# Identification and characterization of an immunodominant SARS-CoV-2-specific CD8 T cell response

**Anastasia Gangaev**
Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, The Netherlands
**Steven L. C. Ketelaars**
Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, The Netherlands
**Olga I Isaeva**
Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, The Netherlands
**Sanne Patiwael**
Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, The Netherlands
**Anna Dopler**
Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, The Netherlands
**Kelly Hoefakker**
Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, The Netherlands
**Sara De Biasi**
Department of Medical and Surgical Sciences for Children and Adults, University of Modena and Reggio Emilia School of Medicine, Italy
**Lara Gibellini**
Department of Medical and Surgical Sciences for Children and Adults, University of Modena and Reggio Emilia School of Medicine, Italy
**Cristina Mussini**
Department of Medical and Surgical Sciences for Children and Adults, University of Modena and Reggio Emilia School of Medicine, Italy
**Giovanni Guaraldi**
Department of Medical and Surgical Sciences for Children and Adults, University of Modena and Reggio Emilia School of Medicine, Italy
**Massimo Girardis**
Department of Medical and Surgical Sciences for Children and Adults, University of Modena and Reggio Emilia School of Medicine, Italy
**Cami M. P. Talavera Ormeno**
Department of Cell and Chemical Biology, Leiden University Medical Center, The Netherlands
**Paul J. M. Hekking**
Department of Cell and Chemical Biology, Leiden University Medical Center, The Netherlands

**Neubury M. Lardy**
Department of Immunogenetics, Sanquin Diagnostics B.V., The Netherlands
**Mireille Toebes**
Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, The Netherlands
**Robert Balderas**
Department of Biological Sciences, BD Bioscience, USA
**Ton N. Schumacher**
Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, The Netherlands
**Huib Ovaa**
Department of Cell and Chemical Biology, Leiden University Medical Center, The Netherlands
**Andrea Cossarizza**
Department of Medical and Surgical Sciences for Children and Adults, University of Modena and Reggio Emilia School of Medicine, Italy
**Pia Kvistborg** ( p.kvistborg@nki.nl )
Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, The Netherlands

---

**Research Article**

**Keywords:** SARS-CoV-2, COVID-19, CD8 T cell reactivity, ORF1ab polyprotein 1ab

**DOI:** https://doi.org/10.21203/rs.3.rs-33197/v2



---

# Abstract

Global efforts are ongoing to develop vaccines against SARS-CoV-2 causing COVID-19. While there is accumulating information on antibody responses against SARS-CoV-2, less is known about CD8 T-cell recognized SARS-CoV-2 epitopes and the functional state of SARS- CoV-2-specific CD8 T cells. To address these issues, we analyzed samples from 18 COVID- 19 patients for CD8 T-cell recognition of 500 peptide HLA class I complexes, restricted by 10 common HLA alleles. Several epitopes derived from ORF1ab were identified, including an immunodominant epitope restricted by HLA-A*01:01. The immunodominance was further supported by high TCR diversity within the CD8 T cells specific for this epitope. Noteworthy, the ORF1ab is not included in the majority of vaccine candidates in development, which may influence their clinical activity. In-depth characterization of identified SARS-CoV-2-specific CD8 T cell responses revealed a lack of cytokine production and a gene expression profile inhibiting T cell re-activation and migration while sustaining cell survival.


# Introduction

The COVID-19 pandemic caused by SARS-CoV-2, is an ongoing global emergency. The first cases of COVID-19 were reported in December 2019, and as of August 24$^{th}$ 2020 there are more than 23.700.000 confirmed cases and 800.000 deaths[1] worldwide. Due to the measures necessary to contain the rapid spread of the infection, this pandemic is having tremendous health and socioeconomic consequences, and there is an urgent need for vaccines.

Accumulating information on the antibody response against SARS-CoV-2 demonstrates that structural proteins such as the spike and nucleoprotein serve as potent antibody targets in a substantial fraction of COVID-19 patients[2-6]. However, much less is currently known regarding the SARS-CoV-2-specific T cell response, in particular which parts of the virus are targeted by T cells. Both CD4 and CD8 T cells can recognize SARS-CoV-2 derived antigens[4,6-13]. Most studies have focused on limited parts of the viral proteome[6,8,12], and in particular the spike protein[4,10,13], which was found to be a very important antigen source for virus specific T cell responses in SARS-CoV-1[14] and MERS[15]. However, accumulating data provide evidence for a broader repertoire of the SARS-CoV-2 derived T cell antigens[6-9,11,12].

The vast majority of studies demonstrating T cell reactivity against SARS-CoV-2 are based on large peptide pools[7-9,12,13]. Only few studies have identified exact T cell epitopes and their restriction elements, and these have mainly been focused on HLA-A*02:01 and a limited set of selected epitopes from the spike protein[2,6,10]. Therefore, broader efforts are needed to provide fundamental insights into which parts of the SARS-CoV-2 proteome induce potent CD8 T cell responses. Such knowledge can be utilized to generate tools which can facilitate high throughput analyses to establish whether infectious or recovered patients harbor SARS- CoV-2 specific T cell responses. This type of analysis is essential to complement the antibody tests to more accurately estimate the proportion of people that have been infected with SARS- CoV-2. Finally, generating a full overview of the T cell antigen landscape from the entire viral

proteome will inform how well the antigens targeted by CD8 T cells reflect the antigen composition of current vaccine candidates, and whether these vaccines include the immunodominant epitopes that are central for successful vaccines[16].

Studies assessing the functionality and state of immune cells are mainly focused on broad characterization of the composition of immune cell linages including bulk CD8 T cells. For example, it has been demonstrated that bulk T cells from COVID-19 patients have impaired effector functions (i.e. cytokine production) and express higher levels of inhibitory receptors compared to bulk T cells from healthy individuals, which is worsening with disease stage[17-19]. However, in-depth characterization of the SARS-CoV-2 specific CD8 T cell response that can provide essential information regarding the role of CD8 T cells in the host defense against SARS-CoV-2 and during disease development is currently lacking.

We probed for CD8 T cell recognition towards 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 were directed towards epitopes derived from the Open Reading Frame 1ab polyprotein (ORF1ab) including an immunodominant epitope restricted by HLA-A*01:01. In-depth characterization of identified SARS-CoV-2 specific CD8 T cell responses revealed high expression of NKG2A, lack of effector function, and a gene expression profile of a tightly regulated T cell response.

# 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 HLA alleles that are common in the Italian population and could be covered with our peptide HLA (pHLA) multimer technology[20]. This collection of HLA-A (HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, 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[21]. For each HLA allele 50 SARS-CoV-2 epitopes that were derived from the entire SARS-CoV-2 proteome were selected (Fig. 1a). Epitope selection was based on predicted binding affinity (NetMHCpan- 4.0)[22] and likelihood of successful proteasomal processing (NetChop-3.1)[23]. 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 (Table S1). Overall, the contributions of the ORF-of-origin to selected epitopes roughly reflected the contribution of the individual ORFs to the proteome (Fig. 1b), demonstrating no bias in the representation of the SARS-CoV-2 proteins for our epitope selection.

## The SARS-CoV-2 specific CD8 T cell response

To investigate which of the included epitopes from SARS-CoV-2 were recognized by CD8 T cells, peripheral blood mononuclear cell (PBMC) samples from 22 hospitalized COVID-19 patients with confirmed SARS-CoV-2 infection were included (Table 1). At the time of sample collection patients were stratified in three groups based on the disease state: severe disease no longer needing oxygen support (n=3), severe disease requiring non-invasive ventilation (n=10), and critical disease requiring intubation and mechanical ventilation at the intensive care unit (n

=9). To probe for CD8 T cell recognition of selected SARS-CoV-2 epitopes, we made use of our in-house developed technology based on multiplexing of pHLA multimers conjugated to fluorescent dyes, which has previously been used to identify tumor-specific CD8 T cells responses in cancer patients[20,24]. At least 1 of the 10 selected HLA alleles was covered for 20 of the 22 patients (Fig.1c and Table 1). Notably, SARS-CoV-2-specific CD8 T cell reactivity was evaluated directly *ex vivo*.

The analysis resulted in the detection of 16 SARS-CoV-2 specific CD8 T cell responses (Fig. 1d and e; representative gating strategy is shown in Fig. S1). SARS-CoV-2-specific CD8 T cell responses were detected in 50% of the patients (after sample acquisition, two patients were excluded from the data analysis due to low number of $CD8^+$ cells: <1000). The average magnitude of the detected responses was 2.4% of total $CD8^+$ cells, with a range of 0.006% to 18.4% of total $CD8^+$ cells (Table S2). Of the 9 CD8 T cell recognized epitopes identified in this study, 5 were unique for SARS-CoV-2 (DTDFVNEFY, PTDNYITTY, QYIKWPWYI, TTDPSFLGRY, YLQPRTFLL) and 4 were shared between SARS-CoV-2 and SARS-CoV-1 (CTDDNALAYY, KTFPPTEPK, RLNEVAKNL, TVATSRTLSY) (Table S1).

