV SGTNGTKRF**Y**NPVLPFNDGV	WT	+++
		Zeta	–
226–245 LVDL	LVDLPIGINITRFQTLLALH LVDLPIGINITRFQTLLΔΔHRS LVDLPIGINITRFQTLΔΔΔHRSY	WT	+++
		Theta	+++
		Beta	+
236–255 TRFQ	TRFQTLLALHRSYLTPGDSS TRFQTLLΔΔHRSYLTPGDSSSG TRFQTLΔΔΔHRSYLTPGDSSSGW TRFQTLLALHRSYLTPGDS**F**	WT	+++
		Theta	+++
		Beta	++
		A23.1	+++
296–315 LSET	LSETKCTLKSFTVEKGIYQT LSETKCTLKSF**A**VEKGIYQT	WT	+++
		Theta	+++
446–465 GGNY	GGNYNYLYRLFRKSNLKPFE GGNYNY**R**YRLFRKSNLKPFE **S**GNYNY**R**YRLF**N**KSNLKPFE	WT	+++
		Delta, B.1.324 Omicron	+++ –
486–505 FNCY	FNCYFPLQSYGFQPTNGVGY FNCYFPLQSYGFQPT**Y**GVGY FNCYFPL**K**SY**S**F**R**PT**Y**GVG**H**	WT	+++
		Alpha, beta, gamma, theta	+++
		Omicron	–
746–765 STEC	STECSNLLLQYGSFCTQLNR STECSNLLLQYGSFCTQL**K**R	WT	+++
		Omicron	+++
961–975 TLVK	TLVKQLSSNFGAISS TLVKQLSS**K**FGAISS	WT	+++
		Omicron	-
1,006–1,025 TYVT	TYVTQQLIRAAEIRASANLA TYVTQQLIRAAEIRAS**S**NLA	WT	+++
		Beta	+++
1,011–1,030 QLIR	QLIRAAEIRASANLAATKMS QLIRAAEIRAS**S**NLAATKMS QLIRAAEIRASANLAA**I**KMS	WT	+++
		Beta	++
		Gamma	++

T cell recognition was quantified as the fold increase in concentration required to yield T cell activity equivalent to the EC₅₀ of the WT peptide. Δ, aa deletion; bold indicates aa substitution; +++ equivalent concentration; ++ 1–2 log increased peptide; + 3 log increased peptide; – no T cell response detected.

epitopes may be unaffected by VOC mutations, we examined ex vivo memory CD4⁺ T cell (CD4⁺ T_M cells) populations specific for S protein epitopes in HCW-PI or in UH 1–5 months postvaccination. First, to analyze the clonal TCR composition of the SARS-CoV-2 infection-induced CD4⁺ T_M cells specific for S protein epitopes we used three representative peptide-HLAII tetramers (Tet) containing epitope sequences from the WT S protein: one epitope that contained no mutations in the current VOCs, SSAN (DPB1*04:01/SSAN), one epitope where mutation in the omicron VOC abrogated T cell clone recognition, GGNY (DPB1*02:01/GGNY) and one where the current VOC mutations did not affect T cell clone recognition, STEC (DRB1*15:01/STEC) (Extended Data Fig. 7a,b). The pHLAII Tet were used in flow cytometry assays alongside a panel of antibodies specific for 21 TCR Vβ segments commonly expressed on CD4⁺ T cells[41]. As expected, in every case, the S epitope-specific CD4⁺ T_M cells were polyclonal, with evidence of oligoclonal expansion (Fig. 5a and Extended Data Fig. 7c). Several TCR Vβ segments were over-represented within the pHLAII Tet⁺ CD4⁺ T_M cells compared with the total CD4⁺ T_M cell repertoire, with smaller frequencies of pHLAII Tet⁺ CD4⁺ T_M cells expressing other TCR Vβs also detected (Fig. 5a).

Second, we performed IFN-γ Elispot assays using CD8⁻ PBMCs from HCW-PI or UH 1–5 months postvaccine to test the response of CD4⁺ T_M cells to five epitopes, four of which (SGTN, LVDL, GGNY and QLIR) had shown decreased CD4⁺ T cell clone recognition against VOC peptides and one, STEC, that had maintained recognition. Compared with the WT peptide, the frequency of CD4⁺ T cells that produced IFN-γ in response to the zeta SGTN peptide containing the point mutation (D80Y) was lower at each peptide concentration tested (Fig. 5b). Likewise, the recognition of LVDL, GGNY and QLIR similarly showed reduced T cell responses for at least one of the variant peptides tested. Notably, no ex vivo IFN-γ Elispot response was detected to the mutated GGNY omicron peptide (G446S, L452R, R457N) (Fig. 5b) in the same individual that possessed a polyclonal population of pHLAII Tet⁺ cells specific for the WT peptide (Fig. 5a). In contrast, similar magnitudes of IFN-γ producing CD4⁺ T cells were detected against the STEC peptide in omicron (N764K) and WT (Fig. 5b). Collectively, these data showed the acute sensitivity of SARS-CoV-2 S-specific CD4⁺ T cells to small aa

changes in their epitope sequence and that our findings using T cell clones were representative of circulating polyclonal epitope-specific CD4⁺ T cell responses.

Discussion

Our detailed analysis of HCW previously infected with WT SARS-CoV-2, including the isolation and extensive use of CD4⁺ T cell clones, provided several new insights into the CD4⁺ T cell response to SARS-CoV-2. Focusing on the S protein, the lack of cross-reactivity with β-HCoVs indicated that all the SARS-CoV-2 S-specific CD4⁺ T cell clones originated from the naïve repertoire rather than pre-existing β-HCoV-specific CD4⁺ T_M cells. A key finding was that the CD4⁺ T cell response in every individual targeted several viral epitopes. This broad response is important because we showed mutations in SARS-CoV-2 VOCs compromise CD4⁺ T cell recognition of some, but currently not all, S epitopes. The breadth of the CD4⁺ T cell response therefore limits the impact of mutations in current VOCs on overall CD4⁺ T cell surveillance.

In line with previous studies on HCW-PI assessing T cell responses to whole antigens ex vivo[4,8–10], we detected robust CD4⁺ T_M cell responses against peptide mixes from SARS-CoV-2 S, M and N proteins. Our data
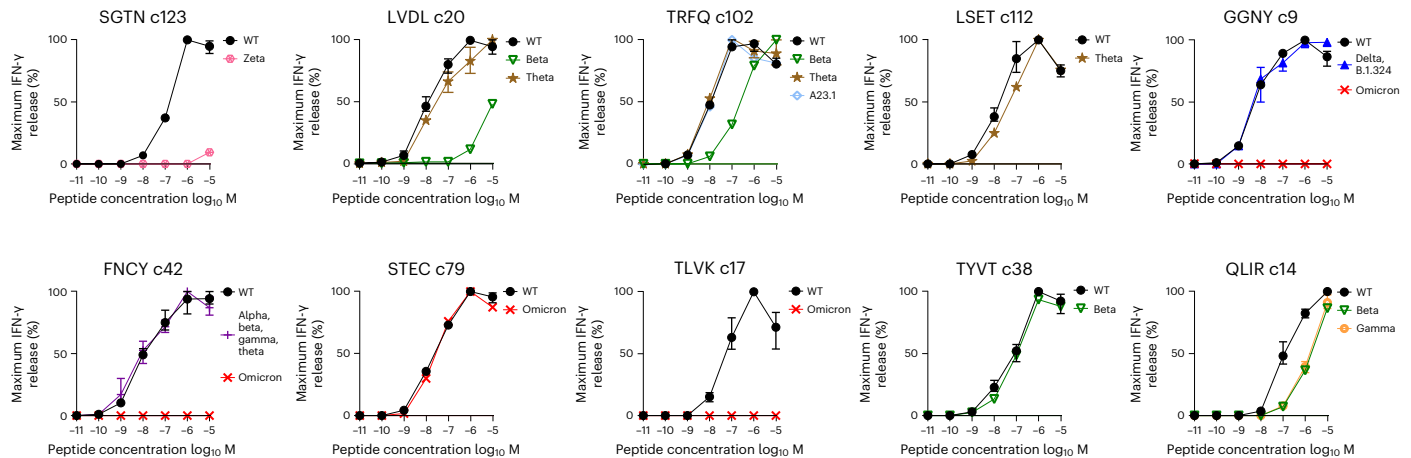
Fig. 4 | Impact of mutations in SARS-CoV-2 VOCs on spike-specific CD4+ T cell clone recognition. Representative ELISA assays for the production of IFN-γ from CD4+ T cell clones cocultured with autologous LCLs loaded with purified SARS-CoV-2 WT S peptides or corresponding mutated peptides from VOCs at 10^{-5} to 10^{-11} M in overnight assays. Results show mean IFN-γ release ±1 s.d. and are representative of three independent experiments.
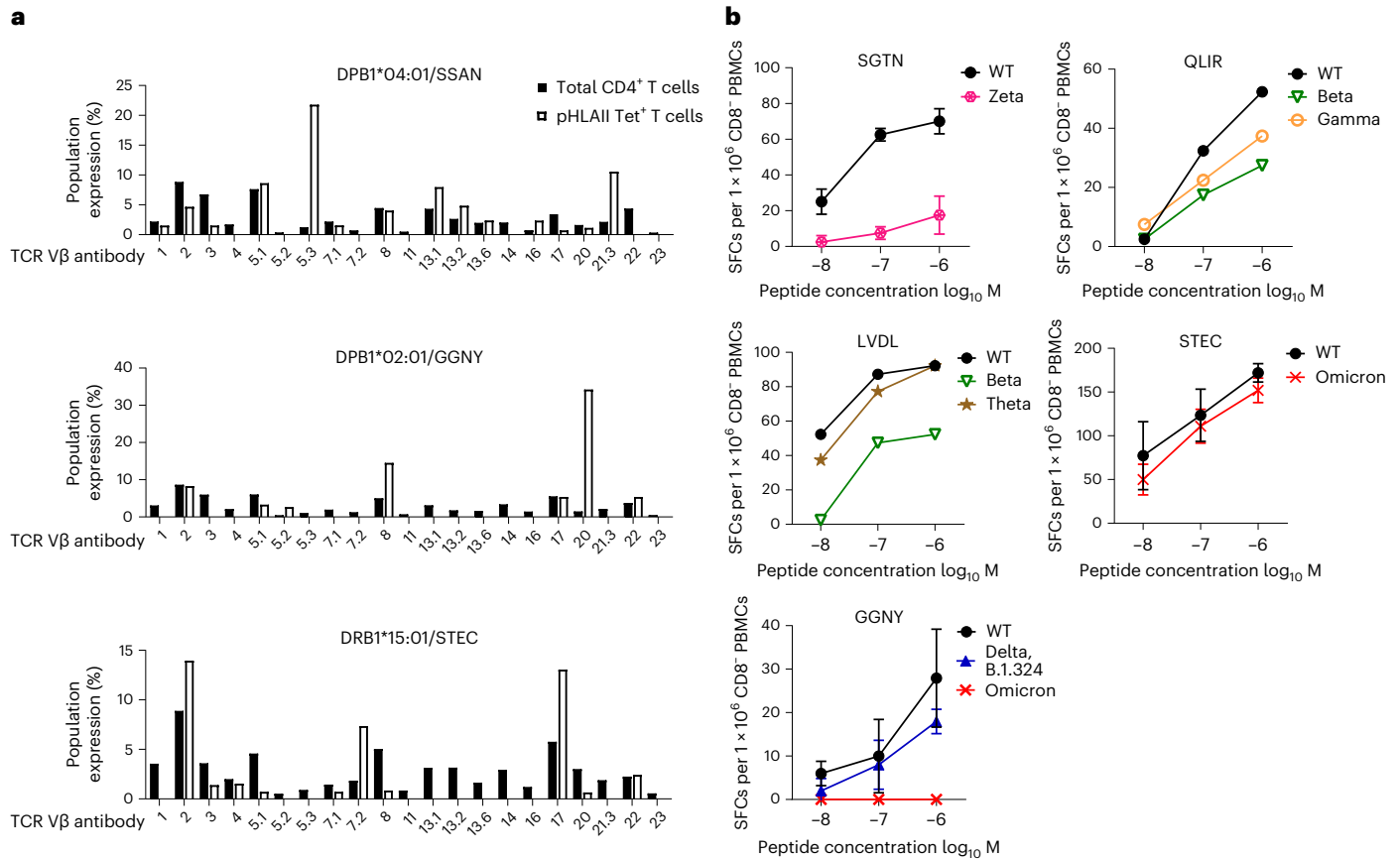

