

IMMUNOLOGY

Celluloepidemiology—A paradigm for quantifying infectious disease dynamics on a population level

My K. Ha^{1,2,3*}, Anna Postovskaya^{3,4,5,6}, Maria Kuznetsova^{1,2,3}, Pieter Meysman^{3,4,5}, Vincent Van Deuren^{3,4,5}, Sabrina Van Ierssel⁷, Hans De Reu^{8,9}, Jolien Schippers^{1,2,3}, Karin Peeters^{1,2}, Hajar Besbassi^{1,2,3}, Leo Heyndrickx¹⁰, Betty Willems¹⁰, Joachim Mariën^{11,12}, Esther Bartholomeus^{2,3}, Koen Vercauteren⁶, Philippe Beutels¹, Pierre Van Damme^{3,13}, Eva Lion^{3,8,9}, Erika Vlieghe^{7,14}, Kris Laukens^{3,4,5}, Samuel Coenen^{15,16}, Reinout Naesens¹⁷, Kevin K. Ariën^{10,18}, Benson Ogunjimi^{1,2,3,19*}

Copyright © 2025 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

To complement serology as a tool in public health interventions, we introduced the “celluloepidemiology” paradigm where we leveraged pathogen-specific T cell responses at a population level to advance our epidemiological understanding of infectious diseases, using SARS-CoV-2 as a model. Applying flow cytometry and machine learning on data from more than 500 individuals, we showed that the number of T cells with positive expression of functional markers not only could distinguish patients who recovered from COVID-19 from controls and pre-COVID donors but also identify previously unrecognized asymptomatic patients from mild, moderate, and severe recovered patients. The celluloepidemiology approach was uniquely capable to differentiate health care worker groups with different SARS-CoV-2 exposures from each other. T cell receptor (TCR) profiling strengthened our analysis by revealing that SARS-CoV-2–specific TCRs were more abundant in patients than in controls. We believe that adding data on T cell reactivity will complement serology and augment the value of infection morbidity modeling for populations.

INTRODUCTION

Serology has been a powerful tool in public health, infectious disease research, and the evaluation of vaccination programs (1). In seroepidemiology, data on antibody prevalence and titers in (human) serum are collected and analyzed to examine the distribution and determinants of infections in populations (2). The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) at the end of 2019 has caused a global pandemic of COVID-19 and, thus, required population-wide epidemiological studies to investigate the causes,

risk factors, outcomes, and eventually relieve the associated burden. However, the serological analysis of patients with COVID-19 is complicated by the variability of the SARS-CoV-2–induced antibody responses. For example, SARS-CoV-2 antibody titers are higher in patients with severe disease and lower in young, asymptomatic, or pauci-symptomatic individuals (3, 4). Although immunoglobulin G (IgG) antibodies targeting SARS-CoV-2 receptor binding domain (RBD) were reported to persist at detectable levels beyond 3 months after symptom onset (5, 6), IgM and IgA responses were short-lived and waned within 2.5 months, especially in asymptomatic cases (3, 5, 7).

On the other hand, SARS-CoV-2–induced cell-mediated immunity is considered more sustained but remains poorly characterized. Within cell-mediated immunity, the role of T cells in the exacerbation of COVID-19 and their potential to provide long-term immunity against SARS-CoV-2 has been increasingly discussed and studied. There is evidence that, compared with that from healthy controls, T cell compartment from patients with COVID-19 displays several alterations in the proliferation, expression of lineage-specifying receptors, and production of cytokines. For example, increases in the expression of inhibitory receptors such as programmed cell death protein 1 (PD-1), T cell immunoglobulin and mucin domain-containing protein 3 (TIM-3), lymphocyte activation gene 3 (LAG-3), and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) have been observed in highly activated or possibly exhausted T cells in severe acute COVID-19 cases (8–12). An increase in CD38⁺HLA-DR⁺ activated CD8⁺ T cell population was also reported in many patients with active SARS-CoV-2 infection (8, 9, 13). T cell activation in patients with acute COVID-19 was found to be skewed toward a T helper 17 (T_H17) functional phenotype, which implies a potential role of T_H17-mediated immunopathology in COVID-19 (14). Regulatory T cells and ICOS⁺CD38⁺ circulating follicular helper T cells have also been reported to be altered in patients with COVID-19 (9, 15). CD4⁺ T

¹Center for Health Economics Research and Modelling Infectious Diseases (CHERMID), Vaccine and Infectious Disease Institute, University of Antwerp, Wilrijk, Belgium.