RLNEVAKNL was previously demonstrated to be a CD8 T cell target in SARS-CoV-1[25]. The 9 identified T cell epitopes were derived from the ORF1ab (n=4), spike protein (n=3), membrane protein (n=1), and nucleoprotein (n=1). Strikingly, the CD8 T cell responses specific for the epitopes derived from ORF1ab were of significantly higher magnitude compared to the CD8 T cell responses towards the spike, membrane and nucleoprotein combined (Fig. 1f, $p = 0.0027$). The high number of detected CD8 T cell responses directed towards the ORF1ab-derived epitopes included in our screen was not due to increased relative expression of ORF1ab compared to other viral proteins. Multiple studies have demonstrated that the ORF1ab is the least translated ORF of the SARS-CoV-2 genome[26-30].

In addition, we analyzed PBMC samples collected prior to October 2019 from 4 healthy donors (HD) covering 7 of the 10 included HLA alleles (i.e. HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-B*07:02, HLA-B*08:01 and HLA-B*15:01). One CD8

T cell response of low magnitude (0.008% of total $CD8^+$ cells) restricted by HLA-B*15:01 was identified in HD1 (Fig. 1e). Of note, the number of epitopes shared between the corona viruses causing the ‘common cold’ and SARS-CoV-2 epitopes included in our analysis (Table S1) was substantially lower compared to previous studies[4,7-10].

## T cell recognized SARS-CoV-2 epitopes are highly conserved across virus isolates

Next, we examined whether the SARS-CoV-2 epitopes recognized by CD8 T cells were spanning the positions of the SARS-CoV-2 genome that contain a high level of single nucleotide polymorphisms (SNPs), the so-called mutational 'hotspots'. Such information is relevant for the broader utilization of the obtained information. For this purpose, we used the SARS-CoV-2 alignment screen tool (version as per 14-05-2020) containing sequencing information of 7667 SARS-CoV-2 isolates[31,32]. None of the 9 CD8 T cell recognized epitopes identified in our analysis overlapped with the hotspots (Fig. S2 and Table S3). Furthermore, between 7667 and 6996 virus isolates contained information regarding the potential consequence of the SNPs on the amino acid level in the regions encoding for the 9 CD8 T cell recognized epitopes. The median fraction of these isolates encoding all 9 CD8 T cell epitopes was 99.98% with a range of 97.13% to 100%. These analyses suggest that the origin of identified CD8 T cell epitopes is highly conserved.

## The ORF1ab encodes an immunodominant epitope restricted by HLA-A*01:01

The vast majority of identified SARS-CoV-2-specific CD8 T cell responses (11 of 16) were restricted by the HLA*01:01 allele (Fig. 2a). Notably, these results were not biased by higher predicted binding affinity for epitopes restricted to HLA-A*01:01 in comparison to other HLA alleles (Fig. S3). Strikingly, TTDPSFGLGRY-specific (TTD-specific) CD8 T cell responses were detected in all 5 HLA-A*01:01-positive patients, and these responses were of substantially higher magnitude compared to the magnitude of all other detected CD8 T cell responses (Fig. 2b, $p = 0.0002$). Importantly, for 3 of the 5 HLA-A*01:01 positive patients, additional responses were identified, and the magnitude of these responses was at least 20-fold lower compared to the response towards the TTD epitope within each patient (Fig. 1c).

The immunodominance hierarchy is partially dictated by the naïve precursor frequency[33,34]. To investigate whether more than one T cell clone gave rise to the immunodominant SARS-CoV-2-specific CD8 T cell response, T cell receptor (TCR) data on 32 TTD-specific CD8 T cells isolated from 4 patients was obtained. Interestingly, the results showed a high level of TCR diversity with 28 of 32 cells harbouring TCR beta chains with unique CDR3 loops (Fig. 2d), albeit, with similarities for both TCR beta and alpha within each patient and across patients (Fig. S4 a and b), demonstrating a large clonal TCR heterogeneity for the TTD-specific CD8 T cell response. A second observation based on these data was the enrichment of TCRs with the TCR beta chain V27 (TRBV27) segment compared to the TCR repertoire of bulk CD8 T cells (Fig. 2e) suggesting that TCRs specific for the TTDPSFLGRY epitope may be more likely to be rearranged during T cell development. Taken together, these data strongly indicate that TTDPSFGLGRY is the immunodominant epitope for the subgroup of patients positive for HLA-A*01:01 which is present in approximately 30% of the European population (depending on the composition of the entire set of HLA alleles)[21].

## Dysfunctional state of SARS-CoV-2 specific CD8 T cell responses

To characterize the identified SARS-CoV-2 specific CD8 T cell response, we first assessed the functional capacity based on activation markers (CD137, HLA-DR) and cytokine production (IFNγ, TNFα, IL-2, IL-17) of SARS-CoV-2-specific CD8 T cells upon peptide stimulation. For this analysis material from 6 patients with a total of 9 of the 16 SARS-CoV-2-specific CD8 T cell responses (including 4 TTD-specific CD8 T cell responses) was used. Surprisingly, no cytokine production was detected after 12 hours of peptide stimulation in patients with ongoing severe and critical disease (representative example shown in Fig. S5). Of note, CD8 T cells from all patients were able to produce IFNγ and TNFα, and to a lesser extend IL-2 in response to phorbol 12-myristate 13-acetate and Ionomycin (Fig. 3a and Fig. S5). Of the 9 SARS-CoV-2-specific CD8 T cell responses that were identified using the pHLA multimer assay, only 3 TTD-specific CD8 T cell responses were detectable based on upregulation of HLA-DR, albeit at a lower magnitude (~6-fold on average) compared to what was detected based on the pHLA multimer assay (Fig. 3b). These results demonstrate that detection of these responses with a functional readout failed or vastly underestimated the magnitude of the SARS-CoV-2 specific CD8 T cell responses. To further examine the state of SARS-CoV-2-specific CD8 T cells, we examined expression levels of 4 inhibitory receptors (NKG2A, PD-1, TIM-3, and 2B4) on all 16 identified responses directly *ex vivo.* The fraction of NKG2A positive cells was significantly increased for SARS-CoV-2-specific CD8 T cells compared to bulk CD8 T cells (Fig. 2c and d). In addition, while we did not observe a significant difference for PD-1 between SARS-CoV-2-specific CD8 T cells and bulk CD8 T cells (Fig. 2c and d), we found a significant correlation between PD-1 expression levels (MFI) on the SARS-CoV-2 specific CD8 T cells and the magnitude of the response (Fig. 2e). Together, these data suggest that the identified SARS-CoV-2 specific CD8 T cell responses in patients with ongoing severe and critical disease were highly dysfunctional.

## Highly regulated activation of the SARS-CoV-2 specific CD8 T cell response

Finally, to perform an unbiased characterization of TTD-specific CD8 T cells, we obtained single-cell RNA sequencing data on CD8 T cells from 5 patients using the 10x Genomics Chromium system. Louvain clustering based on highly variable genes resulted in 6 clusters (Fig. 4a): naïve cells expressing IL7R and TCF7 (C1), proliferating cells expressing high levels of Ki-67 (C2) and activated CD8 T cells expressing GZMB (C3-6). Overall, differences between C3, C4, C5 and C6 were minor, however, differential gene expression analysis revealed C3 as cytotoxic cells expressing GNLY and PRF1 and C5 as antigen-induced differentiating cells expressing CELF2, ZEB2, and NLRC5 (Table S4). Cells from all 5 patients were represented in all clusters, though, the vast majority of cells were from one patient COVID-096 (Fig. 4a).

Next, we focused our analysis on data obtained from a total of 48 oligo-labelled TTD- specific CD8 T cells from 4 patients. TTD-specific CD8 T cells resembled a heterogenous population with the vast majority of cells present in cluster C2, C4 and C6 (Fig. 4a). To further characterize the TTD-specific CD8 T cell response in relation to bulk CD8 T cells, we performed differential gene expression analysis between TTD-

specific CD8 T cells and CD8 T cells from the naïve cluster (C1). The analysis resulted in the detection of 112 upregulated and 68 downregulated genes in TTD-specific CD8 T cells (Table S4). Gene ontology analysis of upregulated genes revealed the presence of several immune response-related processes including regulation of T cell activation, adhesion and proliferation (Fig. 4b). Of note, gene sets previously described to be associated with T cell exhaustion were not differentially expressed (data not shown). Next, we assessed the difference in gene expression in TTD- specific CD8 T cells compared to the activated bulk CD8 cells (C2-C6). A total of 26 and 1508 genes were found to be up- or down-regulated in TTD-specific CD8 T cells, respectively (Table S4). Interestingly, the TTD-specific CD8 T cells showed downregulation of genes involved in T cell activation (CD84, GITR and AHCY) and migration to mucosal and epithelial tissues (CCL28 and CCL3) as well as upregulation of genes that promote T cell survival (CD27), but also inhibit T cell activation (PAG1) and proliferation (DUSP2) (Fig. 4c). Together, these results suggest a specific program of T cell activation in TTD-specific CD8 T cells that are maintained but unable to produce cytokines or migrate to site of infection.

