Identification and characterization of a SARS-CoV-2 specific CD8+ T cell response with immunodominant features

Anastasia Gangaev1,6, Steven L. C. Ketelaars1,6, Olga I. Isaeva1, Sanne Patiwaël1, Anna Dopler1, Kelly Hoefakker1, Sara De Biasi2, Lara Gibellini2, Cristina Mussini2, Giovanni Guaraldi2, Massimo Girardis2, Cami M. P. Talavera Ormeno3, Paul J. M. Hekking3, Neubury M. Lardy4, Mireille Toebes1, Robert Balderas5, Ton N. Schumacher1, Huib Ovaa3, Andrea Cossarizza2 & Pia Kvistborg1

The COVID-19 pandemic caused by SARS-CoV-2 is a continuous challenge worldwide, and there is an urgent need to map the landscape of immunogenic and immunodominant epitopes recognized by CD8+ T cells. Here, we analyze samples from 31 patients with COVID-19 for CD8+ T cell recognition of 500 peptide-HLA class I complexes, restricted by 10 common HLA alleles. We identify 18 CD8+ T cell recognized SARS-CoV-2 epitopes, including an epitope with immunodominant features derived from ORF1ab and restricted by HLA-A*01:01. In-depth characterization of SARS-CoV-2-specific CD8+ T cell responses of patients with acute critical and severe disease reveals high expression of NKG2A, lack of cytokine production and a gene expression profile inhibiting T cell re-activation and migration while sustaining survival. SARS-CoV-2-specific CD8+ T cell responses are detectable up to 5 months after recovery from critical and severe disease, and these responses convert from dysfunctional effector to functional memory CD8+ T cells during convalescence.
The coronavirus disease 2019 (COVID-19) pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is an ongoing global emergency. First cases of COVID-19 were reported in December 2019, and as of February 4, 2021, there are more than 104,400,000 confirmed cases and 2,264,000 deaths worldwide\(^1\). Due to the measures necessary to contain the rapid spread of the infection, this pandemic has having tremendous health and socioeconomic consequences, and there is an urgent need for a better understanding of the natural adaptive immune response.

Accumulating information on the antibody response against SARS-CoV-2 demonstrates that structural proteins, in particular the spike protein and nucleoprotein, serve as potent antibody targets in a substantial fraction of COVID-19 patients\(^2\)-\(^7\). However, accumulating data provide evidence for a broader repertoire of immunogenic T cell targets from SARS-CoV-2\(^8\)-\(^15\). T cell reactivity against SARS-CoV-2 based on peptide pools has been demonstrated by various groups\(^8\)-\(^11\),\(^13\)-\(^16\), and several studies have identified specific epitopes and their restriction elements\(^2\),\(^4\),\(^12\),\(^17\),\(^18\). However, our understanding regarding which T cell recognized SARS-CoV-2 epitopes are most immunogenic during acute COVID-19 disease is currently limited, as most studies that have identified specific SARS-CoV-2 epitopes are based on convalescent samples\(^12\),\(^17\),\(^18\). Knowledge regarding how well the antigens targeted by T cells reflect the antigen composition of current vaccine candidates, and whether these vaccines include immunodominant epitopes, is central for next-generation vaccine development\(^19\),\(^20\).

Several studies have characterized the composition of immune cell lineages in COVID-19 patients, which showed that bulk T cells have impaired effector functions (i.e., cytokine production) and higher expression levels of inhibitory receptors compared to bulk T cells from healthy individuals, which is worsening with disease stage\(^2\),\(^1\)\(^1\)-\(^2\)\(^3\). However, in-depth characterization of the SARS-CoV-2-specific CD8 T cell response during acute COVID-19 disease directly ex vivo is currently lacking. Such information is essential to improve our understanding about the role of CD8 T cells in the host defense against SARS-CoV-2 and during COVID-19 disease development.

In this work, we probe for CD8 T cell recognition toward 500 SARS-CoV-2 epitope human leukocyte antigen (HLA) complexes, restricted by 10 common HLA class I alleles. A substantial fraction of the identified CD8 T cell responses is directed towards epitopes derived from the open reading frame 1ab polypeptide (ORF1ab), including an epitope with immunodominant characteristics restricted by HLA-A\(^*\)01:01. In-depth characterization of identified SARS-CoV-2-specific CD8 T cell responses reveals high expression of NKG2A, lack of cytokine production, and a gene expression profile of constrained T cell re-activation. In addition, we show that SARS-CoV-2-specific CD8 T cell responses convert from dysfunctional effector CD8 T cells to functional memory CD8 T cells during convalescence and persist for up to 5 months post COVID-19 disease.

**Results**

**Epitope selection.** To cover as many HLA alleles as possible in a patient-specific and high-throughput manner, we focused our analysis on 10 common HLA alleles of the Italian population (samples analyzed in this study were collected from COVID-19 patients in Italy) that could be covered with our peptide-HLA (pHLA) multimer technology. This collection of HLA-A (HLA-A\(^*\)01:01, HLA-A\(^*\)02:01, HLA-A\(^*\)03:01, HLA-A\(^*\)11:01, and HLA-A\(^*\)24:02) and HLA-B alleles (HLA-B\(^*\)07:02, HLA-B\(^*\)08:01, HLA-B\(^*\)15:01, HLA-B\(^*\)18:01, and HLA-B\(^*\)51:01) resulted in a coverage of approximately 95% of the Italian population (http://www. allelefrequencies.net). For each HLA allele, 50 SARS-CoV-2 epitopes derived from the entire SARS-CoV-2 proteome were selected (Fig. 1a). Epitope selection was based on likelihood of successful proteasomal processing (NetChop-3.1)\(^2\)\(^4\) and predicted binding affinity to HLA (NetMHCpan-4.0)\(^2\)\(^5\). In addition, SARS-CoV-2 epitopes previously predicted by the science community were included, as well as epitopes shared between SARS-CoV-1 and SARS-CoV-2 for which T cell reactivity has previously been reported (Supplementary Data 1).