Fig. 5 | The impact of mutations in SARS-CoV-2 VOCs on ex vivo CD4+ T cell recognition. a, Flow cytometry analysis of total CD4+ T cells and pHLAII tetramer+ (pHLAII Tet+) cells in PBMCs stained with the indicated pHLAII Tet and antibodies against 21 defined TCR Vb segments. **b**, Elispot assays for the production of IFN-γ in CD8-PBMCs plated at 4×10^{5} cells per well and incubated with purified SARS-CoV-2 WT S peptides or corresponding mutated peptides from VOCs at 10^{-6} to 10^{-8} M. Results are shown as mean SFC ± 1 s.d. per 10^{6} CD8-PBMCs.

showed that every donor possessed a broad CD4+ T cell response after SARS-CoV-2 infection that targeted several epitopes. In total, we defined 21 HLAII-restricted epitopes. This considerably expanded the number of experimentally verified CD4+ T cell epitopes and also specified the HLAII restriction of some previously reported epitopes[11–13,42,43]. Some of our epitopes were presented by HLAII alleles found at high

frequency in global populations, such as the HLA-DRB1*04 subtypes, DRB1*01:01, DRB1*15:01 and DPB1*04:01[44]. CD4+ T cell responses to these epitopes are therefore likely to be widespread following infection or vaccination. Our data considerably strengthens the evidence that SARS-CoV-2 epitopes are skewed towards HLA-DP and HLA-DR restriction alleles[14,25]. This HLAII usage is distinct from other human viruses

investigated using CD4[+] T cell clones[37,45].The reason remains unclear, but it will be important to understand, as our data showed that HLAII genotype was a key determinant of the CD4[+] T cell response to S protein following either SARS-CoV-2 infection or vaccination.

Several endemic β-HCoVs infect humans and cause mild disease. T cells elicited by previous infection with these viruses could modulate the course of disease, if they cross-reacted with SARS-CoV-2. All the CD4[+] T cell clones we studied were generated using SARS-CoV-2 peptides from donors previously infected with SARS-CoV-2, who also had evidence of historical infection with the extant β-HCoVs HKU1 and OC43. All S protein-specific CD4[+] T cell clones efficiently recognized SARS-CoV-2 S protein peptide mix, but not the peptide mixes from OC43 or HKU1 S proteins, which have less than 40% aa similarity to SARS-CoV-2. The only observed cross-reactivity to β-HCoV was against epitopes located within the S2 region of SARS, which has 90% aa similarity with SARS-CoV-2, but to which the UK HCWs studied here had never been exposed. These results strongly suggested that, for our cohort of previously SARS-CoV-2 infected donors, CD4[+] T cell clones specific for the S protein originated from naïve CD4[+] T cells primed by SARS-CoV-2 infection rather than from CD4[+] T_M cells primed by previous β-HCoV infections. Consistent with our data, CD8[+] T_M cells specific for an epitope in the SARS-CoV-2 N protein originated from the naïve CD8[+] T cell repertoire in SARS-CoV-2 infected individuals[46].

Our data do not contradict previous studies that have reported the presence of pre-existing cross-reactive S-specific CD4[+] T cells in unexposed individuals[14-17]. Where cross-reactivity has been investigated at the level of epitopes, it was focused on small regions of the S protein that are highly conserved between SARS-CoV-2 and other β-HCoVs[13-15,18,47] Our data showed that following SARS-CoV-2 infection, the CD4[+] T cell response to the S protein was broadly targeted across the entire protein. Therefore, most CD4[+] T_M cell responses to SARS-CoV-2 S protein targeted epitopes with low sequence similarity to the S proteins of other β-HCoVs and, as expected, did not cross-react.

We only investigated the cross-reactivity of CD4[+] T cells specific for the S protein, which is highly targeted by mutation. Other viral proteins are more conserved across β-HCoVs and T cell cross-reactivity may therefore be more probable. Accordingly, pre-existing T cell immunity against the highly conserved N and ORF1ab-encoded nonstructural proteins (NSP) has been reported in SARS-CoV-2-exposed HCWs with no evidence of virus infection[15,16,48]. The β-HCoV cross-reactive T cells in those donors were frequently directed against epitopes located in the early-expressed replication transcription complex encompassing NSP7, NSP12 and NSP13[48]. Ultimately, the extent to which pre-existing cross-reactive T cell immunity contributes to controlling SARS-CoV-2 infection in an individual might be determined by a complex combination of factors, including the conservation of the epitopes presented by their HLA genotype, their TCR repertoire and their history of previous β-HCoV exposure[13].

Examining S proteins sequences from VOCs, we identified aa substitutions or deletions in 10 of the 17 S protein CD4[+] T cell epitopes; some were common to several VOCs, while others were unique to particular viral isolates. Combining data from the CD4[+] T cell clone experiments and the experiments using ex vivo PBMCs, which contain polyclonal epitope-specific populations, we found variable effects of the epitope mutations on CD4[+] T cell recognition. Triple aa deletion and several aa substitutions within individual epitopes had the greatest impact on CD4[+] T cell recognition[49]. The effect of point mutations was complex. A single mutation could markedly decrease CD4[+] T cell recognition or have no effect. Notably, decreased CD4[+] T cell sensitivity to several epitopes was only apparent at lower peptide concentrations and was not evident using the high concentrations of peptides previously employed[22,23,29,50]. This highlights the requirement for further careful experimental definition of the currently known immunodominant SARS-CoV-2 CD4[+] T cell epitopes, the identification of the essential aa required for

HLAII binding and TCR engagement and the need to consider each epitope individually. The epitope mapping presented here provides a rational basis for VOC risk stratification. Additional mutations recently acquired in the BA.4 and BA.5 omicron variants, such as Δ69−70 and F486V highlight the continued evolution of SARS-CoV-2 and potential for further T cell epitope mutation.

In conclusion, our study demonstrated the fine sensitivity of SARS-CoV-2 S-specific CD4[+] T cells to aa variation in epitope sequence and the potential for SARS-CoV-2 evolution to evade the CD4[+] T_M response. The breadth of SARS-CoV-2 S epitopes targeted in every individual indicated current VOC mutations are likely to have only limited impact on overall CD4[+] T cell surveillance. However, continued evolution of SARS-CoV-2 could lead to further epitope loss and continued monitoring of emerging VOCs is important.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-022-01351-7.