# 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 9 SARS-CoV-2 derived CD8 T cell epitopes. The origin of these epitopes was diverse including the ORF1ab, spike protein, membrane protein and nucleoprotein. Four epitopes were derived from the ORF1ab including the novel immunodominant epitope TTDPSFLGRY that raised CD8 T cell responses in all five patients positive for HLA-A*01:01. The magnitude of the immunodominant CD8 T cell response reached up to 18% of total CD8 T cells and was substantially higher compared to other SARS-CoV-2-specific CD8 T cell responses across and within patients. TCR sequencing data of the immunodominant CD8 T cell response revealed a high level of clonal diversity together with an enrichment for the TCR beta chain V27 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[35,36]. These data provide the first TCR sequences specific for an immunodominant SARS-CoV-2 epitope. Such information paves the way for T cell therapies in COVID-19 patients unable to mount protective T cell immunity in response to e.g. vaccination.

Our analysis was limited by the selection of HLA alleles and number of epitopes, as well as imperfect *in silico* epitope predictions. Nevertheless, accumulating evidence from other groups[7] and our data demonstrate that epitopes derived from other parts of the SARS-CoV-2 proteome than the spike protein alone can induce high-magnitude CD8 T cell responses. Furthermore, 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 if CD8 T cell responses raised against the spike protein alone are sufficient to induce protective T cell immunity. Notably, the ORF1ab is not included in the

majority of the vaccine candidates currently under development[37,38], which may potentially influence their clinical activity.

Based on our data it appears that there may be a bias for HLA alleles capable of presenting epitopes which can induce strong CD8 T cell responses. For 6 of the 11 detected CD8 T cell responses restricted by HLA-A*01:01, the magnitude of the response was >0.5% of total CD8 T cells. In contrast, the magnitude of CD8 T cell responses restricted by all other HLA alleles (5 responses in total) was between 0.006% and 0.085% of total CD8 T cells. This observation suggests that patients with a specific (set of) HLA alleles may be able to raise substantial CD8 T cell response towards SARS-CoV-2. Such a scenario is well described for HIV patients positive for HLA-B*57:01 who are more likely to be long term non-progressors[39]. If this is also the case for COVID-19 patients, one would expect that certain subgroups of patients with specific HLA-alleles harbor SARS-CoV-2 specific CD8 T cell response of higher magnitude, and therefore potentially a more favorable course of disease. However, given the limited size of our dataset, this hypothesis needs to be addressed in larger patient cohorts in the future. Understanding if there is a relationship between HLA haplotypes and severity of disease may also facilitate the identification of additional immunodominant epitopes for other HLA alleles.

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. First, there was a complete lack of cytokine production (IFNγ, TNFα, IL-2 and IL-17) upon peptide stimulation demonstrating a highly dysfunctional state of SARS-CoV-2-specific CD8 T cells. Similarly, CD137 was absent while HLA-DR expression was increased only upon stimulation with a subset of the epitopes tested. Second, the fraction of SARS-CoV-2-specific CD8 T cells expressing NKG2A was significantly higher compared to bulk CD8 T cells. The high expression of the NKG2A receptor may partially explain the inability of SARS-CoV-2 specific CD8 T cells to produce cytokines upon peptide stimulation. Expression of NKG2A can limit immunopathology in influenza[40] and may have a similar role in COVID-19 patients that frequently present with abnormally high levels of IL-6[17,41,42], which partially drives the expression of the NKG2A ligand HLA-E on T cells, B cells, and macrophages[43-45]. Third, gene expression analysis of SARS-CoV-2-specific CD8 T cells revealed a specific program of T cell activation. Together, our findings 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 ongoing severe and critical COVID-19 disease. These data may be a cautious indication that SARS-CoV-2 specific CD8 T cells – unlike CD4 T cells[46] – are less likely to contribute to the immunopathology observed in severely and critically ill COVID-19 patients[47]. However, the current data set is limited in size and can therefore be only hypothesis generating.

To summarize, we have identified a novel immunodominant epitope from SARS-CoV- 2 which is encoded by ORF1ab and restricted by a common HLA allele. Interestingly, our data provide evidence for a tightly regulated T cell activation gene expression program that inhibits the effector functions and sustains the survival of SARS-CoV-2-specific CD8 T cells in patients with active severe and critical COVID-19 disease.

# Methods

## Blood collection and PBMC isolation

The samples from both COVID-19 patients and healthy donors were collected in accordance with the Declaration of Helsinki after approval by the institutional review boards. Each participant signed informed consent. All COVID-19 patients were tested positive for SARS- CoV-2 using reverse transcriptase chain reaction (RT-PCR) from an upper respiratory tract (nose/throat) swab test in accredited laboratories. Peripheral blood was collected in ethylenediaminetetraacetic acid (EDTA) tubes following subsequent isolation of PBMCs using Ficoll-Paque density centrifugation according to standard protocol. PBMCs were suspended in fetal bovine serum (FCS) with 10% dimethyl sulfoxide (DMSO) and stored in liquid nitrogen.

## HLA typing of study participants

PBMCs isolated from COVID-19 patients were thawed and washed with RPMI 1640 supplemented with 10% FCS, 1% Penicillin-Streptomycin solution and Benzonase nuclease (Merck-Millipore, 2500 U/mL), resuspended and incubated at 37C for 30 minutes. For the healthy donor 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, cat. #69506) according to manufacturer's protocol. HLA typing was done using next-generation sequencing according to the manufacturer's protocol (GenDx).

## SARS-CoV-2 epitope selection and peptide synthesis

Fifty SARS-CoV-2 epitope-HLA combinations were selected for each of the top 10 most prevalent HLA alleles in Italy. The selection was primarily based on SARS-CoV-2 epitopes that had the highest predicted binding affinity to the MHC according to NetMHCpan-4.0[22], as well as receiving a prediction score higher than 0.5 using NetChop-3.1[23]. The SARS-CoV-2 proteome was obtained from UniProt (Proteome ID: UP000464024). Predicted 9-11 mer epitopes from the following open reading frames (ORFs) of the SARS-CoV-2 proteome were included in the analysis: ORF 1ab, 3a, 6, 7a, 7b, 8, 9b, 10, 14, envelope (E), membrane (M), nucleoprotein (N) and spike (S) protein. In addition, SARS-CoV-2 epitopes that were predicted to be most immunogenic by the science community were included for analysis (Table S1)[48-51].

A total of 438 unique peptides were synthesized by the Chemical Biology group, Leiden University Medical Centre.

## Generation of UV-cleavable pHLA monomers

The UV-cleavable peptides were synthesized in-house as described previously[52]. Recombinant HLA-A*01:01, A*02:01, A*03:01, A*11:01, A*24:02, B*07:02, B*08:01, B*15:01, B*18:01 and B*51:01 heavy chains and human beta-2 microglobulin (B2M) were produced in *Escherichia coli* and isolated from resulting inclusion bodies[53]. MHC class I refolding reactions and purification by gel filtration HPLC were performed, and HLA-A and B heavy chains and B2M were refolded in the presence of UV-cleavable peptides (Table S5) following subsequent biotinylation as described previously[54].

## Generation of fluorescent pHLA multimers

MHC complexes were loaded with the selected SARS-CoV-2 peptides via UV-induced ligand exchange[52,55]. In brief, pHLA complexes with UV-sensitive peptide were subjected to 254/366 nM UV light for 1h at 4˚C in the presence of a rescue peptide. The following amounts of 14 different fluorescent streptavidin conjugates were added to 10 ml of pHLA monomer (100 mg/ml): 1 ml of SA-BB790 (BD, custom), 1 ml of SA-BB630 (BD, custom), 1 ml of SA-APC- R700 (BD, 565144), 0.6 ml of SA-APC (Invitrogen, S868), 1 ml of SA-BV750 (BD, custom), 2 ml of SA-BV650 (BD, 563855), 2 ml of SA-BV605 (BD, 563260), 2 ml of SA-BV480 (BD, 564876), 2 ml of SA-BV421 (BD, 563259), 1 ml of SA-BUV615 (BD, 613013), 1.5 ml of SA- BUV563 (BD, 565765), 2 ml of SA-BUV395 (BD, 564176), 1.25 ml of SA-BV711 (BD, 563262) and 0.9 ml of SA-PE (Invitrogen, S866). For each pHLA monomer, conjugation was performed with two of these fluorochromes resulting in up to 75 dual fluorescent color codes. Subsequently, milk (1% w/v, Sigma) was added to block and capture unspecific peptide binding residues, and fluorescently labelled pHLA multimers were incubated for 30 min on ice. Finally, D-biotin (26.3 mM, Sigma) in PBS and NaN3 (0.02% w/v) was added to block residual binding sites.