**SARS-CoV-2-specific CD8 T cell responses.** To investigate which of the included epitopes from SARS-CoV-2 were recognized by CD8 T cells, peripheral blood mononuclear cell (PBMC) samples from 31 COVID-19 patients with confirmed SARS-CoV-2 infection were analyzed (Table 1 and Supplementary Table 1). Samples of patients with acute critical (\(n = 14\)), severe (\(n = 10\)), and moderate (\(n = 2\)) COVID-19 disease were collected on average 10 days (range –2 to 29) after hospitalization. Samples of asymptomatic patients (\(n = 5\)) were collected during convalescence (3 months post positive polymerase chain reaction (PCR) test for SARS-CoV-2). SARS-CoV-2-specific CD8 T cell reactivity was evaluated directly ex vivo, using our in-house developed technology based on multiplexing of pHLA multimers conjugated to fluorescent streptavidin reagents. We have previously used this technology to identify tumor-specific CD8 T cell responses in cancer patients\(^2\),\(^6\),\(^2\)\(^7\). A total of 35 SARS-CoV-2-specific CD8 T cell responses specific for 18 different SARS-CoV-2-derived epitopes were detected (Fig. 1b, c and Supplementary Fig. 1). SARS-CoV-2-specific CD8 T cell responses were detected in half of the analyzed patients (17 of 31). The average magnitude of the detected SARS-CoV-2-specific CD8 T cell responses was 1.7% (range: 0.005–19%) of total CD8\(^+\) cells (Supplementary Table 2), and the majority (22 of 35) of the responses was restricted by HLA\(^*\)01:01 (Fig. 1d). CD8 T cell responses specific for epitopes derived from the ORF1ab were of significantly higher magnitude compared to CD8 T cell responses specific for the ORF3a, spike, membrane, and nucleoprotein (Fig. 1e, \(P = 0.0309\), two-tailed Mann–Whitney \(U\) test).

**CD8 T cell recognized SARS-CoV-2 epitopes.** We identified 18 different SARS-CoV-2 epitopes that were recognized by CD8 T cells from COVID-19 patients across 6 of the 10 included HLA alleles (Fig. 1f). Eight of these epitopes were specific for SARS-CoV-2, while 10 epitopes were shared with SARS-CoV-1 and 1 of these epitopes was shared with 3 of the “common cold” coronaviruses (Supplementary Table 3). The origin of these epitopes was diverse, including epitopes derived from the ORF1ab (\(n = 8\)), spike protein (\(n = 6\)), membrane protein (\(n = 2\)), nucleoprotein (\(n = 1\)), and ORF3a (\(n = 1\)). Of note, although a high fraction of the selected epitopes was derived from the ORF1ab (driven by the large size of the ORF1ab within the SARS-CoV-2 proteome) we observed an enrichment for CD8 T cell-recognized epitopes derived from the spike protein (Fig. 1g).

**Identified SARS-CoV-2 epitopes are conserved across virus isolates.** Next, we examined whether CD8 T cell-recognized SARS-CoV-2 epitopes were located in positions of the SARS-CoV-2 proteome with a high level of non-synonymous single-nucleotide variants (SNVs). Such information is relevant for the broader utilization of the obtained information. For this

2 NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-22811-y | www.nature.com/naturecommunications
purpose, we obtained information on 385,154 globally isolated SARS-CoV-2 sequences using the COVID-19 CoV Genomics tool\textsuperscript{28}. Alternative epitopes with a median SNV frequency of 0.041% (range: 0.01\%–8.44\%) have been reported for 17 of the 18 epitopes that were identified in this study (Supplementary Table 3). Hotspot mutations (SNV frequency cut-off of 2.5\% to exclude background mutations and possible sequencing artifacts) were reported for only 1 (IPTNFTISV) of the 17 epitopes. Interestingly, this hotspot mutation is one of the defining mutations of the recently identified SARS-CoV-2 B117 variant (T716I: IPTNFTISV, global alternative SNV frequency 8.44\%; https://www.cogconsortium.uk). It is unlikely that the low impact of the T716I
Fig. 1 Identification of SARS-CoV-2-specific CD8 T cell responses. a Overview of the selected SARS-CoV-2 epitopes derived from the different SARS-CoV-2 ORFs for each of the 5 HLA-A and 5 HLA-B alleles included in this study. b Representative flow cytometric plots of two different SARS-CoV-2-specific CD8 T cell responses detected in patient COVID-153. Magnitude of the response (percentage of double-positive pHLA<sup>+</sup> cells [green] of total CD<sup>+</sup> cells [gray]) is indicated. Representative gating strategy is provided in Supplementary Fig. 1. c Heatmap of detected SARS-CoV-2-specific CD8 T cell responses (n = 35) including information about CD8 T cell–recognized SARS-CoV-2 epitopes (n = 18) and their viral protein origin, the magnitude of the response (pHLA<sup>+</sup> cells of total CD<sup>+</sup> cells), and HLA coverage as well as the disease status for each COVID-19 patient (n = 31) and healthy donor (n = 7). Confirmation of identified responses was done for patients COVID-004, 087, 096, 121, 127, 143, and 153 with similar results. d Overview of the magnitude, target epitope origin, and number of detected SARS-CoV-2-specific CD8 T cell responses (n = 35) across the included HLA alleles. Source data are provided as a Source data file. e Difference in the magnitude of SARS-CoV-2-specific CD8 T cell responses targeting epitopes derived from ORF1ab (n = 20) compared to CD8 T cell responses targeting epitopes derived from all other SARS-CoV-2 proteins (n = 15) combined. All detected responses are included. Box plots indicate the median (line), 25th and 75th percentile (box), min and max (whiskers), and all data points (single circles). Statistical significance was tested with a two-tailed Mann–Whitney U test, P = 0.0309. Source data are provided as a Source data file. f Number of SARS-CoV-2 epitopes (n = 18) that were recognized by CD8 T cells across different HLA alleles. g Bar graphs illustrating the contribution of each ORF to the SARS-CoV-2 proteome (left) in comparison to the fraction of the selected epitopes per ORF based on our predictions (middle) and the contribution of each ORF to the CD8 T cell–recognized epitopes (right). ORF open reading frame, S spike, N nucleoprotein, M membrane, pHLA peptide human leukocyte antigen, PTD PTENFTISV29.