References

1. Rydyznski Moderbacher, C. et al. Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell* **183**, 996–1012.e19 (2020).
2. Bonifacius, A. et al. COVID-19 immune signatures reveal stable antiviral T cell function despite declining humoral responses. *Immunity* **54**, 340–54.e6 (2021).
3. Tregoning, J. S., Flight, K. E., Higham, S. L., Wang, Z. & Pierce, B. F. Progress of the COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape. *Nat. Rev. Immunol.* **21**, 626–636 (2021).
4. Zuo, J. et al. Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection. *Nat. Immunol.* **22**, 620–626 (2021).
5. Jung, J. H. et al. SARS-CoV-2-specific T cell memory is sustained in COVID-19 convalescent patients for 10 months with successful development of stem cell-like memory T cells. *Nat. Commun.* **12**, 4043 (2021).
6. Moss, P. The T cell immune response against SARS-CoV-2. *Nat. Immunol.* **23**, 186–193 (2022).
7. Bertoletti, A., Le Bert, N., Qui, M. & Tan, A. T. SARS-CoV-2-specific T cells in infection and vaccination. *Cell Mol. Immunol.* **18**, 2307–2312 (2021).
8. Cohen, K. W. et al. Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep. Med.* **2**, 100354 (2021).
9. Painter, M. M. et al. Rapid induction of antigen-specific CD4+ T cells is associated with coordinated humoral and cellular immunity to SARS-CoV-2 mRNA vaccination. *Immunity* **54**, 2133–42.e3 (2021).
10. Ewer, K. J. et al. T cell and antibody responses induced by a single dose of ChAdOx1 nCoV-19 (AZD1222) vaccine in a phase 1/2 clinical trial. *Nat. Med.* **27**, 270–278 (2021).
11. Tarke, A. et al. Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. *Cell Rep. Med.* **2**, 100204 (2021).
12. Verhagen, J. et al. Human CD4+ T cells specific for dominant epitopes of SARS-CoV-2 spike and nucleocapsid proteins with therapeutic potential. *Clin. Exp. Immunol.* **205**, 363–378 (2021).
13. Johansson, A. M. et al. Cross-reactive and mono-reactive SARS-CoV-2 CD4+ T cells in prepandemic and COVID-19 convalescent individuals. *PLoS Pathog.* **17**, e1010203 (2021).

14. Low, J. S. et al. Clonal analysis of immunodominance and cross-reactivity of the CD4 T cell response to SARS-CoV-2. *Science* **372**, 1336–1341 (2021).

15. Mateus, J. et al. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* **370**, 89–94 (2020).

16. Le Bert, N. et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **584**, 457–462 (2020).

17. Braun, J. et al. SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* **587**, 270–274 (2020).

18. Loyal, L. et al. Cross-reactive CD4+ T cells enhance SARS-CoV-2 immune responses upon infection and vaccination. *Science* **374**, eabh1823 (2021).

19. Harvey, W. T. et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **19**, 409–424 (2021).

20. Greaney, A. J. et al. Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe* **29**, 44–57.e9 (2021).

21. Cele, S. et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature* **602**, 654–656 (2022).

22. GeurtsvanKessel, C. H. et al. Divergent SARS CoV-2 omicron-reactive T- and B cell responses in COVID-19 vaccine recipients. *Sci. Immunol.* **7**, eabo2202 (2022).

23. Geers, D. et al. SARS-CoV-2 variants of concern partially escape humoral but not T-cell responses in COVID-19 convalescent donors and vaccinees. *Sci. Immunol.* **6**, eabj1750 (2022).

24. Gruell, H. et al. mRNA booster immunization elicits potent neutralizing serum activity against the SARS-CoV-2 Omicron variant. *Nat. Med.* **28**, 477–480 (2022).

25. Nelde, A. et al. SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. *Nat. Immunol.* **22**, 74–85 (2021).

26. Wolter, N. et al. Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. *Lancet* **399**, 437–446 (2022).

27. Tarke, A. et al. Impact of SARS-CoV-2 variants on the total CD4+ and CD8+ T cell reactivity in infected or vaccinated individuals. *Cell Rep. Med.* **2**, 100355 (2021).

28. Keeton, R. et al. Prior infection with SARS-CoV-2 boosts and broadens Ad26.COV2.S immunogenicity in a variant-dependent manner. *Cell Host Microbe* **29**, 1611–1619.e5 (2021).

29. Goel, R. R. et al. mRNA vaccines induce durable immune memory to SARS-CoV-2 and variants of concern. *Science* **374**, abm0829 (2021).

30. Gao, Y. et al. Ancestral SARS-CoV-2-specific T cells cross-recognize the Omicron variant. *Nat. Med.* **28**, 472–476 (2022).

31. Tarke, A. et al. SARS-CoV-2 vaccination induces immunological T cell memory able to cross-recognize variants from alpha to omicron. *Cell* **185**, 847–859.e11 (2022).

32. de Silva, T. I. et al. The impact of viral mutations on recognition by SARS-CoV-2 specific T cells. *iScience* **24**, 103353 (2021).

33. Rudolph, M. G., Stanfield, R. L. & Wilson, I. A. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev. Immunol.* **24**, 419–466 (2006).

34. Draenert, R. et al. Comparison of overlapping peptide sets for detection of antiviral CD8 and CD4 T cell responses. *J. Immunol. Methods* **275**, 19–29 (2003).

35. Long, H. M. et al. CD4+ T-cell responses to Epstein-Barr virus (EBV) latent-cycle antigens and the recognition of EBV-transformed lymphoblastoid cell lines. *J. Virol.* **79**, 4896–4907 (2005).

36. Crompton, L., Khan, N., Khanna, R., Nayak, L. & Moss, P. A. CD4+ T cells specific for glycoprotein B from cytomegalovirus exhibit extreme conservation of T-cell receptor usage between different individuals. *Blood* **111**, 2053–2061 (2008).

37. Cassotta, A. et al. Deciphering and predicting CD4+ T cell immunodominance of influenza virus hemagglutinin. *J. Exp. Med.* **217**, e20200206 (2020).

38. Cueno, M. E. & Imai, K. Structural comparison of the SARS CoV 2 spike protein relative to other human-infecting coronaviruses. *Front Med (Lausanne)* **7**, 594439 (2020).

39. Holland, C. J., Cole, D. K. & Godkin, A. Re-directing CD4+ T cell responses with the flanking residues of MHC class II-bound peptides: the core is not enough. *Front Immunol.* **4**, 172 (2013).

40. Meckiff, B. J. et al. Primary EBV infection induces an acute wave of activated antigen-specific cytotoxic CD4+ T cells. *J. Immunol.* **203**, 1276–1287 (2019).

41. Wei, S., Charmley, P., Robinson, M. A. & Concannon, P. The extent of the human germline T-cell receptor V beta gene segment repertoire. *Immunogenetics* **40**, 27–36 (1994).

42. Poluektov, Y., George, M., Daftarian, P. & Delcommenne, M. C. Assessment of SARS-CoV-2 specific CD4+ and CD8+ T cell responses using MHC class I and II tetramers. *Vaccine* **39**, 2110–2116 (2021).

43. Peng, Y. et al. 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 (2020).

44. Gonzalez-Galarza, F. F. et al. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res.* **48**, D783–D788 (2020).

45. Taylor, G. S., Long, H. M., Brooks, J. M., Rickinson, A. B. & Hislop, A. D. The immunology of Epstein-Barr virus-induced disease. *Annu. Rev. Immunol.* **33**, 787–821 (2015).

46. Peng, Y. et al. An immunodominant NP105-113-B*07:02 cytotoxic T cell response controls viral replication and is associated with less severe COVID-19 disease. *Nat. Immunol.* **23**, 50–61 (2022).

47. Dykema, A. G. et al. Functional characterization of CD4+ T cell receptors crossreactive for SARS-CoV-2 and endemic coronaviruses. *J. Clin. Invest.* **131**, e146922 (2021).

48. Swadling, L. et al. Pre-existing polymerase-specific T cells expand in abortive seronegative SARS-CoV-2. *Nature* **601**, 110–117 (2022).

49. Phillips, R. E., Harcourt, G. C. & Price, D. A. CD4+ T cells: the great escape. *Nat. Med.* **7**, 777–778 (2001).

50. Naranbhai, V. et al. T cell reactivity to the SARS-CoV-2 omicron variant is preserved in most but not all individuals. *Cell* **185**, 1041–1051.e6 (2022).

Methods

Donor characteristics and ethical statement

The study was approved by the North West–Preston Research Ethics Committee, United Kingdom (20/NW/0240) and all participants gave written informed consent and received no compensation. Blood was collected in June to September from 20 healthcare workers (HCW), ten males and ten females aged 28–64 years, at 3–6 months postinfection during the first wave of SARS-CoV-2 WT (D614G) infection in the United Kingdom. Control samples were collected from SARS-CoV-2 UH individuals (*n* = 14; 6 males and 8 females aged 24–63 years) confirmed to be seronegative for SARS-CoV-2 (Extended Data Fig. 1), or collected before the pandemic from healthy donors as part of an ethically approved study (South Birmingham Research Ethics Committee 14/WM/1254). Subsequent samples were collected from those UH individuals who remained uninfected following vaccination with Pfizer BioNtech BNT162b2 or AstraZeneca ChAdOx-1 (*n* = 9; 3 males and 6 females aged 39–63 years) 1–5 months after vaccination.

Sample preparation

Plasma and PBMCs were isolated from heparinized blood using standard Ficoll-Hypaque centrifugation. The resulting PBMC layer was washed twice with RPMI and either used directly or cryopreserved before use. Where stated, PBMCs were depleted of CD8+ T cells (CD8−PBMCs) using anti-CD8 Dynabeads (Invitrogen) to over 94%, as measured by flow cytometry. Autologous LCLs were generated by transformation with B95.8 EBV as previously described[51] and HLAII typing by next generation sequencing was performed at the Anthony Nolan Histocompatibility Laboratories.