## Combinatorial encoding of pHLA multimers and surface marker staining

PBMCs were thawed and washed with RPMI 1640 supplemented with 10% human serum, 1% Penicillin-Streptomycin solution and Benzonase nuclease (Merck-Millipore, 2500 U/mL), resuspended and incubated at 37C for 30 minutes. The following amounts of fluorescently labelled pHLA multimers were used to stain CD8 T cells: 1 ml of SA-BB790-pHLA, SA-BB630-pHLA, SA-APC-R700-pHLA, SA-BV750-pHLA, SA-BV650-pHLA, SA-BV605- pHLA, SA-BV480-pHLA, SA-BV421-pHLA, SA-BUV615-pHLA, SA-BUV563-pHLA, SA- BUV395-pHLA, SA-BV711-pHLA, SA-PE-pHLA and 2 ml of SA-APC-pHLA. The cells were stained in Brilliant Staining Buffer Plus (BD, 563794) according to manufacturer's protocol. Final staining volume was 100 up to 194 ml depending on the amount of fluorescent pHLA multimers for each individual sample. Cells were incubated for 15 min at 37 °C. Subsequently cells were stained with 2 ml of a(nti)CD8-BUV805 (BD, 564912), 1 ml of aCD4- APC-H7 (BD, 641398), aCD14-APC-H7 (BD, 560180), aCD16-APC-H7 (BD, 560195), aCD19-APC-H7 (BD, 560252), a2B4-FITC (BD, 550815), aTim-3-BV786 (BD, 742857), aPD1-BUV737 (BD, 612791), aNKG2A-PE-Cy7 (Beckman, B10246) and 0.5 ml of LIVE/DEAD

Fixable IR Dead Cell Stain Kit (Invitrogen, L10119) and incubated on ice for 20 min. Samples were analyzed on the BD FACSymphony A5.

## Identification of antigen-specific CD8 T cell responses

Analysis of antigen-specific CD8 T cell responses was carried out without prior knowledge about clinical patient characteristics to avoid experimental bias. The following gating strategy was applied to identify $CD8^+$ T cells: (i) selection of live (IRDye low-dim) single-cell lymphocytes [forward scatter (FSC)-W/H low, side scatter (SSC)-W/H low, FSC/SSC-A], (ii) selection of anti-$CD8^+$ and 'dump' (anti-CD4, anti-CD14, anti-CD16, anti-CD19) negative cells. Antigen-specific CD8 T cell responses that were positive for two none of the other pHLA multimer channels were identified using Boolean gating. The full gating strategy used on the BD FACSymphony A5 is shown in Fig. S1. Cut-off values for the definition of positive responses were ≥ 0.005% of total CD8 T cells and ≥ 5 events. A minimum of 1,000 CD8 T cells were acquired per sample. To reduce researcher-bias caused by manual gating, only positive responses that were confirmed by three independent people were defined as real. Data was analysed using either the BD FACSDiva v.8.0.1 or the FlowJo 10.6.2 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 of SARS-CoV-2-specific CD8 T cells

PBMCs were thawed and washed with RPMI 1640 supplemented with 10% human serum, 1% Penicillin-Streptomycin solution and Benzonase nuclease (Merck-Millipore, 2500 U/mL), resuspended and incubated at 4C for 60 minutes. Cells were washed and cultured at 37C for 12 hours in 96-wells U-bottom plates at $1x10^5$ PBMCs per well in the presence of GolgiPlug (1:1000, BD, 555029), and either equimolar amounts of DMSO alone (negative control), or phorbol 12-myristate 13-acetate (50 ng/ml) and Ionomycin (1 µg/ml) (positive control) or SARS-CoV-2 peptides (2 µg/ml of each peptide). Cells were washed, resuspended in 50 ml FACS buffer and stained for 20 min on ice with 1 ml of aCD8-BUV805 (BD, 564912), 0.5 ml of aCD4-APC-H7 (BD, 641398), aCD14-APC-H7 (BD, 560180), aCD16-APC-H7 (BD, 560195), aCD19-APC-H7 (BD, 560252), aNKG2A-PE-Cy7 (Beckman, B10246), 0.125 ml of aHLA-DR-BUV661 (BD, 565074) and 1 ml of CD137-PE/Dazzle-594 (BioLegend, 309825).

After incubation, cells were washed with PBS and resuspended in 50 ml PBS and 0.125 ml LIVE/DEAD Fixable IR Dead Cell Stain Kit (Invitrogen, L10119). After a 10 min incubation on ice, cells were washed, fixed and permeabilized using the Foxp3 Transcription Factor Staining Buffer Set (eBioscience, 00-5523-00) according to manufacturer's protocol. Intracellular cytokines were stained for 20 min on ice in 50 ml staining volume with 1 ml of aIFNγ-APC (BD, 554702), aTNFα-FITC (BD, 554512), aIL17-PE (BioLegend, 512306) and 0.25 ml of aIL2-BV750 (BD, 566361). Cells were washed twice with FACS buffer after staining subsequently measured on the BD FACSymphony A5. The following gating strategy was applied to identify $CD8^+$ T cells: (i) selection of live (IRDye low-dim) single-cell lymphocytes [forward scatter

(FSC)-W/H low, side scatter (SSC)-W/H low, FSC/SSC-A], (ii) selection of anti-CD8$^+$ and 'dump' (anti-CD4, anti-CD14, anti-CD16, anti-CD19) negative cells. The gates applied for the identification of TNFα, IFNγ, IL-2, IL-17, CD137, HLA-DR and NKG2A positive cells of total CD8$^+$ T cells were defined based on the cells cultured with DMSO for each individual. Reactivity and activation upon peptide stimulation was defined based on a fold-change of ≥ 1.5 in expression of the cytokine and activation markers included in the analysis.

## Flow cytometer settings

The following 21-color instrument settings were used on the BD FACSymphony A5: blue laser (488 nm at 200 mW): FITC, 530/30BP, 505LP; BB630, 600LP, 610/20BP; BB790, 750LP, 780/60BP. 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, 431/28BP; BV480, 455LP, 470/20BP; BV605, 565LP, 605/40BP; BV650, 635LP, 661/11BP; BV711, 711/85, 685; BV750, 735LP, 750/30BP, BV786, 780/60BP, 750LP. UV laser (355 nm at 75 mW): BUV395, 379/28BP, BUV563, 550LP, 580/20BP; BUV615, 600LP, 615/20BP; BUV661, 630LP, 670/25BP; BUV737, 735/44BP, 770LP; BUV805, 770LP, 819/44BP.

Yellow-green laser (561 nm at 150 mW): PE, 586/15BP; PE/Dazzle-594, 600LP, 610/20BP; 2.PE-Cy7, 750LP, 780/60BP. Appropriate compensation controls were included in each analysis.

## Statistical analysis

Differences in magnitude of identified CD8 T cell responses stratified based on antigen source or recognized epitope were assessed using the non-parametric Mann-Whitney *U*-test. Ordinary one-way ANOVA test was used to assess differences in binding affinity of predicted epitopes restricted to different HLA alleles. The data cut-off for all analyses was 21 May 2020. Statistical analysis was performed using Excel 16.36 and PRISM 8 (Version 8.4.0).

## Sorting of SARS-CoV-2 specific CD8 T cells

PBMCs were thawed and washed with cold RPMI 1640 supplemented with 10% human serum, 1% Penicillin-Streptomycin solution and Benzonase nuclease (Merck-Millipore, 2500 U/mL), resuspended and incubated on ice for 60 minutes. SARS-CoV-2-specific CD8 T cells were stained with TotalSeq$^{TM}$-streptavidin-labelled (Biolegend, 405271, 405273, 405275, 405277) pHLA multimers on ice for 30 min in a final staining volume of 100ul. Subsequently cells were stained with 2 ml of aCD8-BV421 (BD, 562428), 1 ml of aCD4-FITC (BD, 345768), aCD14- FITC (BD, 345784), aCD16-FITC (BD, 335035), aCD19-FITC (BD, 345776), 0.5 ml of LIVE/DEAD Fixable IR Dead Cell Stain Kit (Invitrogen, L10119), and TotalSeq$^{TM}$-C anti-human hashtag antibodies (BioLegend, 394661, 394665, 394667, 394669, 394673) on ice for 20 min. Stained cells from 5 patients were pooled, washed and sorted on the FACSAria Fusion into 0.04% bovine serum albumin (BSA, w/v) in PBS at 4C. Forward and side scatter settings were used to select for

lymphocytes and to exclude doublets. Viable CD8 T cells were identified and sorted based on low LIVE/DEAD IRdye staining and high CD8 expression.