Table 1 Characteristics of COVID-19 patients included in the analysis of SARS-CoV-2-specific CD8 T cell responses.

| Outcome, n (%) | Critical | Severe | Moderate | Asymptomatic | Healthy |
|----------------|----------|--------|----------|--------------|---------|
| Total number of patients (n) | 14 | 10 | 2 | 5 | 7 |
| Median age, years (range) | 70 (34 to 77) | 62 (37 to 83) | 64 (36 to 88) | 33 (23 to 51) | 69 (41 to 77) |
| Gender, n (%) | | | | | |
| Female | 2 (14%) | 3 (42%) | 0 (0%) | 0 (0%) | 3 (42%) |
| Male | 12 (86%) | 7 (58%) | 2 (100%) | 5 (100%) | 4 (58%) |
| Treatment, n (%) | | | | | |
| None | 0 (0%) | 0 (0%) | 2 (100%) | 5 (100%) | 7 (100%) |
| Anakinra | 3 (21%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |
| Tocilizumab | 11 (79%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |
| Average length of hospitalization, weeks (range) | 5 (2 to 28) | 4 (1 to 16) | 2 (1 to 3) | N/A | N/A |
| Outcome, n (%) | | | | | |
| Deceased | 4 (28%) | 1 (10%) | 1 (50%) | 0 (0%) | 0 (0%) |
| Alive | 10 (72%) | 9 (90%) | 1 (50%) | 7 (100%) | 8 (100%) |
| Average time of sample collection, days (range) | | | | | |
| After hospitalization | 12 (–2 to 24) | 7 (1 to 29) | 3 (2 to 3) | N/A | N/A |
| After recovery | N/A | N/A | N/A | 12 | |
| HLA coverage, n (%) | | | | | |
| Alleles, 1 | 15 (50%) | 6 (60%) | 1 (50%) | 1 (50%) | 3 (43%) |
| Alleles, 2 | 5 (35%) | 1 (10%) | 0 (0%) | 4 (80%) | 1 (14%) |
| Alleles, 3 | 2 (15%) | 3 (30%) | 1 (50%) | 1 (20%) | 3 (43%) |
| Number of patients with DR, n (%) | 6 (43%) | 6 (60%) | 1 (50%) | 4 (80%) | 1 (14%) |
| Total number of DR, n | 14 | 12 | 1 | 8 | 1 |
| Median magnitude of DR, % (range) | 2.5% (0.006 to 19%) | 2.059% (0.017 to 18%) | 0.038% | 0.025 (0.005 to 0.13%) | 0.008% |

Detailed information for individual patients is provided in Supplementary Table 1. DR detected SARS-CoV-2-specific CD8 T cell response, HLA human leukocyte antigen, N/A not available/applicable.

Amino acid change will hamper the recognition by T cell receptors (TCRs) specific for IPTNFTISV<sup>29–31</sup>; however, this needs to be experimentally tested in future studies. Overall, our analysis demonstrates that the identified SARS-CoV-2-derived CD8 T cell epitopes are highly conserved across virus variants.

The ORF1ab encodes an epitope with immunodominant features. Strikingly, CD8 T cell responses specific for the TTDPFLGGRY (TTD) epitope were detected in all HLA-A*01:01+ patients (8/8) with acute COVID-19 disease (Supplementary Table 3). The magnitude of these CD8 T cell responses was remarkably high: on average >7% of total CD<sup>+</sup> cells (range: 0.074–19%, Fig. 2a and Supplementary Fig. 2). Furthermore, these CD8 T cell responses were of significantly higher magnitude compared to other detected SARS-CoV-2-specific CD8 T cell responses in all COVID-19 patients with acute COVID-19 disease (Fig. 2a, P = 0.0003, two-tailed Mann–Whitney U test). Additional SARS-CoV-2-specific CD8 T cell responses were identified for 5 of the 8 HLA-A*01:01+ COVID-19 patients, and the highest magnitude of these responses was on average 70-fold (range: 1.2–216) lower compared to the TTD-specific CD8 T cell response (Fig. 2b, c, P = 0.0136, two-tailed Mann–Whitney U test). The immunodominance hierarchy is partially determined by the naive precursor frequency<sup>32,33</sup>. To further investigate whether more than one T cell clone gave rise to the TTD-specific CD8 T cell response, TCR sequencing data on 542 TTD-specific CD8 T cells obtained from 5 patients was analyzed. In line with the immunodominance pattern based on the presence and the high magnitude of the TTD-specific CD8 T cell responses in all HLA-A*01:01+ patients, we observed a high level of TCR diversity with 150 unique TCR beta-complementarity-determining region 3 (TRB-CDR3) sequences that were expressed in only 1 cell vs. 75 unique TRB-CDR3 sequences that were expressed in ≥2 cells (information for TRB-CDR3 sequences was available for 431/542 cells, Fig. 2d). In addition, we observed an enrichment of TCRs with the TCR beta chain V<sub>27</sub> segment within the TTD-specific TCR sequences compared to the TCR repertoire of bulk CD8 T cells (Fig. 2e) suggesting that TCRs specific for the TTD epitope may more likely to be rearranged during T cell development. Together, these data strongly indicate that (depending on the composition of the entire set of HLA alleles) TTDPSFLGGRY is an immunodominant epitope for the subgroup of HLA-A*01:01+ patients, which is present in approximately 30% of the European population (http://www.allelefrequencies.net).
Dysfunctional state of SARS-CoV-2-specific CD8 T cells. To characterize the identified SARS-CoV-2-specific CD8 T cell response, we first assessed the functionality based on cytokine production of interferon-γ (IFNγ), tumor necrosis factor (TNF), interleukin (IL)-2, and IL-17 upon stimulation with cognate peptide. The functional assay was focused on the TTD-specific CD8 T cell responses that were detected in 5 COVID-19 patients with acute COVID-19 disease. The TTD-specific CD8 T cell responses were found to have no (4 patients) or very low levels of cytokine production (1 patient) after 12 h of stimulation (Fig. 3a, b and Supplementary Fig. 3a–c). From one patient, sufficient cell numbers were available to include a cytomegalovirus (CMV)
epitope, and CD8 T cells reactive toward this epitope were found to be functional (Supplementary Fig. 3d, e). Of note, the technical control using phorbol 12-myristate 13-acetate and Ionomycin led to the production of IFNγ, TNF, and IL-2 (Supplementary Fig. 3f). Importantly, these results demonstrate that detection of these responses with a functional readout failed or vastly underestimated the magnitude of the TTD-specific CD8 T cell responses during acute disease. To further examine the SARS-CoV-2-specific CD8 T cells, we measured the expression levels (ex vivo) of the inhibitory receptors NKG2A and programmed cell death protein 1 (PD-1) on SARS-CoV-2-specific CD8 T cells compared to bulk CD8 T cells from COVID-19 patients with acute disease. The fraction of NKG2A+ cells was significantly higher for SARS-CoV-2-specific CD8 T cells compared to bulk CD8 T cells but not for PD-1 (Fig. 3c, d, NKG2A: P = 0.0427, PD-1: P = 0.7651, two-tailed Mann–Whitney U test). In addition, we observed a positive correlation between the magnitude of the response and the expression levels (mean fluorescence intensity (MFI)) of NKG2A and PD-1 (Fig. 3e, NKG2A: r = 0.5, P = 0.0199, PD-1: r = 0.5, P = 0.0361, Spearman correlation analysis). Together, these data suggest that the identified TTD-specific CD8 T cell responses in patients with ongoing severe and critical COVID-19 disease were highly dysfunctional and that SARS-CoV-2-specific CD8 T cell responses displayed a high level of activation/inhibitory receptors.