²Antwerp Center for Translational Immunology and Virology (ACTIV), Vaccine and Infectious Disease Institute, University of Antwerp, Wilrijk, Belgium. ³Antwerp Unit for Data Analysis and Computation in Immunology and Sequencing (AUDACIS), University of Antwerp, Antwerp, Belgium. ⁴ADReM Data Lab, Department of Mathematics and Computer Science, University of Antwerp, Antwerp, Belgium. ⁵Biomedical Informatics Research Network Antwerp (biomina), University of Antwerp, Antwerp, Belgium. ⁶Clinical Virology Unit, Department of Clinical Sciences, Institute of Tropical Medicine, Antwerp, Belgium. ⁷Department of General Internal Medicine, Infectious Disease and Tropical Medicine, Antwerp University Hospital, Edegem, Belgium. ⁸Laboratory of Experimental Hematology (LEH), Vaccine and Infectious Disease Institute, University of Antwerp, Wilrijk, Belgium. ⁹Flow Cytometry and Cell Sorting Core Facility (FACSUA), University of Antwerp, Wilrijk, Belgium. ¹⁰Virology Unit, Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium. ¹¹Department of Ecology and Evolutionary Biology, University of Antwerp, Antwerp, Belgium. ¹²The Virus Ecology Group, Institute of Tropical Medicine, Antwerp, Belgium. ¹³Centre for the Evaluation of Vaccination (CEV), Vaccine and Infectious Disease Institute, University of Antwerp, Wilrijk, Belgium. ¹⁴Global Health Institute, University of Antwerp, Wilrijk, Belgium. ¹⁵Laboratory of Medical Microbiology (LMM), Vaccine and Infectious Disease Institute (VAXINFECTIO), University of Antwerp, Wilrijk, Belgium. ¹⁶Center for General Practice, Department of Family Medicine and Population Health (FAMPOP), University of Antwerp, Wilrijk, Belgium. ¹⁷Department of Clinical Biology, Antwerp Hospital Network, Antwerp, Belgium. ¹⁸Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium. ¹⁹Department of Pediatrics, Antwerp University Hospital, Edegem, Belgium.

*Corresponding author. Email: my.ha@uantwerp.be (M.K.H.); benson.ogunjimi@uantwerp.be (B.O.)

cells in patients who recovered from mild COVID-19 were found to gain a typical memory phenotype with high expression of IL-7R α (16). Other studies have targeted the functions of T cells in recognizing and acting against the structural SARS-CoV-2 viral proteins spike, membrane, and nucleocapsid (NP) (17–20). T cell receptor (TCR) sequencing (TCR-seq) has also shown that T cells can offer high sensitivity and specificity in the detection of past SARS-CoV-2 infection (21, 22). Therefore, deeper and more comprehensive immune profiling of antigen-specific TCRs could enable better discrimination between infected and uninfected individuals, as well as improve our understanding of cell-mediated immunity against pathogens and its potential relevance in inferring susceptibility differences in the population. We believe that the systematic study of T cell responses against pathogens on a population level could introduce an unprecedented paradigm referred to as “celluloepidemiology.” The celluloepidemiology approach could leverage the assessment of human immune response dynamics against various pathogenic agents, given the diversity of biomarkers expressed and secreted by antigen-specific T cells in contrast to the unidimensional nature of antibodies (23).

In this study, we present a feasible experimental setup and the essential complementary computational approaches to perform celluloepidemiology on a SARS-CoV-2 model. We will compare celluloepidemiology with seroepidemiology at four levels: (i) sensitivity (i.e., the capacity to measure the genuine proportion of previously infected individuals), (ii) specificity (i.e., the capacity to identify individuals who have not been infected), (iii) morbidity/disease categorization (i.e., the capacity to distinguish between different disease groups among infected individuals), and (iv) health-related consequences [i.e., the capacity to detect immune alterations after an infection, such as in long Covid cases; (24)].