## Single-cell RNA-seq and TCR-seq capturing, library construction and sequencing

The single-cell suspension was split in two samples that were successively loaded onto a Chromium Single Cell Chip (10x Genomics) according to the manufacturer's protocol for co- encapsulation with barcoded Gel Beads at a capture rate of 1200 and 960 individual cells, respectively. The following 10x genomics kits were used to produce the Gel Bead-In Emulsions (GEMs) and the resulting sequence libraries (Gene expression library, Feature Barcode library, TCR library) according to manufacturer's protocol: Chromium Next GEM Single Cell 5' Library and Gel Bead kit v1.1 (10x Genomics, PN-1000167), Chromium Next Gem Single Cell V(D)J Reagent Kit v1.1 (10x Genomic, PN-1000165), Chromium Single Cell 5' Library Construction Kit (10x Genomics, PN-1000020), Chromium Single Cell 5' Feature Barcide Library Kit (10x Genomics, PN1000080) and Chromium Next GEM Chip G Single Cell kit (10x Genomics, PN-100127). The three libraries were combined in relative fractions of 0.785, 0.085 and 0.130 in order to generate sufficient reads per cell for each type of library. The final library pool was sequenced on a NextSeq Mid Flowcell, with 150 cycle chemistry kit in paired-end fashion 26-8-130bp.

## Single-cell RNA-seq data analysis

The Cell Ranger Software Suite (v.3.1.0) was used to perform sample de-multiplexing, barcode processing and single-cell 5' unique molecular identifier (UMI) counting. Single-cell RNA- seq data analysis was performed with Scanpy 1.5.1[56]. Data was analyzed with Python 3.7.6. pandas 1.0.1, and NumPy 1.18.1 were used for data manipulation, and Seaborn 0.10.0 and Matplotlib 3.1.3 were used for plotting. The following criteria were applied to each cell of all five patients: gene count between 200 and 2500, mitochondrial gene percentage <0.25 and ribosomal gene percentage >0.2. Data was then normalized to depth 10 000, and ln(1+x) was calculated. After filtering and normalization number of counts per cell and percentage of mitochondrial genes were regressed out from the data using scanpy.pp.regress_out. Data was subsequently scaled with scanpy.pp.scale using default parameters. PCA was computed on highly variable genes. A neighborhood graph of observations was computed with 50 principal components and n_neighbors=10. UMAP plots were plotted using scaled data. Louvain clustering was performed with scanpy.tl.louvain with default parameters. Marker genes were found using scanpy.tl.rank_genes_groups on the non-scaled data (use_raw=True) with t-test. Marker genes were filtered based on a minimum fold change of >2 or <0.5 and maximum Benjamini-Hochberg FDR value of 0.05. Gene ontology analysis using selected differentially expressed gene sets was performed using DAVID Bioinformatics Resources 6.8 using default settings (https://david.ncifcrf.gov/home.jsp).

## TCR V(D)J sequencing and analysis

Full length TCR V(D)J segments were enriched from amplified cDNA from 5' libraries via PCR amplification using the Chromium Single-Cell V(D)J Enrichment kit (10x Genomic, PN- 1000005) according to the manufacturer's protocol (10x Genomics). TCR sequences for each single T cell were assembled by Cell Ranger vdj pipeline (v.3.1.0), leading to the identification of CDR3 sequences and the re-arranged TCR gene. TCR repertoire analysis was performed with Scirpy 0.3. TCR diversity and TCR clonal size were estimated using scirpy.tl.alpha_diversity and ir.pl.clonal_expansion (performing the normalization), respectively. V(D)J gene usage was estimated with scirpy.pl.vdj_usage. Abundance of particular TRB V segments was estimated with scirpy.pl.group_abundance, performing the normalization. Multiple sequence alignment was performed with T-Coffee Expresso using the default settings[57].

# Declarations

Acknowledgments

We would like to thank Kees Korbee for helping to set up a work environment according to the safety regulations, Ron Kerkhoven, Marja Nieuwland and Iris de Rink for the processing of the 10x Genomics data, Frank van Diepen and Martijn van Baalen for flow cytometry support with the BD FACSymphony which was partly funded by the Louise Vehmeijer Stichting. We are also grateful for the help from Stephanie Timmer and Patty Lagerweij for getting the COVID-19 patient samples shipped to Amsterdam.

This is manuscript is dedicated to Huib Ovaa.

Author Contributions

Conceptualization: PK; Investigation: AG, SK, SP, AD, KH, PK; Data analysis: AG, SK, OI, PK; Reagents and samples: SDB, HO, RB, MT, TNS, AC; Clinical activities: AC, CM, GG, MG; Interpretation of data: AG, SK, TNS, AC, PK.

Competing Interests Statement

The authors declare no competing interests.

# References

1. https://www.worldometers.info/coronavirus/.
2. Shomuradova, A. S. *et al.* SARS-CoV-2 epitopes are recognized by a public and diverse repertoire of human T-cell receptors. *medRxiv* 1–31 (2020). doi:10.1101/2020.05.20.20107813
3. Long, Q.-X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19.*Nat Med* **62,** 1–15 (2020).
4. Ni, L. *et al.* Detection of SARS-CoV-2-Specific Humoral and Cellular Immunity in COVID-19 Convalescent Individuals. *Immunity* **52,** 971–977.e3 (2020).

5. Wajnberg, A. *et al.* Humoral immune response and prolonged PCR positivity in a cohort of 1343 SARS-CoV 2 patients in the New York City region. *medRxiv* 1–17 (2020). doi:10.1101/2020.04.30.20085613

6. Sekine, T. *et al.* Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *Cell* (2020). doi:https://doi.org/10.1016/j.cell.2020.08.017
7. Grifoni, A. *et al.* Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* 1–33 (2020). doi:10.1016/j.cell.2020.05.015
8. Le Bert, N. *et al.* SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* 1–25 (2020). doi:10.1038/s41586-020-2550-z
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,** eabd2071–11 (2020).
10. Chour, W. *et al.* Shared Antigen-specific CD8+ T cell Responses Against the SARS- COV-2 Spike Protein in HLA A*02:01 COVID-19 Participants. *medRxiv* 1–17 (2020). doi:10.1101/2020.05.04.20085779
11. Peng, Y. *et al.* Broad and strong memory CD4 +and CD8 +T cells induced by SARS- CoV-2 in UK convalescent COVID-19 patients. *bioRxiv* **8,** 420–36 (2020).
12. Thieme, C. J. *et al.* The SARS-CoV-2 T-cell immunity is directed against the spike, membrane, and nucleocapsid protein and associated with COVID 19 severity. *medRxiv* 1–17 (2020). doi:10.1101/2020.05.13.20100636
13. Neidleman, J. *et al.* SARS-CoV-2-specific T cells exhibit unique features characterized by robust helper function, lack of terminal differentiation, and high proliferative potential. *bioRxiv* **332,** 687–45 (2020).
14. Li, C. K.-F. *et al.* T cell responses to whole SARS coronavirus in humans. *Immunol.***181,** 5490–5500 (2008).

15. Channappanavar, R., Fett, C., Zhao, J., Meyerholz, D. K. & Perlman, S. Virus-specific memory CD8 T cells provide substantial protection from lethal severe acute respiratory syndrome coronavirus infection. *Journal of Virology* **88,** 11034–11044 (2014).
16. Yewdell, J. W. Confronting Complexity: Real-World Immunodominance in Antiviral CD8+ T Cell Responses. *Immunity* **25,** 533–543 (2006).
17. Diao, B. *et al.* Reduction and Functional Exhaustion of T Cells in Patients With Coronavirus Disease 2019 (COVID-19). *Immunol.* **11,** 727–7 (2020).
18. Zheng, M. *et al.* Functional exhaustion of antiviral lymphocytes in COVID-19 patients. *Cellular & Molecular Immunology* **382,** 1–3 (2020).
19. Mathew, D. *et al.* Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* eabc8511–29 (2020). doi:10.1126/science.abc8511