Transcriptomics show constrained T cell re-activation profile. To further characterize the state of the TTD-specific CD8 T cell responses during acute COVID-19 disease, we performed single-cell RNA sequencing on bulk and TTD-specific CD8 T cells from COVID-19 patients with acute severe and critical disease using the 10x Genomics Chromium system (Supplementary Fig. 4a). We obtained data on a total of 3064 CD8 T cells from 6 patients with acute COVID-19 disease (Table 2). The obtained cells were sorted and sequenced in two independent experiments (batch I and batch II). In the first experiment, we obtained data on bulk CD8 T cells from 5 patients (batch I: COVID-087, 096, 117, 143, 153, Supplementary Fig. 4b) and on TTD-specific CD8 T cells from 4 of these patients (COVID-087 was negative for HLA-A*01:01 and did therefore not harbor TTD-specific CD8 T cells). In the second experiment, we obtained data from 1 additional patient including bulk and TTD-specific CD8 T cells (batch II: COVID-131, Supplementary Fig. 4b). Louvain clustering based on highly variable genes resulted in six and eight clusters in batch I and batch II, respectively (Fig. 4a and Supplementary Data 2). We compiled a consensus list of well-established T cell differentiation gene markers (Supplementary Data 3), which allowed us to discriminate naive (batch I and batch II: C1) from effector (batch I: C2–C6 and batch II: C2–C7) and memory CD8 T cells (C8 only clearly separated in batch II, Fig. 4b).

Data on 542 TTD-specific CD8 T cells was obtained for 5 of the 6 patients (Table 2), with the vast majority of cells derived from 2 patients (batch I: COVID-096 and batch II: COVID-131). TTD-specific CD8 T cells were predominantly present in cluster C2, C4, and C6 in batch I, while TTD-specific CD8 T cells in batch II formed a separate cluster (C2, Fig. 4c). This difference in clustering is likely based on the difference in numbers of TTD-specific CD8 T cells (batch I: n = 48, batch II: n = 494). Subsequently, we performed differential gene expression analysis between TTD-specific CD8 T cells and bulk CD8 T cells from the naive cluster (C1). The analysis resulted in 175 and 2594 differentially expressed genes in TTD-specific CD8 T cells in batch I and batch II, respectively (Supplementary Fig. 4c and Supplementary Data 4). Gene ontology analysis revealed the presence of several immune response-related processes, including T cell activation, regulation of T cell activation, adhesion, and proliferation in the TTD-specific CD8 T cells compared to the naive cells (Supplementary Data 4). Next, we assessed the difference in gene expression between TTD-specific CD8 T cells and the effector/memory bulk CD8 T cells (C2–C8). We focused our analysis on genes that were identified in both data sets to be either upregulated (n = 12) or downregulated (n = 209) in TTD-specific CD8 T cells compared to bulk effector/memory CD8 T cells (Supplementary Fig. 4d and Supplementary Data 5). Upregulated genes were associated with T cell survival (e.g., CD27) and TCR signaling (e.g., PAG1 and COTL1) while downregulated genes were associated with T cell activation (e.g., TNFRSF9 [CD137] and LMNA), migration to epithelial tissue and site of inflammation (e.g., CXCR2, CCL3), proliferation (e.g., MYBLL1, CD2K, and USP37), and effector function (e.g., GNLY). Together, these results suggest that activated TTD-specific CD8 T cells display a gene expression program of maintained cell survival but restricted T cell (re)activation, proliferation, and migration, which is in line with the lack of ability to produce cytokines in response to peptide stimulation.

SARS-CoV-2-specific CD8 T cells convert to functional memory cells. To investigate whether SARS-CoV-2-specific CD8 T cell responses persist after recovery from acute COVID-19 disease, we first analyzed the kinetics of the identified CD8 T cell responses from 3 COVID-19 patients during acute disease and convalescence (4–5 months post hospital discharge). All of the six previously identified SARS-CoV-2-specific CD8 T cell responses were detected during convalescence (Fig. 5a, b). The magnitude of the SARS-CoV-2-specific CD8 T cell responses (n = 5) detected in COVID-96 and -143 was on average 11-fold lower during convalescence compared to acute COVID-19 disease. Nevertheless, the magnitude of the TTD-specific CD8 T cell response in...
COVID-143 remained considerably high (>1% of total CD8 T cells) during convalescence. For patient COVID-117, the magnitude of the TTD-specific CD8 T cell response was unaltered during convalescence. Phenotypic characterization of these responses by flow cytometry revealed that the percentage of SARS-CoV-2-specific CD8 T cells expressing activation and effector T cell markers (HLA-DR, CD95, and CXCR3) was significantly decreased. In contrast, the percentage of SARS-CoV-2-specific CD8 T cells expressing the T cell differentiation markers CD45RA and CCR7 (albeit at a lower level compared to bulk naive T cells) was significantly increased during convalescence (Fig. 5c, d and Supplementary Fig. 5c). In addition, SARS-CoV-2-specific CD8 T cells displayed lower expression levels (MFI) of NKG2A on total CD8 T cells and the magnitude of SARS-CoV-2-specific CD8 T cell responses. Two-tailed Spearman correlation analysis was performed, NKG2A: P = 0.0199, r = 0.5, PD-1: P = 0.0361, r = 0.5. TTD TTDPSFLGRY, pHLA peptide human leukocyte antigen. *P < 0.05. Source data are provided as a Source data file.
During acute disease, TTD-specific CD8 T cells from both patients gained the capacity to secrete IFNγ and TNF during convalescence (Fig. 5g, h). Together, these results demonstrate that SARS-CoV-2-specific CD8 T cell responses can be detected up to 5 months post recovery from acute severe and critical COVID-19 disease and that these cells convert from dysfunctional effector CD8 T cells to functional memory CD8 T cells during convalescence.