Synthetic peptides and protein

For stimulation of T cell responses against whole protein sequences, peptide mixes containing 15mer peptides overlapping by 11 aa covering the full length of WT (D614G) SARS-CoV-2 S1 and S2 (PM-WCPV-S-2), M (PM-WCPV-VME-2) and N (PM-WCPV-NCAP-2), the S1 and S2 regions of SARS (PM-CVHSA-S-1), MERS (PM-MERS-CoV-1), HKU1 (PM-HKU1-S-1) and OC43 (PM-OC43-S-1), and control HLAII CEFX peptide mix (PM-CEFX-3) were purchased from JPT. HLAI peptide mix was generated inhouse by combining 30 known EBV CD8+ T cell epitope peptides[45].

For analysis of responses at the epitope level, individual 20mer peptides overlapping by 15 aa covering the full sequences of SARS-CoV-2 S, M and N were purchased from Alta Biosciences. Upon epitope identification, purified epitope peptides (>85% purity) and corresponding peptides (20mers and one 15mer for which the variant 20mer peptide could not be synthesized) from SARS-CoV-2 VOCs were synthesized by Alta Biosciences and Genscript. All peptides were resuspended in DMSO at a concentration of 5 mg ml⁻¹. The concentration of all purified epitope peptides was confirmed using a Pierce Quantitative Colorimetric Peptide assay (Thermo Scientific) that binds to the peptide amide backbone, and peptides were adjusted to equivalent concentrations. The SARS-CoV-2 S protein is a soluble prefusion stabilized form, containing proline substitutions at positions F817P, A892P, A899P, A942P, K986P and V987P[52]. Epitope sequences recognized by the isolated CD4+ T cell clones are free of these stabilizing mutations.

Interferon-γ Elispot assay

Whole or CD8−PBMCs (0.2–0.4 × 10⁶ cells) were resuspended in standard medium (RPMI supplemented with 8% batch-tested FCS, 100 IU ml⁻¹ penicillin and 100 µg ml⁻¹ streptomycin) and added to duplicate or triplicate wells of IFNγ Elispot Pro Kit (Mabtech) plates containing 1 µg ml⁻¹ peptide mix, 5 µg ml⁻¹ purified epitope peptide or titrated concentration as stated, DMSO (negative control) and PHA or anti-CD3 (positive controls). Samples were incubated at 37 °C for 16–18 h. Plates were developed in accordance with the manufacturer's instructions and read on a Bioreader 5000 Pro F Gamma (Bio-Sys GmbH). To quantify antigen-specific responses, mean DMSO values were deducted from all test wells and the results were expressed as spot forming cells (SFC) per 10⁶ cells.

Polyclonal T cell generation and analysis

Polyclonal CD4+ T cell lines, generated by stimulation of CD8−PBMCs with 1 µg ml⁻¹ S, M or N peptide mix (JPT, 15 aa peptides overlapping by 11 aa), were cultured in RPMI supplemented with 5% batch-tested human serum (Gibco) and 50 IU ml⁻¹ IL-2. Following 1–4 biweekly repeat stimulations with the same peptide mix, 50,000 polyclonal T cells were incubated for 16–18 h in V-bottom microtest plate wells with 1 µg ml⁻¹ individual 20mer peptides overlapping by 10 aa covering S, M or N or DMSO (negative control). IFNγ release into the supernatant was tested by enzyme-linked immunosorbent assay (ELISA; Invitrogen) according to the manufacturer's protocol. Responses against the test peptides were considered positive if greater than twice the mean of the control wells.

T cell clone isolation and assays

CD4+ T cell clones were isolated from polyclonal cultures as previously described[53]. Briefly, polyclonal CD4+ T cells were selected on rechallenge with 1 µg ml⁻¹ appropriate peptide mix using an IFNγ cell enrichment kit (Miltenyi Biotech) followed by MACS separation using anti-PE beads. Enriched cells were cloned by limiting dilution seeding to establish T cell clones originating from single cells. Growing microcultures were screened for reactivity against individual 20mer peptides and selected clones were expanded using standard methods[35].

CD4+ T cell clones (2,000 or 5,000 cells per well) were incubated in V-bottom 96-well microtest plate wells with 5 × 10⁴ cells per well of autologous LCL or allogeneic LCLs with HLAII types partially matched to the autologous LCL. In some assays, LCLs were either pre-exposed for 1 h to 5 µM 20mer epitope peptide, 1 µg ml⁻¹ peptide mix, or for 3 h to 1 ng ml⁻¹ soluble prefusion-stabilized SARS-CoV-2 S protein[52] or equivalent volumes of DMSO (negative control) before washing and addition to the wells, or added with a titrated concentration of individual 20mer peptides. Half maximal effective concentration (EC₅₀) values were defined for each peptide as the concentration eliciting 50% maximum IFNγ release produced in response to the optimal WT peptide. T cell clone responses to mutated peptides from VOCs were quantified as the fold increase in concentration required to yield T cell activity equivalent to the EC₅₀ of the WT peptide. In blocking assays, peptide-loaded or DMSO-exposed LCLs were incubated for 1 h with purified monoclonal antibodies against HLA-DR (L243, Biolegend), HLA-DQ (SPV-L3, Biotium) and HLA-DP (B7/21, Life Technologies) before addition of T cells. In all other assays, T cells were added immediately. The supernatant medium harvested after 16–18 h was assayed for IFNγ by ELISA.

pHLAII tetramer staining and TCR Vβ repertoire analysis

PBMCs from HCW-PI or UH individuals postvaccine were stained with optimized concentrations of peptide-HLAII tetramers (pHLAII Tet) (NIH tetramer core) containing S epitopes SSAN (DPB1*04:01/SSAN), GGNY (DPB1*02:01/GGNY) or STEC (DRB1*15:01/STEC), appropriate for the HLAII genotype of each individual. A total of 1.5–2 × 10⁶ PBMCs were used per tube to enable collection of low frequency pHLAII Tet+ events. PBMCs were washed in PBS and stained with pHLAII Tet in batch-tested human serum for 1 h at 37 °C with regular resuspension. After incubation, cells were washed in PBS and stained at RT for 30 mins with BV510 anti-CD14 (MΦP9), BV510 anti-CD19 (SJ25C1), LIVE/DEAD Fixable Aqua Dead Cell Stain (Invitrogen), BV650 anti-CD3 (OKT3) and PerCP anti-CD4 (RPA-T4), plus appropriate combinations of the following TCR Vβ antibodies covering over 70% of normal human expressed TCR Vβ repertoire[41]: FITC anti-TCR Vβ3 (CH92), FITC anti-TCR Vβ4 (WJF24), FITC anti-TCR Vβ5.1 (IMMU 157), FITC anti-TCR Vβ5.2 (36213), FITC anti-TCR Vβ8 (56C5.2), FITC anti-TCR Vβ13.1 (H131), FITC anti-TCR Vβ20 (ELL1.4), FITC anti-TCR Vβ22 (IMMU 546), APC

anti-TCR Vβ1 (REA662), APC anti-TCR Vβ5.3 (REA670), APC anti-TCR Vβ7.1 (REA871), APC anti-TCR Vβ7.2 (REA677), APC anti-TCR Vβ13.6 (REA554), APC anti-TCR Vβ14 (REA557), APC anti-TCR Vβ23 (REA497), AF647 anti-TCR Vβ13.2 (H132), APC-Vio770 anti-TCR Vβ2 (REA654), APC-Vio770 anti-TCR Vβ11 (REA559), APC-Vio770 anti-TCR Vβ16 (REA553), APC-Vio770 anti-TCR Vβ17 (REA915), APC-Vio770 anti-TCR Vβ21.3 (REA894). Cells were washed and resuspended in PBS following staining and data were acquired on a BD LSR Fortessa ×20 flow cytometer (BD Biosciences) using FACSDiva (v.9.0). All data was processed using FlowJo analysis software (v.10.6.1).

Serological analysis
Quantitative IgG antibody titers against SARS-CoV-2 S and N and β-HCoV family S proteins were measured using a multiplex serology assay (V-PLEX COVID-19 Coronavirus Panel 2 (IgG) kit, catalog no. K15369U), according to the manufacturer's instructions. In brief, 96-well plates were blocked and washed. Samples were prediluted 1:4,000 in provided sample diluent and added to the wells in duplicate alongside the reference standard and assay kit controls. Following incubation, washing and addition of anti-IgG detection antibodies, read buffer was added to all wells and plates were measured immediately using a MESO Quickplex SQ 120 System (Meso Scale Discovery). Data were generated by Methodological Mind software (v.1.0.36), adjusted for sample dilutions and analyzed using MSD Discovery Workbench (v.4.0).

Bioinformatic analysis
Spike amino acid alignments were performed using the MUSCLE algorithm[54] with default settings. Protein sequence identity was calculated using Expasy[55]. Correlation plots were prepared in R v.4.0.3 using corrplot (v.0.84).

Statistics and reproducibility
No statistical methods were used to predetermine cohort sizes and researchers were not blinded to the serostatus of donors before Elispot and serological assays. All statistical tests were performed in GraphPad Prism (v.9.3.1).

Reporting summary
Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
Data in this study are available within the article and from the corresponding author on reasonable request. Source data are provided with this paper.