20. Kvistborg, P. *et al.* Anti-CTLA-4 therapy broadens the melanoma-reactive CD8+ T cell response. *Science Translational Medicine* **6,** 254ra128–254ra128 (2014).
21. http://www.allelefrequencies.net.
22. Jurtz, V. *et al.* NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *The Journal of Immunology* **199,** 3360–3368 (2017).
23. Nielsen, M., Lundegaard, C., Lund, O. & Kesmir, C. The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics* **57,** 33–41 (2005).
24. Kvistborg, P. *et al.* TIL therapy broadens the tumor-reactive CD8(+) T cell compartment in melanoma patients. *Oncoimmunology* **1,** 409–418 (2012).
25. Wang, B. *et al.* Identification of an HLA-A*0201-restricted CD8+ T-cell epitope SSp- 1 of SARS-CoV spike protein. *Blood* **104,** 200–206 (2004).
26. Davidson, A. D. *et al.* Characterisation of the transcriptome and proteome of SARS- CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Med* **12,** 68–39 (2020).
27. Finkel, Y. *et al.* The coding capacity of SARS-CoV-2. *eLife* **479–480,** 600–29 (2020).
28. Kim, D. *et al.* The Architecture of SARS-CoV-2 Transcriptome. *Cell* **181,** 914– 921.e10 (2020).
29. Taiaroa, G. *et al.* Direct RNA sequencing and early evolution of SARS-CoV-2.*bioRxiv* **34,** 4121–17 (2020).
30. Xie, X. *et al.* An Infectious cDNA Clone of SARS-CoV-2. *Cell Host and Microbe* **27,**841–848.e3 (2020).
31. van Dorp, L. *et al.* Emergence of genomic diversity and recurrent mutations in SARS- CoV-2. *Genet. Evol.* 104351 (2020). doi:10.1016/j.meegid.2020.104351
32. https://macman123.shinyapps.io/ugi-scov2-alignment-screen/.
33. 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. *The Journal of Immunology* **171,** 3928–3935 (2003).
34. Kotturi, M. F. *et al.* Naive precursor frequencies and MHC binding rather than the degree of epitope diversity shape CD8+ T cell immunodominance. *Immunol.* **181,** 2124–2133 (2008).
35. Miles, J. J., Douek, D. C. & Price, D. A. Bias in the αβ T-cell repertoire: implications for disease pathogenesis and vaccination. *Immunology and Cell Biology* **89,** 375–387 (2011).
36. Turner, S. J., Doherty, P. C., McCluskey, J. & Rossjohn, J. Structural determinants of T-cell receptor bias in immunity. *Nature Reviews Immunology* **6,** 883–894 (2006).
37. Amanat, F. & Krammer, F. SARS-CoV-2 Vaccines: Status Report. *Immunity* **52,** 583– 589 (2020).

38. Thanh Le, T. *et al.* The COVID-19 vaccine development landscape. *Nature Reviews Drug Discovery* **19,** 305–306 (2020).
39. Flores-Villanueva, P. O. *et al.* Control of HIV-1 viremia and protection from AIDS are associated with HLA-Bw4 homozygosity. *Proc Natl Acad Sci USA* **98,** 5140–5145 (2001).
40. Zhou, J., Matsuoka, M., Cantor, H., Homer, R. & Enelow, R. I. Cutting edge: engagement of NKG2A on CD8+ effector T cells limits immunopathology in influenza pneumonia. *The Journal of Immunology* **180,** 25–29 (2008).
41. Giamarellos-Bourboulis, E. J. *et al.* Complex Immune Dysregulation in COVID-19 Patients with Severe Respiratory Failure. *Cell Host and Microbe* **27,** 992–1000.e3 (2020).
42. Mazzoni, A. *et al.* Impaired immune cell cytotoxicity in severe COVID-19 is IL-6 dependent. *Clin. Invest.* **105,** 93–11 (2020).
43. Morandi, F., Airoldi, I. & Pistoia, V. IL-27 driven upregulation of surface HLA-E expression on monocytes inhibits IFN-γ release by autologous NK cells. *Journal of Immunology Research* **2014,** 938561–7 (2014).
44. Pereira, B. I. *et al.* Senescent cells evade immune clearance via HLA-E-mediated NK and CD8+T cell inhibition. *Nature Communications* **10,** 2387–13 (2019).
45. Pyle, C. J., Uwadiae, F. I., Swieboda, D. P. & Harker, J. A. Early IL-6 signalling promotes IL-27 dependent maturation of regulatory T cells in the lungs and resolution of viral immunopathology. *PLoS Pathog* **13,** e1006640–27 (2017).
46. Chen, Z. & Wherry, J. E. T cell responses in patients with COVID-19. *Nature Reviews Immunology* 1–8 (2020). doi:10.1038/s41577-020-0402-6
47. Cao, X. COVID-19: immunopathology and its implications for therapy. *Nature Reviews Immunology* **20,** 269–270 (2020).
48. Poran, A. *et al.* Sequence-based prediction of vaccine targets for inducing T cell responses to SARS-CoV-2 utilizing the bioinformatics predictor RECON. **186,** 7264– 30 (2020).
49. Campbell, K. M., Steiner, G., Wells, D. K., Ribas, A. & Kalbasi, A. Prediction of SARS-CoV-2 epitopes across 9360 HLA class I alleles. *bioRxiv* **90,** 1–12 (2020).
50. Grifoni, A. *et al.* A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host and Microbe* **27,** 671–680.e2 (2020).
51. Prachar, M. *et al.* COVID-19 Vaccine Candidates: Prediction and Validation of 174 SARS-CoV-2 Epitopes. *bioRxiv* **7,** 13404–14 (2020).
52. Rodenko, B. *et al.* Generation of peptide-MHC class I complexes through UV- mediated ligand exchange. *Nat Protoc* **1,** 1120–1132 (2006).
53. Garboczi, D. N., Hung, D. T. & Wiley, D. C. HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. *Proc Natl Acad Sci USA* **89,** 3429–3433 (1992).
54. Toebes, M. *et al.* Design and use of conditional MHC class I ligands. *Nat Med* **12,** 246–251 (2006).

55. Hadrup, S. R. *et al.* Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat Meth* **6,** 520–526 (2009).
56. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* **19,** 15–5 (2018).
57. Armougom, F. *et al.* Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Research* **34,** W604–8 (2006).
58. Bakker, A. H. *et al.* Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Natl. Acad. Sci. U.S.A.* **105,** 3825–3830 (2008).
59. Chang, C. X. L. *et al.* Conditional ligands for Asian HLA variants facilitate the definition of CD8+ T-cell responses in acute and chronic viral diseases. *J. Immunol.* **43,** 1109–1120 (2013).

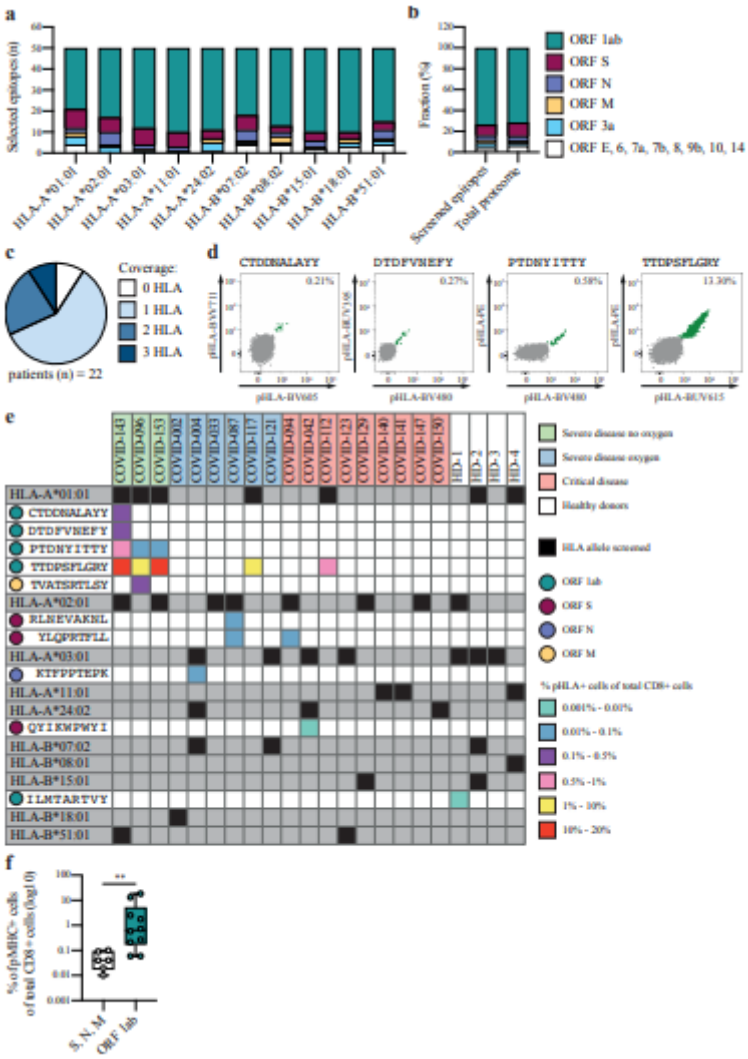
## Table

# Table 1. Characteristics of COVID-19 patients and healthy donors.

Covered HLA alleles are indicated in bold. The outcome of COVID-19 patient 42 is unknown as this patient was transferred to a different hospital. HLA: human leukocyte antigen, N/A: not available/applicable