### Table 2 Number of CD8 T cells and TTD-specific CD8 T cells obtained from single-cell RNA-sequencing per patient.

| Batch       | Patient     | CD8 T cells (n) | CD8 T cells with TCR (n) | TTD-specific CD8 T cells (n) | TTD-specific cells with TCR (n) |
|-------------|-------------|-----------------|--------------------------|-----------------------------|-------------------------------|
| Batch I     | COVID-096   | 1074            | 799                      | 33                          | 23                            |
| Batch I     | COVID-087   | 35              | 21                       | 0                           | 0                             |
| Batch I     | COVID-153   | 33              | 25                       | 8                           | 5                             |
| Batch I     | COVID-143   | 26              | 18                       | 5                           | 2                             |
| Batch I     | COVID-117   | 12              | 11                       | 2                           | 2                             |
| Batch II    | COVID-131   | 1884            | 1479                     | 494                         | 407                           |

**Fig. 4 Single-cell transcriptome analysis of SARS-CoV-2-specific CD8 T cells.**

- **a** UMAP representations of cell clusters (batch I: n = 6 and batch II: n = 8) of the single-cell gene expression data of CD8 T cells (batch I: n = 1180 and batch II: n = 1884) isolated from COVID-19 patients (batch I: n = 5 and batch II: n = 1). Clusters were identified using the Louvain algorithm. Full list of differentially expressed genes for each cluster is provided in Supplementary Data 2.
- **b** UMAP representations of enriched gene signatures (Supplementary Data 3) used to identify naive, memory, and effector CD8 T cells in batch I and batch II.
- **c** UMAP representations of single-cell gene expression data of TTD-specific CD8 T cells (batch I: n = 48 and batch II: n = 494) isolated from COVID-19 patients (batch I: n = 4 and batch II: n = 1).
- **d** Differentially expressed genes that were found to be significantly (FDR < 0.05) upregulated (ln fold change of >1) or downregulated (ln fold change < −1) in TTD-specific CD8 T cell compared to non-naive CD8 T cells in batch I and batch II. The full list of genes is provided in Supplementary Data 5. Statistical significance was tested with a two-tailed *t*-test; multiple correction was performed with the Benjamini-Hochberg procedure. Original *P* values are displayed on a log10 x-axis. TTD TTDPSFLGRY.
Discussion

We employed a systematic effort to identify CD8 T cell–recognized SARS-CoV-2 epitopes by covering 10 common HLA alleles and a selection of in silico predicted peptides from the entire SARS-CoV-2 proteome resulting in 500 unique pHLA complexes. Our analyses resulted in the identification of 18 SARS-CoV-2-derived CD8 T cell epitopes. The origin of these epitopes was diverse, including the ORF1ab, spike protein, membrane protein, nucleoprotein, and the ORF3a. Eight of the 18 identified epitopes were derived from the ORF1ab including one epitope (TTDPSFLGGRY), which displayed strong immunodominant characteristics. First, CD8 T cell responses specific for this epitope were detected in all HLA-A*01:01+ patients with acute COVID-19 disease. Second, the magnitude of TTD-specific CD8 T cell responses was remarkably high (average: 7%, range: 0.074–19%) and substantially higher compared to the other SARS-CoV-2-specific CD8 T cell responses across and within most patients. Notably, such remarkably high magnitude of SARS-CoV-2-specific CD8 T cell responses has not been reported in previous studies. Third, TCR sequencing data of the TTD-specific CD8 T cell responses revealed a high level of clonal diversity based on TRB-CDR3 regions, together with an enrichment for the TRBV27 segment suggesting that these TCR rearrangements may occur more frequently similar to what is known for, e.g., TCRs specific for the immunodominant CMV epitope NLVPMVATV.
sequences specific for an immunodominant SARS-CoV-2 epitope. Such information can pave the way for T cell therapies in COVID-19 patients unable to mount protective T cell immunity in response to, e.g., vaccination. Although our analysis was limited by the selection of HLA alleles and the number of epitopes, as well as imperfect in silico epitope predictions, accumulating evidence from other groups and our data demonstrate that epitopes derived from other parts of the SARS-CoV-2 proteome than the spike protein can induce high-magnitude CD8 T cell responses. It is currently not known whether any of the epitopes derived from the spike protein have immunodominant properties in the absence of, e.g., the ORF1ab, and it remains to be established whether CD8 T cell responses raised against the spike protein alone are sufficient to induce long-lasting protective T cell immunity. The spike protein is the focus of the currently approved vaccines against SARS-CoV-2, and future studies assessing vaccine-induced T cell responses will be able to answer such outstanding questions. A number of recent studies provide evidence suggesting that individual genetic variation in HLA type may predict COVID-19 outcome across a population. Such a scenario is well described for, e.g., HLA-B*57:01+ HIV patients who are more likely to be long-term non-progressors. One study on COVID-19 patients showed that HLA-A*01:01 and HLA-A*02:01 were associated with higher risk for severe disease in Italy; however, this analysis was not based on separate HLA alleles but rather combinations. Based on our data, it appears that there may be a bias for HLA alleles capable of presenting epitopes, which can induce CD8 T cell responses of profound magnitude. We detected 27 SARS-CoV-2-specific CD8 T cell responses in samples collected during acute COVID-19 disease. The average magnitude of the 20 HLA-A*01:01-restricted CD8 T cell responses that were detected in patients with acute COVID-19 disease was 2.97% (range: 0.017–19%) of total CD8 T cells. In contrast, the average magnitude of the 7 detected CD8 T cell responses restricted by other HLA alleles was 0.084% (range: 0.006–0.22%) of total CD8 T cells. This difference in magnitude was not only driven by CD8 T cell responses specific for the TTD epitope, as the average magnitude of the 12 other (not TTD-specific) HLA-A*01:01-restricted CD8 T cell responses was 2-fold higher compared to CD8 T cell responses restricted by other HLA alleles (0.182 vs. 0.082%). This observation suggests that patients with a specific (set of) HLA allele(s) may be able to raise a substantial CD8 T cell response toward SARS-CoV-2. However, given the limited size of patients in our and other studies, this hypothesis needs to be addressed in larger patient cohorts in the future.

Interestingly, in-depth characterization of SARS-CoV-2-specific CD8 T cells based on a stimulation assay, ex vivo flow cytometry, and transcriptome analysis revealed a regulated activation program that maintains CD8 T cell survival while limiting their effector function during acute COVID-19 disease. First, we observed a lack or strong reduction of cytokine production (IFNγ, TNF, IL-2, and IL-17) upon stimulation with the TTD epitope demonstrating a highly dysfunctional state of SARS-CoV-2-specific CD8 T cells during acute severe and critical COVID-19 disease. These data are in line with other studies demonstrating that the SARS-CoV-2-specific CD8 T cell response is functionally impaired in COVID-19 patients with severe disease in comparison to patients with acute mild and moderate disease. Second, the percentage of SARS-CoV-2-specific CD8 T cells expressing NKG2A was significantly higher compared to bulk CD8 T cells, and the magnitude of SARS-CoV-2-specific CD8 T cell responses correlated with increased expression of NKG2A and PD-1. The expression of the NKG2A receptor, in particular, may in part explain the inability of SARS-CoV-2-specific CD8 T cells to produce cytokines. Expression of NKG2A can limit immunopathology in influenza by inhibiting effector function.