References

51. Frisan, T., Levitsky, V. & Masucci, M. Generation of lymphoblastoid cell lines (LCLs). *Methods Mol. Biol.* **174**, 125–127 (2001).
52. Hsieh, C. L. et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).
53. Long, H. M. et al. Cytotoxic CD4+ T cell responses to EBV contrast with CD8 responses in breadth of lytic cycle antigen choice and in lytic cycle recognition. *J. Immunol.* **187**, 92–101 (2011).
54. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
55. Duvaud, S. et al. Expasy, the Swiss bioinformatics resource portal, as designed by its users. *Nucleic Acids Res.* **49**, W216–W227 (2021).

Author contributions
H.M.L. designed and supervised the study. E.X.C.T., E.J., T.A.H., B.K., N.K., S.J.D. and H.M.L. performed in vitro assays. E.X.C.T., E.J., T.A.H., B.K., N.K., S.J.D., P.M., G.S.T. and H.M.L. performed data analysis. P.P. and H.M.P. conducted sample collection. M.L.N. and M.C. generated and provided essential reagents. G.S.T. and H.M.L. led on data interpretation and wrote the paper. All authors commented on the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41590-022-01351-7.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-022-01351-7.

Correspondence and requests for materials should be addressed to Heather M. Long.

Peer review information *Nature Immunology* thanks the anonymous reviewers for their contribution to the peer review of this work. Ioana Visan was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.
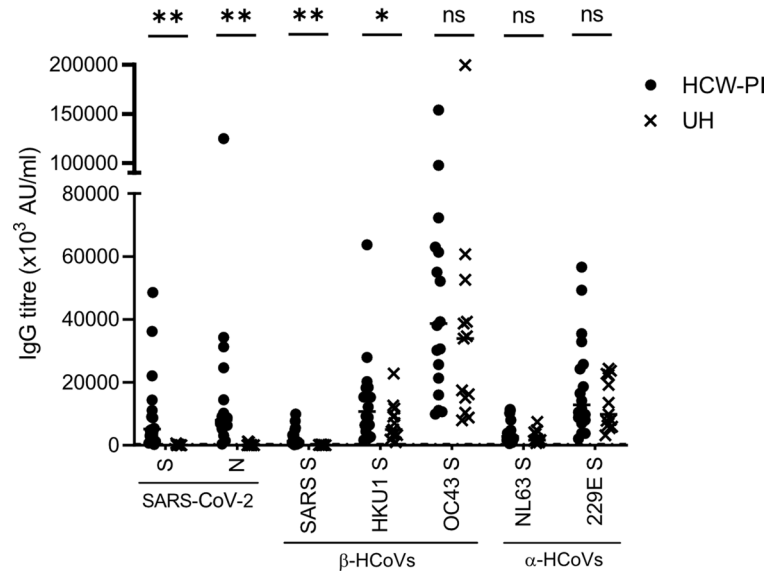
Extended Data Fig. 1 | Antibody responses to coronaviruses in healthcare workers post-infection and uninfected healthy individuals. Multiplex serology assays showing IgG antibody titres against SARS-CoV-2 S and N proteins and S protein from HCoVs in HCW-PI 3–6 months post-infection (n = 18) and UH individuals (n = 13). Significance was determined by two-sided Mann-Whitney U test, **p < 0.0001, *p = 0.0379.
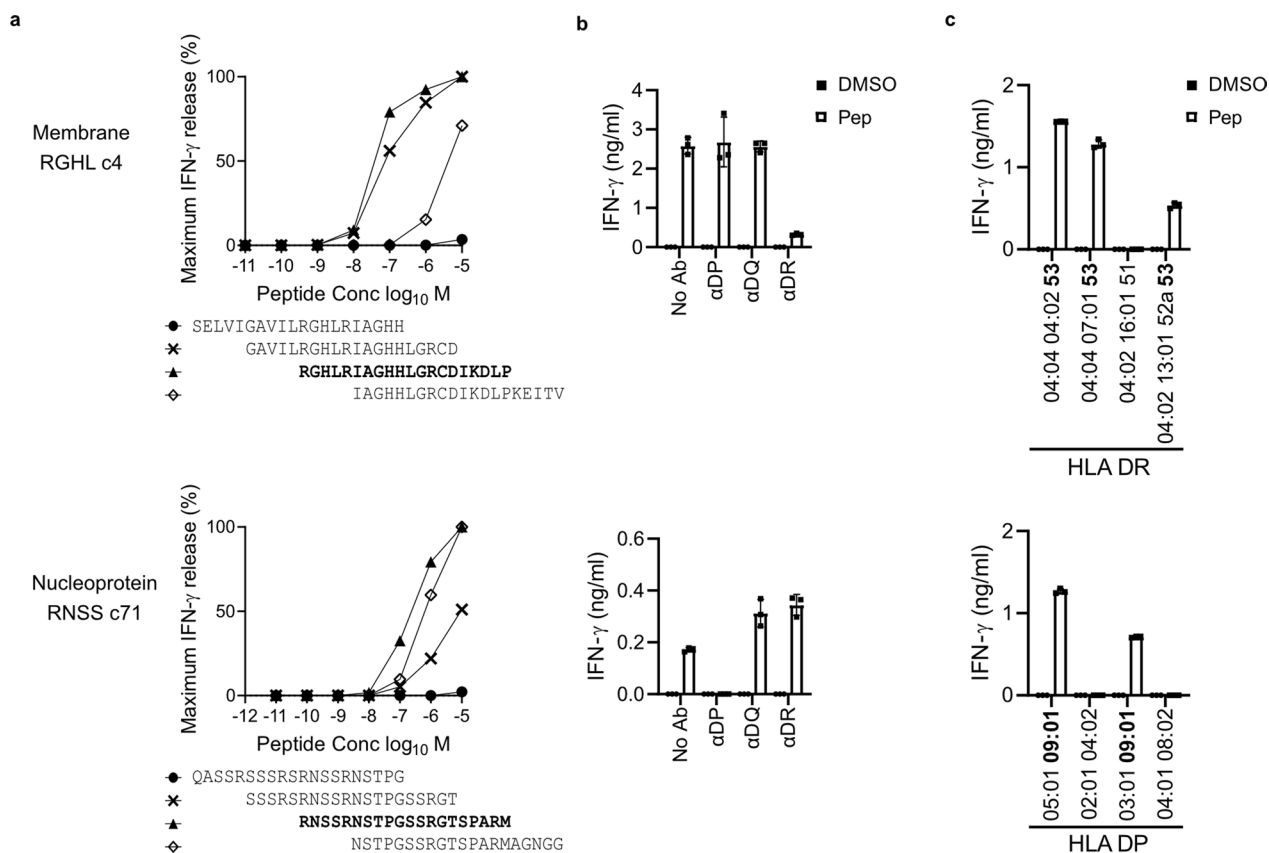
Extended Data Fig. 2 | Characterisation of novel membrane and nucleoprotein CD4+ T cell epitopes. ELISA assays for the production of IFNγ fro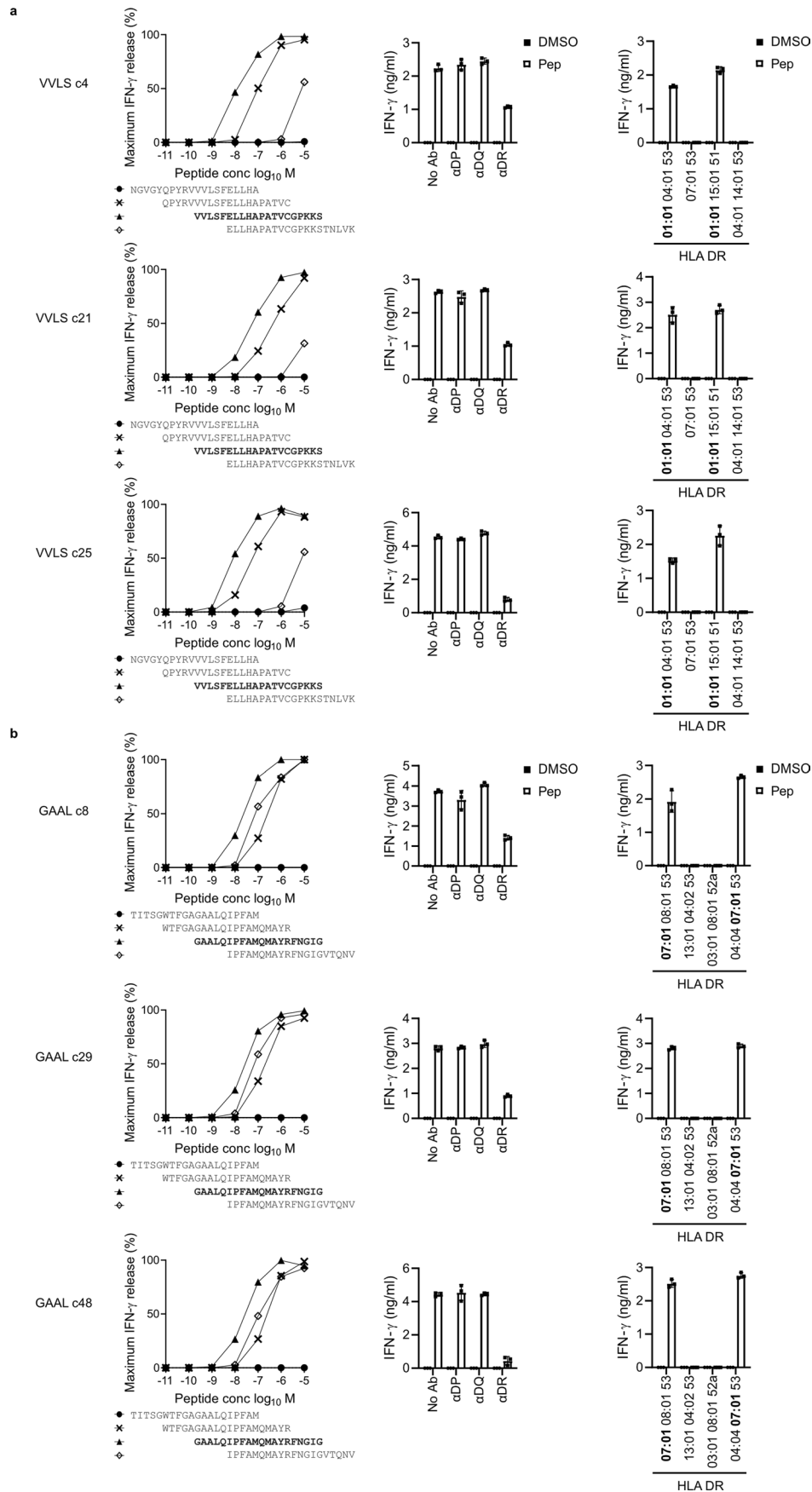m CD4+ T cell clones cocultured in overnight assays with **(a)** autologous lymphoblastoid cell line (LCL) loaded with individual 20mer peptides overlapping by 15aa (10^{-5} to 10^{-11}M), **(b)** autologous LCL pre-pulsed with epitope peptide or DMSO solvent and either tested alone (no antibody; No Ab), or in the presence of blocking antibodies against HLA-DP, HLA-DQ or HLA-DR, and **(c)** autologous LCL and allogeneic LCLs with HLAII types partially matched to the autologous LCL, either pre-pulsed with 5µM 20mer epitope peptide or DMSO solvent (neg). (a-c) Results show mean IFNγ release ±1SD and are representative of 3 experiments.

Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Mapping of multiple CD4$^+$ T cell clones against S epitopes. ELISA assays for the production of IFNγ from **(a)** three VVLS-specific and **(b)** three GAAL-specific CD4$^+$ T cell clones cocultured in overnight assays with autologous lymphoblastoid cell line (LCL) loaded with individual 20mer peptides overlapping by 15aa (10^{-5} to 10^{-11}M) (left panels), autologous LCL pre-pulsed with epitope peptide or DMSO solvent and either tested alone (no antibody; No Ab), or in the presence of blocking antibodies against HLA-DP, HLA-DQ or HLA-DR (middle panels) and autologous LCL and allogeneic LCLs with HLAII types partially matched to the autologous LCL, either pre-pulsed with 5μM 20mer epitope peptide or DMSO solvent (neg) (right panels). (a-c) Results show mean IFNγ release ±1SD and are representative of 3 experiments.
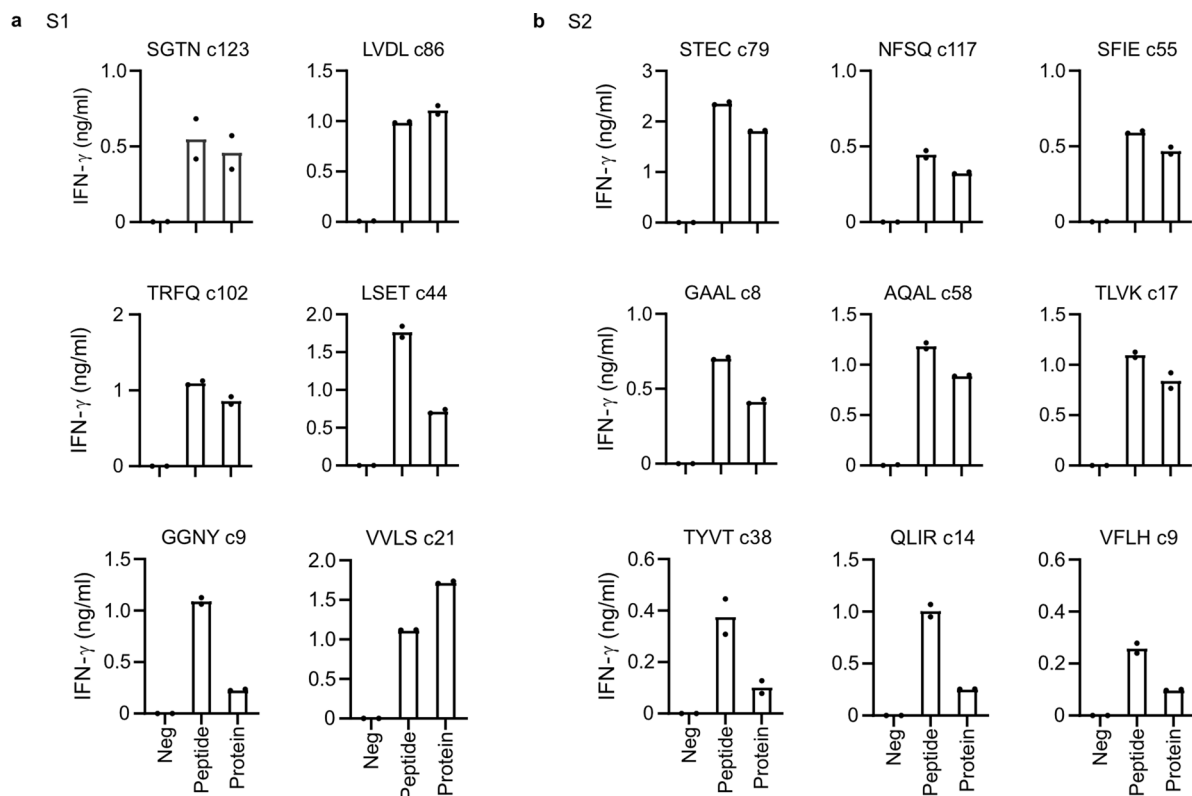
Extended Data Fig. 4 | Spike-specific CD4⁺ T cell clone recognition of S protein. ELISA assays for the production of IFNγ from CD4⁺ T cell clones cocultured overnight with autologous LCL pre-pulsed with DMSO solvent (neg), epitope peptide or 1 ng/ml S tetrameric protein. Data is shown for CD4⁺ T cell clones specific for the 15S epitopes not presented in Fig. 2. Results show mean IFNγ release and are representative of 2 experiments.
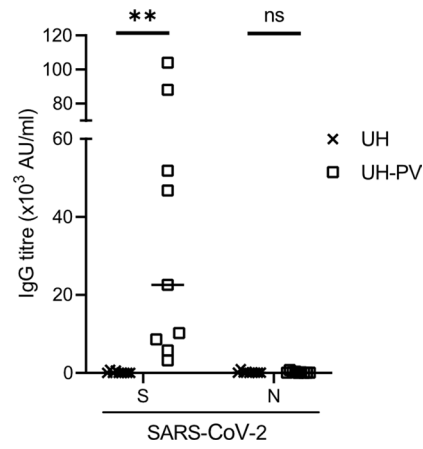
Extended Data Fig. 5 | SARS-CoV-2 S and N-specific IgG antibody titres in uninfected healthy individuals pre and post-vaccination. Multiplex serology assays showing IgG antibody titres against SARS-CoV-2 S and N proteins in UH individuals (n = 9) and the same UH individuals 1–5 months post SARS-CoV-2 vaccination (UH-PV). Significance was determined by two-sided Wilcoxon test **p = 0.0039.

a S1

SARS-CoV-2 `SGTNGTKRFDNPVLP--------FNDGV`

βHCoV
- SARS `T-------FGNPVIP--------FKDGI`
- MERS `TGT----TPQKLFVANYSQDVKQFANGF`
- HKU1 `STL----WYQKPFLSD------FNNGI`
- OC43 `SRL----WFKPPFLSD------FINGI`

SARS-CoV-2 `TRFQTLLALHRSYLTPGDSS`

βHCoV
- SARS `TNFRAIL----TAFSPAQDI`
- MERS `KYYSIIPHSIRSIQSDR-KA`
- HKU1 `SHYYVLPLTCNAISSNTDNE`
- OC43 `SHYYVMPLTCNSKLT-----`

SARS-CoV-2 `GGNYNY----------------------------------------LYRLFRKSNLKPFE`

βHCoV
- SARS `TGNYNY----------------------------------------KYRYLRHGKLRPFE`
- MERS `TKPLKY----------------------------------------SYINKCSRFLSDDR`
- HKU1 `NRRYGF---NNFN------LSSHSVVYSRYCFSVNNTFCPCAKPSFASSCKSHKPPS-`
- OC43 `NKRFGFIEDSVFKPRPAGVLTNHDVVYAQHCFKAPKNFCPC---KLNGSCVGSGPGKN`

SARS-CoV-2 `LVDLPIGINITRFQTLLALH`

βHCoV
- SARS `IFKLPLGINITNFRAIL---`
- MERS `FATLPVYDTIKYYSIIPHSI`
- HKU1 `LFSLYLGTLLSHYYVLPLTC`
- OC43 `LFNVYLGMALSHYYVMPLTC`

SARS-CoV-2 `LSETKCTLKSFTVEKGIYQT`

βHCoV
- SARS `LAELKCSVKSFEIDKGIYQT`
- MERS `LSQLHCSYESFDVESGVYSV`
- HKU1 `FSEIQCKTKSLLPNTGVYDL`
- OC43 `MSEIKCKTQSIAPPTGVYEL`

b S2

SARS-CoV-2 `GAALQIPFAMQMAYRFNGIG`

βHCoV
- SARS `GAALQIPFAMQMAYRFNGIG`
- MERS `SSFAAIPFAQSIFYRLNGVG`
- HKU1 `G----IPFSLNVQYRINGLG`
- OC43 `G----VPFYLNVQYRINGLG`