| Donor ID | Age (years) | Therapy | Disease state | Days on therapy prior to sampling | Patient outcome | HLA-A1 | HLA-A2 | HLA-B1 | HLA-B2 |
|---|---|---|---|---|---|---|---|---|---|
| COVID-143 | 73 | Tocilizumab | Severe (no oxygen) | 10 | Discharged | **A*01:01** | **A*02:01** | B*44:03 | **B*51:01** |
| COVID-096 | 43 | Tocilizumab | Severe (no oxygen) | 7 | Discharged | **A*01:01** | A*23:01 | B*49:01 | B*57:01 |
| COVID-153 | 52 | Tocilizumab | Severe (no oxygen) | 7 | Discharged | **A*01:01** | **A*02:01** | B*35:01 | B*35:02 |
| COVID-002 | 88 | None | Severe (oxygen) | N/A | Died | A*03:02 | A*32:01 | **B*18:01** | B*44:02 |
| COVID-004 | 36 | None | Severe (oxygen) | N/A | Discharged | **A*03:01** | **A*24:02** | **B*07:02** | B*35:01 |
| COVID-009 | 60 | Tocilizumab | Severe (oxygen) | 3 | Discharged | A*02:05 | A*32:01 | B*13:02 | B*14:02 |
| COVID-033 | 75 | Tocilizumab | Severe (oxygen) | 0 | Discharged | **A*02:01** | A*68:01 | B*35:01 | B*49:01 |
| COVID-087 | 52 | Tocilizumab | Severe (oxygen) | 8 | Discharged | **A*02:01** | A*03:02 | B*39:01 | B*50:01 |
| COVID-116 | 80 | Tocilizumab | Severe (oxygen) | 1 | Discharged | **A*01:01** | A*26:01 | B*13:02 | B*44:02 |
| COVID-117 | 67 | Tocilizumab | Severe (oxygen) | 1 | Discharged | **A*01:01** | A*31:01 | B*15:17 | B*38:01 |
| COVID-121 | 34 | Tocilizumab | Severe (oxygen) | 8 | Discharged | **A*03:01** | A*74:01 | **B*07:02** | B*42:01 |
| COVID-127 | 72 | Tocilizumab | Severe (oxygen) | 2 | Died | A*29:02 | A*31:01 | B*35:01 | B*44:03 |
| COVID-152 | 66 | Tocilizumab | Severe (oxygen) | 1 | Died | **A*11:01** | A*30:01 | B*13:02 | B*53:01 |
| COVID-094 | 75 | Tocilizumab | Critical | 13 | Died | **A*02:01** | A*32:01 | B*51:08 | B*56:01 |
| COVID-042 | 50 | Anakinra | Critical | 2 | Transferred | **A*03:01** | **A*24:02** | B*35:02 | B*35:02 |
| COVID-112 | 70 | Tocilizumab | Critical | 10 | Discharged | **A*01:01** | A*26:01 | B*35:02 | B*35:03 |
| COVID-123 | 70 | Tocilizumab | Critical | 13 | Died | **A*03:01** | A*32:01 | B*27:05 | **B*51:01** |
| COVID-129 | 76 | Tocilizumab | Critical | 15 | N/A | **A*02:01** | A*29:02 | **B*15:01** | B*44:03 |
| COVID-140 | 56 | Tocilizumab | Critical | 6 | Died | **A*11:01** | A*68:02 | B*35:01 | B*53:01 |
| COVID-141 | 77 | Tocilizumab | Critical | 6 | N/A | **A*11:01** | A*68:01 | B*35:03 | B*50:01 |
| COVID-147 | 71 | Tocilizumab | Critical | 6 | N/A | **A*02:01** | A*26:01 | B*38:01 | B*44:02 |
| COVID-150 | 74 | Anakinra | Critical | 6 | Died | **A*24:02** | A*29:01 | B*13:02 | B*35:03 |
| HD1 | 73 | N/A | N/A | N/A | N/A | **A*02:01** | **A*03:01** | **B*15:01** | B*35:01 |
| HD2 | 77 | N/A | N/A | N/A | N/A | **A*01:01** | **A*03:01** | **B*07:02** | B*40:01 |
| HD3 | 73 | N/A | N/A | N/A | N/A | **A*03:01** | **A*03:01** | B*35:01 | B*35:01 |
| HD5 | 54 | N/A | N/A | N/A | N/A | **A*01:01** | **A*11:01** | **B*08:01** | B*38:01 |

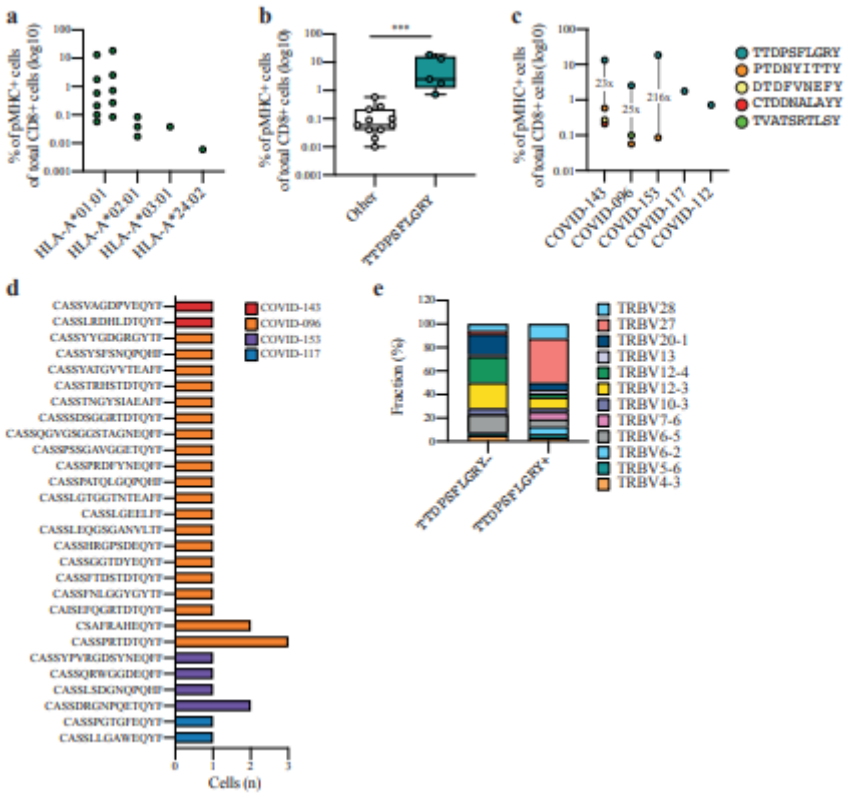
# Figures




## Figure 1

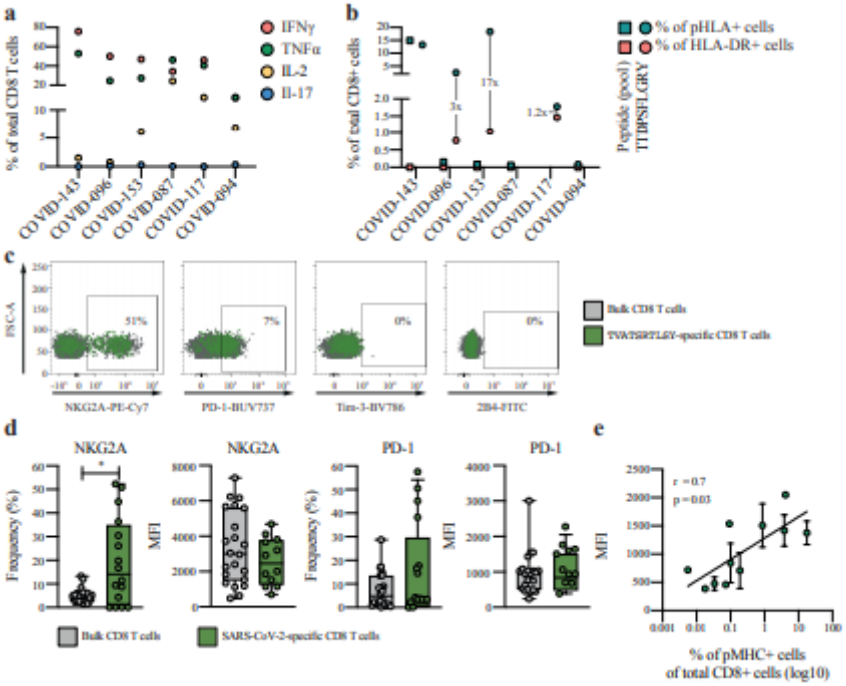
Identification of the SARS-CoV-2-specific CD8 T cell response. a) Distribution of selected SARS-CoV-2 epitopes within the different ORFs of the SARS- CoV-2 proteome for each of the 5 HLA-A and 5 HLA-B alleles included in this study. S: spike, N: nucleoprotein, M: membrane, E: envelope. b) Proportion of screened SARS-CoV-2 epitopes compared to the fraction of the total SARS-CoV-2 proteome for each ORF of the SARS-CoV-2 virus. (Details shown in Table S1). c) HLA allele coverage with the 10 selected HLA alleles is shown for the patient cohort (n=22). d) Representative flow cytometry plots of 4 different SARS-CoV-2-specific CD8 T cell responses (green) detected in patient COVID-143. Magnitude of SARS-CoV-2-specific CD8 T cell responses represents the percentage of double-positive pHLA+ cells of total CD8+ cells (grey). A representative example of the full gating strategy is provided in Fig. S1. e) Heatmap of all detected SARS-CoV-2-specific CD8 T cell responses (n=16) including the information about specific epitopes (n=9) and their viral protein origin, the magnitude of the response and HLA coverage and disease status for each individual (COVID-19 patients, n=18, healthy donors, n=4). f) The magnitude of SARS-CoV-2-specific CD8 T cell responses detected in COVID-19 patients directed towards epitopes from ORF1ab (n=10) compared to other SARS- CoV-2 proteins (n=6) combined. Statistical significance was