This receptor may have a similar role in COVID-19 patients who frequently present with abnormally high levels of IL-6, which drives the expression of the NKG2A ligand HLA-E on T cells, B cells, and macrophages. Third, several studies showed that bulk T cells have impaired effector functions (i.e., cytokine production) and higher expression levels of inhibitory receptors compared to bulk T cells from healthy individuals, which is worsening with disease stage. In line with the functional data, our results demonstrate that identified SARS-CoV-2-specific CD8 T cell responses harbor a regulated activation program that limits their functional capacity and migration in patients with acute severe and critical COVID-19 disease. These data may be a cautionary suggestion that SARS-CoV-2-specific CD8 T cells—unlike CD4 T cells—are less likely to contribute to the immunopathology observed in patients with severe and critical COVID-19 but rather have a protective role, which is in line with findings from other studies. Our longitudinal data analysis demonstrates that SARS-CoV-2-specific CD8 T cell responses persist up to 5 months post hospital discharge (as expected at a lower magnitude compared to acute disease for most responses) and that these responses convert to functional memory T cells. These findings are in line with previous studies demonstrating functional memory SARS-CoV-2-specific CD8 T cell responses at the time of convalescence. Such memory T cell responses can persist up to 6 years in recovered SARS-CoV-1 patients, and it is of key importance to understand if...
this will also be the case for the SARS-CoV-2-specific T cell response.

To summarize, we have identified a SARS-CoV-2 epitope with immunodominant features, which is encoded by the ORF1ab and restricted by a common HLA allele. Interestingly, our data provide evidence for a dysfunctional state of SARS-CoV-2-specific CD8 T cell responses during acute infection and that these responses convert to functional memory T cells during convalescence.

**Methods**

**Patient material.** Peripheral blood samples of COVID-19 patients and HDs were collected in accordance with the Declaration of Helsinki after approval by the institutional review boards (Ethical Committee of Area Vasta Emilia Romagna, protocol number 177/2020, March 10, 2020, and subsequent amendments). Each participant gave informed consent. All COVID-19 patients were tested positive for SARS-CoV-2 using reverse transcriptase PCR from an upper respiratory tract (nose/throat) swab test in accredited laboratories. Patients were assigned in groups based on the COVID-19 disease status during sample collection according to the World Health Organization guidelines.34 Samples were collected as follows: during hospitalization/acute disease for COVID-19 patients with moderate, severe, or critical disease, 3 months post hospital discharge, and with a negative SARS-CoV-2 PCR test, for convalescent patients who recovered from severe or critical COVID-19 disease (Table 1 and Supplementary Table 1). Peripheral blood was collected in ethylenediaminetetraacetic acid tubes following subsequent isolation of PBMCs using Ficoll-Paque density centrifugation according to standard protocol. PBMCs were suspended in fetal bovine serum (FBS, Sigma, F7524) and stored in liquid nitrogen.

**HLA typing.** PBMCs isolated from COVID-19 patients were thawed and washed with RPMI 1640 (Life Technologies, 21875-034) supplemented with FBS (Sigma, F7524, 10% v/v), penicillin-streptomycin (Life Technologies, 15140-122, 1% v/v), and benzamide nuclease (Merck-Millipore, 70746-4, 2500 U/mL), resuspended and incubated at 37 °C for 30 min. For the HD samples, DNA was isolated directly from whole blood. PBMCs were counted and up to 1,000,000 cells were aliquoted for subsequent DNA isolation. DNA was isolated using the DNeasy Blood & Tissue Kit (Qiagen, 69506) according to the manufacturer’s protocol (Qiagen). HLA typing was done using next-generation sequencing according to the manufacturer’s protocol (GenDx).

**SARS-CoV-2 epitopes.** The entire SARS-CoV-2 proteome obtained from UniProt (Proteome ID: UP000464024) was considered for predicting SARS-CoV-2 epitopes using R version (3.6.3) and data.table (v.12.6). List of selected SARS-CoV-2 epitopes is provided in Supplementary Data 1. Potential SARS-CoV-2 epitopes were first selected based on a predicted NetChop-3.124 proteasomal processing score ≥0.5, followed by a final selection step based on the highest predicted binding affinity to major histocompatibility complex (MHC) according to NetMHCpan-4.0.25 In addition, SARS-CoV-2 epitopes that were predicted to be most immunogenic by the science community were included for analysis.45,57 Predicted 9-11mer epitopes from the following ORFs of the SARS-CoV-2 proteome were included in the final analysis: ORF 1ab, 3a, 6, 7a, 7b, 8b, 10, 14, envelope, membrane, nucleocapsid, and spike protein. Fifty SARS-CoV-2 epitope–HLA combinations were selected for each of the top 10 most prevalent HLA alleles in Italy, which resulted in a total of 438 unique epitopes that were synthesized by the Chemistry group, Leiden University Medical Centre. Predicted scores for proteasomal processing (NetChop-3.1) and binding affinity (NetMHCpan-4.0) of the 50 selected and 18 CD8 T cell-reco gnized epitopes for each allele are shown in Supplementary Fig. 6. The COVID-19 CoV Genomics tool (v.1.6.0) was used to predict NetMHCpan-4.0 scores for the selected SARS-CoV-2 peptides via UV-induced ligand exchange60,61. In brief, phyla complexes with UV-sensitive peptide were subjected to 254/366 nm UV light for 1 h at 4 °C in the presence of a rescue peptide. Fourteen fluorescence–strepavidin reagents (Supplementary Table 5) were conjugated to phyla monomers (100 μg/mL). For each phyla monomer, conjugation was performed with 2 fluorochrome–strepavidin reagents resulting in dual fluorescent color codes for up to 75 epitopes. Subsequently, milk (Sigma, LP0031, 1% w/v) was added to block and capture unspécific peptide-binding residues, and fluorescently labeled phyla monomers were incubated for 30 min (Sigma, B4501, 26.3 mM) and Na2O3 (0.02% w/v) in phosphate-buffered saline (PBS) was added to block residual binding sites.

**Flow cytometry assays.** For the phyla multimer assay and the phenotypic characterization, PBMC samples were thawed and washed with RPMI 1640 (Life Technologies, 21875-034) supplemented with human serum (Sigma, H3667, 10% v/v), penicillin–streptomycin (Life Technologies, 15140-122, 1% v/v), and benzamide nuclease (Merck-Millipore, 70746-4, 2500 U/mL); resuspended; and incubated at 37 °C for 30 min before staining. Antigen-specific CD8 T cells were stained for 5 minutes at 37 °C with phyla multimers (Supplementary Table 6) encoding unique dual fluorescent code combinations for up to 75 epitopes. Subsequently, cells were stained for 20 min on ice with antibodies (Supplementary Table 6). LIVE/DEAD Fixable IR Dead Cell Stain Kit (Invitrogen, L10199) staining was performed for 20 min on ice either during antibody staining (phyla multimer assay, 1/200) or for 10 min on ice after antibody staining (phenotypic characterization, 1/400). Individual staining was performed in the presence of Brilliant Staining Buffer Plus (BD, 562070) according to the manufacturer’s protocol (BD), and samples were washed twice before acquisition.