SARS-CoV-2 `TLVKQLSSNFGAISSVLNDI`

βHCoV
- SARS `TLVKQLSSNFGAISSVLNDI`
- MERS `KLASELSNTFGAISASIGDI`
- HKU1 `SLLQQLFNKFGAISSSLQEI`
- OC43 `NLLQQLSNRFGAISASLQEI`

SARS-CoV-2 `QLIRAAEIRASANLAATKMS`

βHCoV
- SARS `QLIRAAEIRASANLAATKMS`
- MERS `QLVRSESAALSAQLAKDKVN`
- HKU1 `QLSDISLVKFGAALAMEKVN`
- OC43 `QLSDSTLVKFSAAQAMEKVN`

SARS-CoV-2 `AQALNTLVKQLSSNFGAISS`

βHCoV
- SARS `AQALNTLVKQLSSNFGAISS`
- MERS `AQALSKLASELSNTFGAISA`
- HKU1 `AQALNSLLQQLFNKFGAISS`
- OC43 `AEALNNLLQQLSNRFGAISA`

SARS-CoV-2 `TYVTQQLIRAAEIRASANLA`

βHCoV
- SARS `TYVTQQLIRAAEIRASANLA`
- MERS `AFVAQQLVRSESAALSAQLA`
- HKU1 `AYVSQQLSDISLVKFGAALA`
- OC43 `AYVSQQLSDSTLVKFSAAQA`

SARS-CoV-2 `VFLHVTYVPAQEKNFTTAPA`

βHCoV
- SARS `VFLHVTYVPSQERNFTTAPA`
- MERS `YFMHVGYYPSNHIEVVSAYG`
- HKU1 `LFMHFSYKPISFKTVLVSPG`
- OC43 `YFIHFSYVPTKYVTARVSPG`

Extended Data Fig. 6 | Alignment of SARS-CoV-2 spike CD4⁺ T cell epitopes with the corresponding amino acid sequences of other β−HCoVs. Alignments show the aa sequences of SARS-CoV-2 S-derived CD4⁺ T cell epitopes in **(a)** S1 and **(b)** S2 with the corresponding sequences of other β-HCoVs, for those epitopes not shown in Fig. 3. Amino acids highlighted in yellow are conserved with SARS-CoV-2.
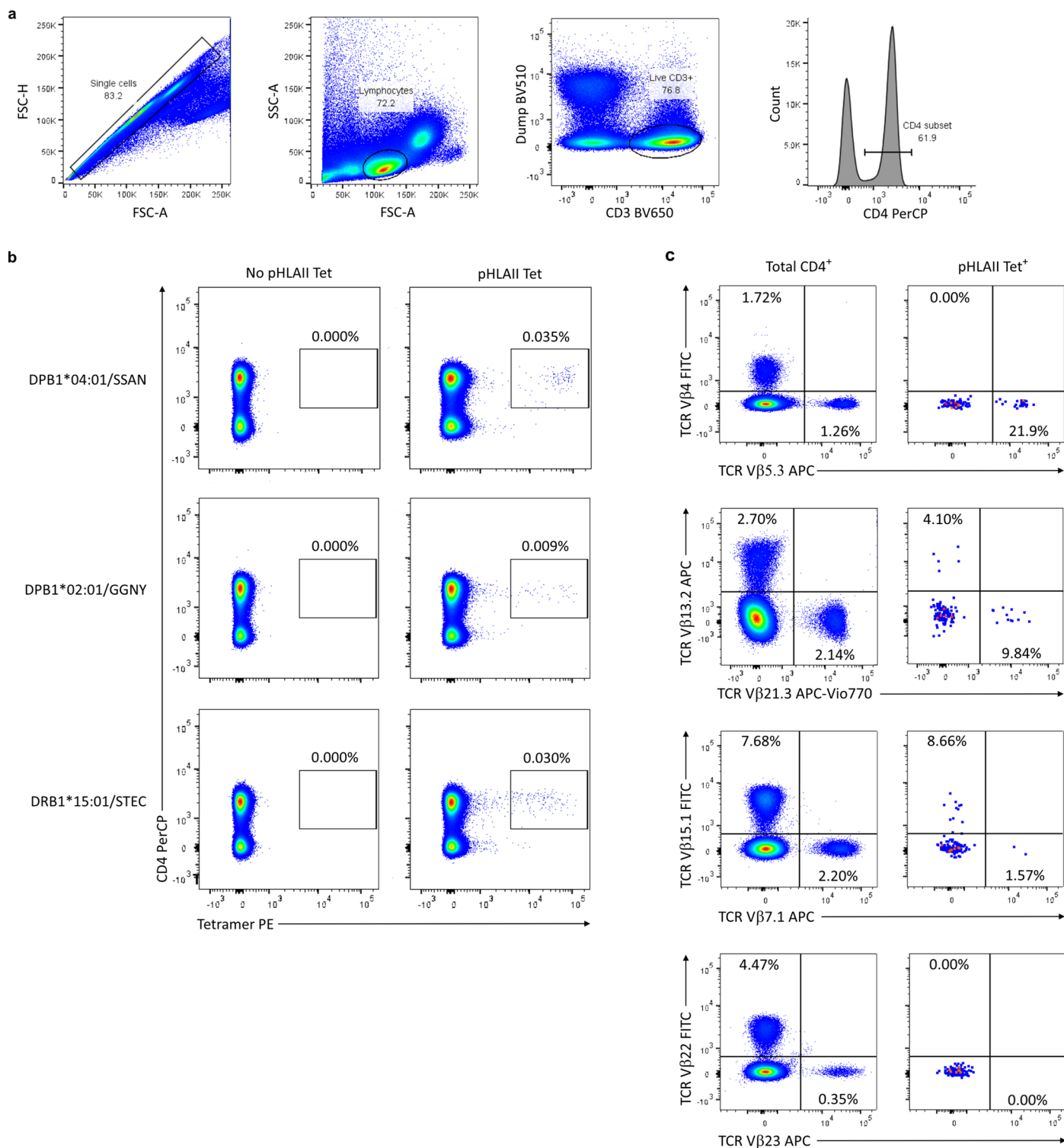
Extended Data Fig. 7 | pHLAII tetramer and TCR Vβ analysis. Flow cytometry of PBMCs stained with pHLAII Tet and antibodies against defined TCR Vβ segments. **(a)** Gating strategy used for analysis. **(b)** Representative flow cytometry plots of pHLAII Tet+ events within the total CD4+ T cell population of PBMCs either exposed to no pHLAII Tet or DPB1*04:01/SSAN, DPB1*02:01/GGNY or DRB1*15:01/STEC pHLAII Tets. **(c)** Representative flow cytometry plots of individual TCR Vβ antibody staining on total CD4+ T cells and pHLAII Tet+ cells in PBMCs stained with the DPB1*04:01/SSAN tetramer.

Corresponding author(s): Dr Heather M Long

Last updated by author(s): Sep 29, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Elispot data was collected using Bioreader 5000 Fgamma (Bio-Sys GmbH) ELISA data was collected using iMark Microplate Reader (Bio-Rad) Antibody titres were measured using MESO Quickplex SQ 120 System (Meso Scale Discovery) Flow Cytometry data was acquired using BD LSR Fortessa X20 (BD Biosciences)
Data analysis	Amino acid alignments were performed using the MUSCLE algorithm (EMBL-EBI) with default settings Protein sequence identity was calculated using Expasy SIM alignment tool for protein sequences with default settings Correlation plots were prepared in R version 4.0.3 using corrplot (version 0.84) Antibody data was analysed using Methodolocial Mind software (version 1.0.36) and MSD Discovery Workbench (version 4.0) Statistical analysis was performed using GraphPad Prism (version 9.3.1) Flow Cytometry data was collected using BD FACSDiva (version 9.0) and analysed using FlowJo (version 10.6.1)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data in this study are available within the article and from the corresponding author on reasonable request. Statistical source data are provided with this paper.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size was not calculated using statistical measures prior to the study. This study was performed on a novel virus, on which we had no data to estimate the required sample size. The sample size is justified based on the range of HLA class II alleles represented in the restrictions of the characterised SARS-CoV-2 epitopes.
Data exclusions	In the Elispot assays, any plates in which the positive controls did not work were excluded.
Replication	Elispot, ELISA and antibody titre assays were all performed using duplicate or triplicate wells. Elispot assays and flow cytometry assays using ex vivo PBMCs were performed once due to limited availability of cells, but were repeated on multiple individuals therefore quantifying biological variability. All ELISA and antibody titre assays were independently repeated and results were always consistent. All attempts at replication were successful.
Randomization	The two groups studied were determined by prior SARS-CoV-2 infection
Blinding	The investigators were not blinded during the data collection and analysis

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	BV510 anti-CD14 clone: MoP9; Cat: 563079; Lot: 2115084; Company: BD Biosciences BV510 anti-CD19 clone: SJ25C1; Cat: 562953; Lot: 1137014; Company: BD Biosciences BV650 anti-CD3 clone: OKT3; Cat: 317324; Lot: B336299; Company: Biolegend PerCP anti-CD4 clone: RPA-T4; Cat: 300528; Lot: B212868; Company: Biolegend FITC anti-TCR Vb3 clone: CH92; Cat: IM2372; Lot: 22; Company: Beckman Coulter FITC anti-TCR Vb4 clone: WJF24; Cat: B07084; Lot: 03; Company: Beckman Coulter FITC anti-TCR Vb5.1 clone: IMMU 157; Cat: IM1552; Lot: 15; Company: Beckman Coulter FITC anti-TCR Vb5.2 clone: 36213; Cat: IM1482; Lot: 18; Company: Beckman Coulter