tested with a two-tailed Mann-Whitney U-test. p=0.0027. S: spike, N: nucleoprotein, M: membrane, E: envelope.




## Figure 2

Characterization of the immunodominant SARS-CoV-2-specific CD8 T cell response. a) Number of detected SARS-CoV-2-specific CD8 T cell responses and their magnitude for each individual HLA-A allele: HLA-A*01:01 (n=11), HLA-A*01:01 (n=3), HLA- A*03:01 (n=1), HLA-A*24:02 (n=1) that were detected in COVID-19 patients. b) Differences in the magnitude of SARS-CoV-2-specific CD8 T cell responses specific for the immunodominant epitope TTDPSFLGRY (n=5) compared to other (n=11) detected SARS-CoV-2-specific CD8 T cell responses in COVID-19 patients. Statistical significance was tested with a two-tailed Mann-Whitney U-test. p=0.0002. c) Overview of detected SARS-CoV-2-specific CD8 T cell response in COVID-19 patients (n=5) positive for HLA-A*01:01. The numbers indicate the fold difference in the magnitude between the TTDPSFLGRY-specific CD8 T cell response and the response with the second-highest magnitude within the patient.Overview of unique TCR beta chains (n=28) expressed in TTDPSFLGRY-specific CD8 T cells (n=32) detected in COVID-19 patients (n=4). The number of single cells expressing each individual TCR beta chain is indicated on the x-axis. d) Fractions of cells expressing different TCRBV regions that were found in TTDPSFLGRY+ CD8 T cells (n=32) compared to TTDPSFLGRY-CD8 T cells (n=842).




## Figure 3

Functional assessment of SARS-CoV-2-specific CD8 T cells. a) Cytokine production (IFNγ, TNFα, IL-2 and IL-17) in CD8 T cells from COVID-19 patients (n=6) upon stimulation with PMA/IO. b) Detection of SARS-CoV-2-specific CD8 T cell responses in COVID-19 patients (n=6) using the pHLA multimer assay (green) compared to HLA-DR expression upon peptide stimulation (red). The magnitude of SARS-CoV-2-specific CD8 T cell responses using the pHLA multimer assay was assessed using single peptides. The magnitude of SARS- CoV-2-specific CD8 T cell responses using the peptide stimulation assay (background signal in a DMSO negative control was subtracted from the readout) was assessed with peptide pools (square) or single peptides (circle): COVID-143: peptide pool (CTD+DTD+PTD+TTD); COVID-096: peptide pool (PTD+TVA) and TTD; COVID- 153: PTD and TTD; COVID-087: peptide pool (RLN+YLQ); COVID-117: TTD; COVID-094: YLQ). c) Representative flow cytometry plots of inhibitory receptor expression (NKG2A, PD-1, Tim-3, 2B4) in SARS-CoV-2-specific CD8 T cells (green) and bulk CD8 T cells (grey) in patient COVID-096. Percentages of SARS-CoV-2-specific CD8 T cells expressing individual inhibitory receptor (% of total pHLA+ inhibitory receptor+ CD8+ cells) is shown. d) Frequencies and MFI of PD-1+ and NKG2A+ cells of bulk CD8+ (grey, n =22) compared to pHLA+ (green, n = 16) cells. Statistical significance was tested with a two- tailed Mann-Whitney U-test. Only significant p values are shown, *p=0.041. e) Correlation analysis between PD-1 expression levels (MFI and range) and the magnitude of detected SARS-CoV-2-specific CD8 T cell responses (n = 11). Spearman correlation analysis was performed, p=0.03, r=0.7.

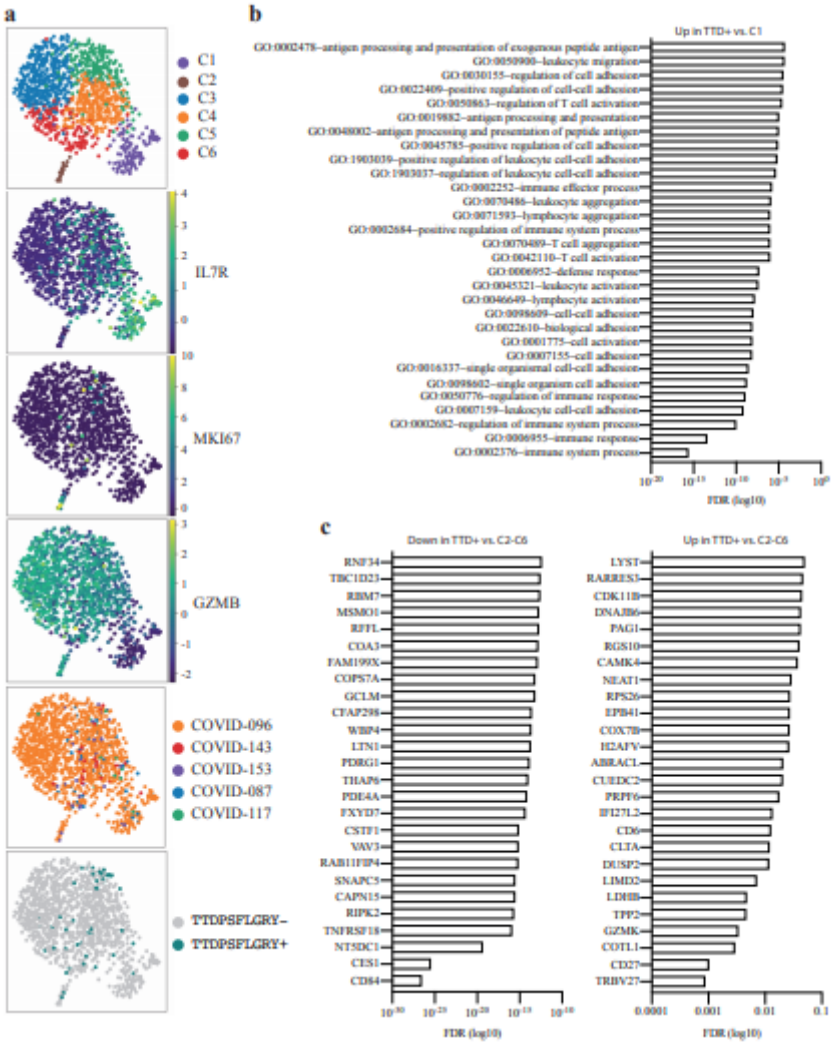



**Figure 4**

Single-cell transcriptome analysis of SARS-CoV-2-specific CD8 T cells. a) UMAP presentation of single-cell gene expression data of CD8 T cells (n=1180) and TTD-specific CD8 T cells (n=48) isolated from COVID-19 patients (n=5) illustrating clusters (n=6) identified using Louvain algorithm (full list of differentially expressed genes for each cluster is provided in Table S4), expression levels of selected genes for different clusters (IL7R, MKI67 and GZMB), distribution of single cells for each patient and TTD-specific CD8 T cells. b) Top 30 biological processes that were found to be upregulated by TTD-specific CD8 T cell compared to cluster C1. Inclusion criteria of genes included in the analysis were FDR<0.05 and fold change of >2. The full list is provided in Table S4. c) Top 26 differentially expressed genes that were found to be significantly (FDR<0.05) up- (fold change of >2) or down-regulated (fold change <0.5) in TTD-specific CD8 T cell compared to clusters C2-C6. The full list of genes is provided in Table S4.

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GangaevetalSupplemental.pdf