**Analysis of SARS-CoV-2-specific CD8 T cell responses.** The following gating strategy shown in Supplementary Fig. 1 was applied to identify CD8+ T cells: (i) selection of live (IRDye low-dim) single-cell lymphocytes [forward scatter (FSC)-/side scatter (SSC)-low, FSC/SSC-A1; (ii) selection of anti-CD8 and “dum” (anti-CD4, anti-CD14, anti-CD16, anti-CD19) negative cells. Antigen-specific CD8 T cell responses that were positive for only two and none of the other fluorescent phyla multimers were identified using Boolean gating. Full gating strategy is shown. A minimum of 1000 CD8 T cells was required for further analysis. Cut-off values for the definition of positive antigen-specific CD8 T cell responses were ≥2 events and ≥0.005% of total CD T cells. To avoid experimental bias, analysis was carried out without prior knowledge about clinical patient characteristics, and to reduce researcher bias caused by manual gating, only positive responses that were confirmed by three independent researchers were defined as real. Data were analyzed using either the BD FACSDiva (v.8.0.1) or the FlowJo (v.10.6.2/10.7) software. To monitor the reproducibility of the assay system, reference samples with up to 10 CD8 T cell responses present at varying frequencies were included in each analysis.

**Peptide stimulation assay.** PBMCs were thawed, washed, and incubated at 37 °C for 30 min or at 4 °C for 60 min (if cells were simultaneously used for single-cell RNA sequencing) in RPMI 1640 (Life Technologies, 21875-034) supplemented with human serum (Sigma, H3667, 10% v/v), penicillin–streptomycin (Life Technologies, 15140-122, 1% v/v), and benzamide nuclease (Merck-Millipore, 2500 U/mL). After washing, equal amounts of PBMCs (2 × 10^7 cells per condition) were cultured for 12 h at 37 °C in the presence of GolgiPlug (BD, 555029, 1/1000) and either the TTPDPEFLGKYR peptide (2 μg/mL) or equimolar amounts of DMSO (negative control). Phorbol 12-myristate 13-acetate (50 ng/mL) and ionomycin (1 μg/mL) were used as technical control. Cells were washed and stained for 20 min on ice with surface marker antibodies (Supplementary Table 6). After washing, cells were stained for 10 min on ice with the LIVE/DEAD Fixable IR Dead Cell Stain Kit (Invitrogen, L10119, 1/400). Subsequently, cells were washed, fixed, and permeabilized using the Foop3 Transcription Factor Staining Buffer Set (ebioscience, 00-5523-00) according to the manufacturer’s protocol. Intracellular cytokines were stained for 20 min on ice with antibodies (Supplementary Table 6). Cells were washed twice before acquisition. The following gating strategy shown in Supplementary Fig. 3a–c was applied to identify CD8+ T cells using FlowJo (v.10.6.2/10.7): (i) selection of live (IRDye low-dim) single-cell lymphocytes [FSC-W/H low, SSC-C/A, IFN-γ, IL-2, IL-10, IL-17, IL-4, tumor necrosis factor, IFN-γ, IL-2, IL-17, IL-4, tumor necrosis factor], gates were set based on the DMSO control. Cut-off values for the definition of cytokine-producing CD8 T cells responses stimulated with the cognate SARS-CoV-2 peptide were ≥2 events and ≥2-fold difference in the magnitude of TNF-α, IFN-γ, IL-2, and IL-17+ CD8+ cells compared to the DMSO control.

**Flow cytometer settings.** All samples were analyzed on the BD FACSymphony A5. The following 21-color instrument settings were used on the BD FACSymphony A5: blue laser (488 nm at 200 mW); FITC, 530/30BP, 503LP, B630, 670LP; PE, 575/20BP, BB515, 585/40BP, 750/20BP, BB610, 750/60BP, BB700, 750/60BP. Red laser (637 nm at 140 mW): APC, 670/30BP, APC-R700, 690LP, 630/45BP, IRDye and APC-H7, 750LP, 780/60BP. Violet laser (405 nm at 100 mW): BV421, 420LP, 431/
28PB; BV-480, 455LP, 470/20BP; BV-605, 565LP, 605/40BP; BV-650, 655LP, 661/11BP; BV-711, 711/85, 685; BV-750, 735LP; BV-786, 780/60BP; 570LP; UV laser (520/30BP; BV-805, 805LP, 805/25BP; BV-856, 550LP, 580/20BP; BV-865, 605LP, 615/20BP; BV-866, 630LP, 670/25BP; BV-8737, 735/45BP, 770LP; BV-8805, 770LP, 819/45BP). Yellow-green laser (561 nm at 150 mW); PE, 385/15BP; PE-Dazzle, 595LP, 600LP, 610/20BP; PerCP-eFl710, 710/50BP, 685LP; PE-Cy7, 750LP, 780/60BP. Appropriate compensation controls were included in each analysis.

**Statistical analysis.** Differences in the magnitude of identified CD8 T cell responses (stratified based on antigen source or recognized epitope) or the expression of inhibitory receptors between CD8+ and PHila+ cells were assessed using non-parametric Mann–Whitney U test. Phenotypic changes of SARS-CoV-2-specific CD8 T cells during acute COVID-19 disease and convalescence were assessed using Wilcoxon matched-pairs signed-rank test. Differences were considered significant if P < 0.05. Only significant P values are displayed. Data cut-off for all analyses was 4 January 2021. Statistical analysis was performed using Excel (v.16.36) and PRISM 8 (v.8.4.0).