FITC anti-TCR Vb8 clone: 56C5.2; Cat: IM1233; Lot: 35; Company: Beckman Coulter
FITC anti-TCR Vb13.1 clone: H131; Cat: 362403; Lot: B270870; Company: Biolegend
FITC anti-TCR Vb20 clone: ELL1.4; Cat: IM1562; Lot: 24; Company: Beckman Coulter
FITC anti-TCR Vb22 clone: IMMU 546; Cat: IM1484; Lot: 23; Company: Beckman Coulter
APC anti-TCR Vb1 clone: REA662; Cat: 130-110-019; Lot: 5220507353; Company: Miltenyi Biotech
APC anti-TCR Vb5.3 clone: REA670; Cat: 130-110-166; Lot: 5220507356; Company: Miltenyi Biotech
APC anti-TCR Vb7.1 clone: REA871; Cat: 130-114-349; Lot: 5220507334; Company: Miltenyi Biotech
APC anti-TCR Vb7.2 clone: REA677; Cat: 130-110-170; Lot: 5220507357; Company: Miltenyi Biotech
APC anti-TCR Vb13.6 clone: REA554; Cat: 130-109-071; Lot: 5220507350; Company: Miltenyi Biotech
APC anti-TCR Vb14 clone: REA557; Cat: 130-108-805; Lot: 5220507352; Company: Miltenyi Biotech
APC anti-TCR Vb23 clone: REA497; Cat: 130-107-497; Lot: 5220507349; Company: Miltenyi Biotech
AF647 anti-TCR Vb13.2 clone: H132; Cat: 568087; Lot: 1342662; Company: BD Biosciences
APC-Vio770 anti-TCR Vb2 clone: REA654; Cat: 130-110-098; Lot: 5220507355; Company: Miltenyi Biotech
APC-Vio770 anti-TCR Vb11 clone: REA559; Cat: 130-127-301; Lot: 5220507327; Company: Miltenyi Biotech
APC-Vio770 anti-TCR Vb16 clone: REA553; Cat: 130-108-769; Lot: 5220507351; Company: Miltenyi Biotech
APC-Vio770 anti-TCR Vb17 clone: REA915; Cat: 130-108-769; Lot: 5220507351; Company: Miltenyi Biotech
APC-Vio770 anti-TCR Vb21.3 clone: REA894; Cat: 130-130-105; Lot: 5220600456; Company: Miltenyi Biotech

Validation	All the antibodies used in the study have been validated by the manufacturers. Validation information can be found on the company websites as follows:

BV510 anti-CD14: https://www.bdbiosciences.com/en-gb/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/bv510-mouse-anti-human-cd14.740163
BV510 anti-CD19: https://www.bdbiosciences.com/en-gb/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/bv510-mouse-anti-human-cd19.562953
BV650 anti-CD3: https://www.biolegend.com/en-us/products/brilliant-violet-650-anti-human-cd3-antibody-7667
PerCP anti-CD4: https://www.biolegend.com/en-us/products/brilliant-violet-650-anti-human-cd3-antibody-7667
FITC anti-TCR Vb3: https://www.mybeckman.uk/reagents/coulter-flow-cytometry/antibodies-and-kits/single-color-antibodies/tcr-vb-3/im2372
FITC anti-TCR Vb4: https://www.mybeckman.uk/reagents/coulter-flow-cytometry/antibodies-and-kits/single-color-antibodies/tcr-vb-4/b07084
FITC anti-TCR Vb5.1: https://www.mybeckman.uk/reagents/coulter-flow-cytometry/antibodies-and-kits/single-color-antibodies/tcr-vb-5-1/im1552
FITC anti-TCR Vb5.2: https://www.mybeckman.uk/reagents/coulter-flow-cytometry/antibodies-and-kits/single-color-antibodies/tcr-vb-5-2/im1482
FITC anti-TCR Vb8: https://www.mybeckman.uk/reagents/coulter-flow-cytometry/antibodies-and-kits/single-color-antibodies/tcr-vb-8/im1233
FITC anti-TCR Vb13.1: https://www.biolegend.com/en-us/products/fitc-anti-human-tcr-vbeta13-1-antibody-10215
FITC anti-TCR Vb20: https://www.mybeckman.uk/reagents/coulter-flow-cytometry/antibodies-and-kits/single-color-antibodies/tcr-vb20/im1562
FITC anti-TCR Vb22: https://www.mybeckman.uk/reagents/coulter-flow-cytometry/antibodies-and-kits/single-color-antibodies/tcr-vb22/im1484
APC anti-TCR Vb1: https://www.miltenyibiotec.com/GB-en/products/tcr-vb1-antibody-anti-human-reafinity-rea662.html#apc:100-tests-in-200-ul
APC anti-TCR Vb5.3: https://www.miltenyibiotec.com/GB-en/products/tcr-vb5-3-antibody-anti-human-reafinity-rea670.html#apc:100-tests-in-1-ml
APC anti-TCR Vb7.1: https://www.miltenyibiotec.com/GB-en/products/tcr-vb7-1-antibody-anti-human-reafinity-rea871.html#apc:100-tests-in-200-ul
APC anti-TCR Vb7.2: https://www.miltenyibiotec.com/GB-en/products/tcr-vb7-2-antibody-anti-human-reafinity-rea677.html#apc:100-tests-in-1-ml
APC anti-TCR Vb13.6: https://www.miltenyibiotec.com/GB-en/products/tcr-vb13-6-antibody-anti-human-reafinity-rea554.html#apc:30-tests-in-300-ul
APC anti-TCR Vb14: https://www.miltenyibiotec.com/GB-en/products/tcr-vb14-antibody-anti-human-reafinity-rea557.html#apc:100-tests-in-1-ml
APC anti-TCR Vb23: https://www.miltenyibiotec.com/GB-en/products/tcr-vb23-antibody-anti-human-reafinity-rea497.html#apc:30-tests-in-300-ul
AF647 anti-TCR Vb13.2: https://www.bdbiosciences.com/en-gb/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/alexa-fluor-647-mouse-anti-human-tcr-v-13-2.568087
APC-Vio770 anti-TCR Vb2: https://www.miltenyibiotec.com/GB-en/products/tcr-vb2-antibody-anti-human-reafinity-rea654.html#apc-vio-770:30-tests-in-300-ul
APC-Vio770 anti-TCR Vb11: https://www.miltenyibiotec.com/GB-en/products/tcr-vb11-antibody-anti-human-reafinity-rea559.html#apc-vio-770:100-tests-in-200-ul
APC-Vio770 anti-TCR Vb16: https://www.miltenyibiotec.com/GB-en/products/tcr-vb16-antibody-anti-human-reafinity-rea553.html#apc-vio-770:100-tests-in-1-ml
APC-Vio770 anti-TCR Vb17: https://www.miltenyibiotec.com/GB-en/products/tcr-vb16-antibody-anti-human-reafinity-rea553.html#apc-vio-770:100-tests-in-1-ml
APC-Vio770 anti-TCR Vb21.3: https://www.miltenyibiotec.com/GB-en/products/tcr-vb21-3-antibody-anti-human-reafinity-rea894.html#apc-vio-770:100-tests-in-200-ul

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	B lymphoblastoid cell lines (LCLs) from the donors used in the study were generated in-house from PBMCs by transformation with B95.8 EBV.

Authentication	All LCLs were tested by PCR using primers specific for HLAII alleles possessed by each individual's HLAII genotype. All LCLs were positive for the expected HLA alleles.
Mycoplasma contamination	All cell lines tested negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Individuals with prior SARS-CoV-2 infection included 10 males and 10 females between the ages of 28 and 64. Uninfected healthy individuals included 6 males and 8 females between the ages of 24 and 63. The uninfected healthy individuals who remained uninfected following vaccination included 3 males and 6 females between the ages of 39 and 63.
Recruitment	Individuals with prior SARS-CoV-2 infection were all healthcare workers. Donors were recruited in the early period of the COVID-19 pandemic, between 1.5 and 6 months post infection. Uninfected individuals were healthy volunteers with no history of SARS-CoV-2 infection, recruited from staff working at the University of Birmingham. All donors were recruited by personal request from the research team. This biased the selection to individuals of working age who were known to the research team, resulting in an overall age range restricted to 24-64. Birmingham is an ethnically diverse city and the recruited individuals reflected this diversity. All participants gave written informed consent and received no compensation.
Ethics oversight	The study was approved by the North West - Preston Research Ethics Committee, UK. Written informed consent was given by all donors.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	PBMCs were isolated from heparinised whole blood by standard Ficoll-Hypaque centrifugation. PBMCs were cryopreserved in RPMI containing 10% DMSO and 20% FCS prior to use. Cells were defrosted on the day of the experiments and used immediately.
Instrument	BD LSR Fortessa X20 (BD Biosciences)
Software	BD FACSDiva (Version 9.0) was used to collect the data. FlowJo version 10.6.1 was used to analyse the data
Cell population abundance	All staining was performed on whole PBMCs with no prior selection. Frequencies of the cells of interest are shown on each figure.
Gating strategy	Single cells were gated in a forward scatter-height (FSC-H) versus forward scatter-area (FSC-A) plot to remove doublets. Lymphocytes were gated using a side scatter-area (SSC-A) versus FSC-A plot. A single dump channel was used to exclude dead cells, CD14+ events and CD19+ events, allowing gating of viable CD3+ cells. To analyse TCR Vbeta segment expression on total CD4+ T-cells, CD4+ cells were gated using a histogram of CD4 expression

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.