**Single-cell RNA and TCR sequencing.** PBMCs were thawed, washed, and inoculated for 60 min on ice in cold RPMI 1640 (Life Technologies, 21875-034) supplemented with human serum (Sigma, H3667, 10% v/v), penicillin–streptomycin (Life Technologies, 15140-122, 1% v/v), and benzamide nuclease (Merck Millipore, 200746-5ML). Total SQSE-streptavidin oligo barcoded reagents (Supplementary Table 5) were conjugated to pHLA monomers (100 μg/mL) as described above and used to stain SARS-CoV-2-specific CD8 T cells for 30 min on ice (Supplementary Table 6). Subsequently, cells were stained for 20 min on ice with antibodies (Supplementary Table 6) and LIVE/DEAD Fixable IRDye Low-Dim Cell Stain Kit (Invitrogen, L10119, 1/200). Stained cells from individual patients were pooled and washed before sorting on the FACSaria Fusion. The following gating strategy shown in Supplementary Fig. 4a was applied to identify and sort CD8+ cells into PBS supplemented with bovine serum albumin (0.04% w/v) at 4 °C: (i) selection of live (IRDye low-dim) single-cell lymphocytes [FSC-W/H low, SSC-W/H low, FSC/SSC-All high], (ii) selection of anti-CD8+/CD3+ cells, (iii) exclusion of dead cells and ghosts, (iv) exclusion of monocytes and granulocytes. Flow cytometry data generated in this study are deposited in FlowRepository with the identifier FR-FCM-ZIKA. Source data are provided with this paper.

**Data availability.** Single-cell RNA-sequencing data generated in this study are deposited in the Gene Expression Omnibus (GEO) repository under the accession code GSE169503. Flow cytometry data generated in this study are deposited in FlowRepository with the identifier FR-FCM-ZIKA. Source data are provided with this paper.

Received: 28 August 2020; Accepted: 26 March 2021; Published online: 10 May 2021

**References.**

1. Worldometer. [https://www.worldometers.info/coronavirus](https://www.worldometers.info/coronavirus) (2020).

2. Shomuradova, A. S. et al. SARS-CoV-2 epitopes are recognized by a diverse and realistic repertoire of human T cell receptors. *Immunity*, 53, 1245–1257 (2020).

3. Wajnberg, A. et al. Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science*, 370, 1227–1230 (2020).

4. Sehine, T. et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *Cell*, 183, 158–168 (2020).

5. Lau, E. H. Y. et al. Neutralizing antibody titres in SARS-CoV-2 infections. *Nat. Commun.*, 12, 63 (2021).

6. Ni, L. et al. Detection of SARS-CoV-2-specific humoral and cellular immunity in COVID-19 convalescent individuals. *Immunity*, 52, 971–977 (2020).

7. Ren, L. et al. The kinetics of humoral response and its relationship with the disease severity in COVID-19. *Commun. Biol.*, 3, 780 (2020).

8. Grifoni, A. et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell*, 181, 1–13 (2020).

9. Weiskopf, D. et al. Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci. Immunol.*, 5, eabd2071 (2020).

10. 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).

11. Thiemer, C. J. et al. Robust T cell response toward spike, membrane, and nucleocapsid SARS-CoV-2 proteins is not associated with recovery in critical COVID-19 patients. *Cell Rep. Med.*, 1, 100092 (2020).

12. Schulen, I. et al. Characterization of pre-existing and induced SARS-CoV-2-specific CD8+ T cells. *Nat. Med.*, 27, 78–85 (2021).

13. Oja, A. E. et al. Divergent SARS-CoV-2-specific T- and B-cell responses in severe but not mild COVID-19 patients. *Eur. J. Immunol.*, 50, 1998–2012 (2020).

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

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

16. Neidleman, J. et al. SARS-CoV-2-specific T cells exhibit phenotypic features of helper function, lack of terminal differentiation, and high proliferation potential. *Cell Rep. Med.*, 1, 100081 (2020).

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

18. Ferretti, A. P. et al. Unbiased screens show CD8+ T cells from COVID-19 patients recognize shared epitopes in SARS-CoV-2 that largely reside outside the spike protein. *Immunity*, 53, 1–13 (2020).

19. Yewdell, J. W. Confronting complexity: real-world immunodominance in antiviral CD8+ T-cell responses. *Immunity*, 25, 533–543 (2006).
20. Su, S., Du, L. & Jiang, S. Learning from the past: development of safe and effective COVID-19 vaccines. Nat. Rev. Microbiol. 368, 1–9 (2020).
21. Ding, B. et al. Reduction and functional exhaustion of T cells in patients with coronavirus disease 2019 (COVID-19). Front. Immunol. 11, 827 (2020).
22. Zheng, M. et al. Functional exhaustion of antiviral lymphocytes in COVID-19 patients. Cell. Mol. Immunol. 17, 533–535 (2020).
23. Mathew, D. et al. Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. Science 369, eabs511 (2020).
24. Nielsen, M., Lundegaard, C., Lund, O. & Kesmir, C. The role of the proteasome in generating cyotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. Immunogenetics 57, 33–41 (2005).
25. Jurtz, V. et al. NetMHCpan-4.0: improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. J. Immunol. 199, 3360–3368 (2017).
26. Kivistö, P. et al. Anti-CTLA-4 therapy broadens the melanoma-reactive CD8+ T cell response. Sci. Transl. Med. 254, 254ra126 (2014).
27. Kivistö, P. et al. TIL therapy broadens the tumor-reactive CD8(+) T cell compartment in melanoma patients. Oncoimmunology 1, 409–418 (2012).
28. Shu, Y. & McCauley, J. GISAID: Global Initiative on Sharing All Influenza Data - from vision to reality. Eurosurveillance 22, 30497 (2014).
29. Frankländ, S., De Boer, R. J., Lund, O., Nielsen, M. & Kesmir, C. Amino acid similarity accounts for T cell cross-reactivity and for ‘holes’ in the T cell receptor repertoire. Proc. Natl Acad. Sci. USA 108, 1831 (2008).
30. Rubio-Godoy, V. et al. Positional scanning-synthetic peptide library-based analysis of self- and pathogen-derived peptide cross-reactivity with tumour-reactive Melan-A-specific CTL. J. Immunol. 169, 5696–5707 (2002).
31. Riley, T. P. et al. T cell receptor cross-reactivity expanded by dramatic peptide-MHC adaptability. Nat. Chem. Biol. 14, 934–942 (2018).
32. De Boer, R. J., Homann, D. & Perelson, A. S. Different dynamics of CD4+ and CD8+ T cell responses during and after acute lymphocytic choriomeningitis virus infection. J. Immunol. 171, 3928–3935 (2003).
33. Kotturi, M. F. et al. Naive precursor frequencies and MHC binding rather than MHC frequencies in 82 Chinese individuals with coronavirus disease-2019 (COVID-19). J. Immunol. Res. 6, 883–894 (2006).
34. Amanat, F. & Krammer, F. SARS-CoV-2 vaccines: status report. Nat. Clin. Pract. Immunol. 6, 880 (2010).
35. Mathew, D. et al. Deep immune pro...
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-22811-y.

Correspondence and requests for materials should be addressed to P.K.

Peer review information Nature Communications thanks Antonio Bertoletti and the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.