RESULTS

Comparison of ex vivo immunological phenotypes between patients who recovered from COVID-19, household members, controls, and pre-COVID donors

To obtain the best T cell surface markers that would then allow us to look into antigen-specific T cell responses at a population level, we analyzed peripheral blood mononuclear cells (PBMCs) of 30 patients who recovered from COVID-19 and 15 pre-COVID blood donors stimulated by SARS-CoV-2-specific major histocompatibility complex (MHC) class I- and class II-specific peptide pools (thus, without expected cross-reactivity with other coronaviruses), as well as SARS-CoV-2 membrane protein peptide pools (SC2-MP) using an extensive 33-isotope mass cytometry [cytometry by time-of-flight (CyTOF)] panel (table S1) and applied the Boruta algorithm (25) to look for the best differentiation between these groups. Figure S1 displays the selected T cell surface markers. High-importance values indicate markers whose expressions differ greatly between controls, patients, and patient subgroups (i.e., mild, moderate, and severe COVID-19). The Boruta-selected membrane markers were CTLA-4, human leukocyte antigen (HLA-DR), PD-1, TIM-3, CD27, CD28, CD38, Fas, T cell immunoreceptor with Ig and ITIM domains (TIGIT), CD154, OX40, CD137, and CD69 (fig. S1). These markers were divided into two panels: The first panel, taking into account representative markers of cells' lineage and functional status, consisted of CD45RA, CCR7, CD27, CD28, CD38, Fas, TIGIT, CD154, OX40, CD137, and CD69 among other markers; and the second panel included the remaining CTLA-4, HLA-DR, PD-1, and TIM-3.

In a second phase, we tried to assess whether previous SARS-CoV-2 infection would continue to leave an impression on the immune system after more than 3 months. To achieve this, we analyzed the PBMCs of 478 individuals, including 166 patients who recovered from COVID-19, 29 household members of patients who recovered from COVID-19, 259 controls, and 24 pre-COVID donors (group definitions are presented in Methods). Household members [i.e., individuals who lived in the same household with patients with COVID-19 and thus had close contact and high chance of exposure to SARS-CoV-2 but did not have confirmed infection, either by polymerase chain reaction (PCR) or IgG test] were involved in this analysis to examine the effects of SARS-CoV-2 exposure on T cells after more than 3 months. The PBMCs were stimulated by MHC class I and class II SARS-CoV-2-specific and SC2-MP. Flow cytometry measurements were performed using the first panel (table S2) to broadly examine the immunophenotype profiles of different groups.

First, FlowSOM was implemented to semiautomatically gate CD4⁺ and CD8⁺ T cell subsets based on CD45RA and CCR7 expressions (fig. S2). Differences in T cell frequencies between different groups indicate that the naïve T cell subset was reduced in patients who recovered from COVID-19, the effector memory subset was increased in patients who recovered from COVID-19 and household members, whereas the central memory and terminally differentiated effector memory subsets were more present in patients who recovered from severe COVID-19 than in other groups. It is, however, worth noting that the distribution of memory and naïve T cell compartments is strongly affected by age. In our cohort, severe COVID-19 cases were all older than 57 years at the time of recruitment, whereas other groups included younger donors (table S5). In fig. S1, we performed statistical significance test on the age of different groups and showed that median age of severe COVID-19 cases was significantly higher than that of other groups. However, this does not disregard our observation of CD4⁺ and CD8⁺ naïve and memory compartments but rather adds to the epidemiological perspective of this study.

Patients who recovered from COVID-19 had more SARS-CoV-2-specific T cells than controls

To explore further the complementary values of celluloepidemiology to seroepidemiology, we compared SARS-CoV-2-specific IgG and T cell data between patients who recovered from COVID-19 and controls. Figure 1A displays that, after stimulation with SC2-MP, CD4⁺ and CD8⁺ T cells in recovered patients expressed significantly higher levels of CD38, Fas, TIGIT, CD154, OX40, CD137, and CD69 ($P \leq 0.0001$) than those in controls, emphasizing that there is a great number of T cells getting activated, exhausted, or apoptotic upon reencounter with SARS-CoV-2 peptides in recovered patients even more than 3 months after COVID-19 onset. Household members had lower Fas expression than patients ($P \leq 0.0001$) but higher TIGIT expression than controls ($P \leq 0.01$), indicating non-apoptotic exhaustion of T cells following reencounter with SARS-CoV-2 ex vivo. In terms of serology, fig. S3A reveals that recovered patients had significantly higher levels of SARS-CoV-2 NP, RBD, and S1-S2-specific IgG ($P \leq 0.0001$) than controls. Household member had higher IgG reactivity than controls ($P \leq 0.01$) but lower than patients ($P \leq 0.01$). Simultaneously, we developed T cell classifiers that would allow us to adequately differentiate patients from controls. We applied random forest algorithm on the flow cytometry data. Expressions of eight functional markers (i.e., CD38, CD69, Fas, OX40, CD137, CD154, LAG-3, and TIGIT) in eight T cell populations were

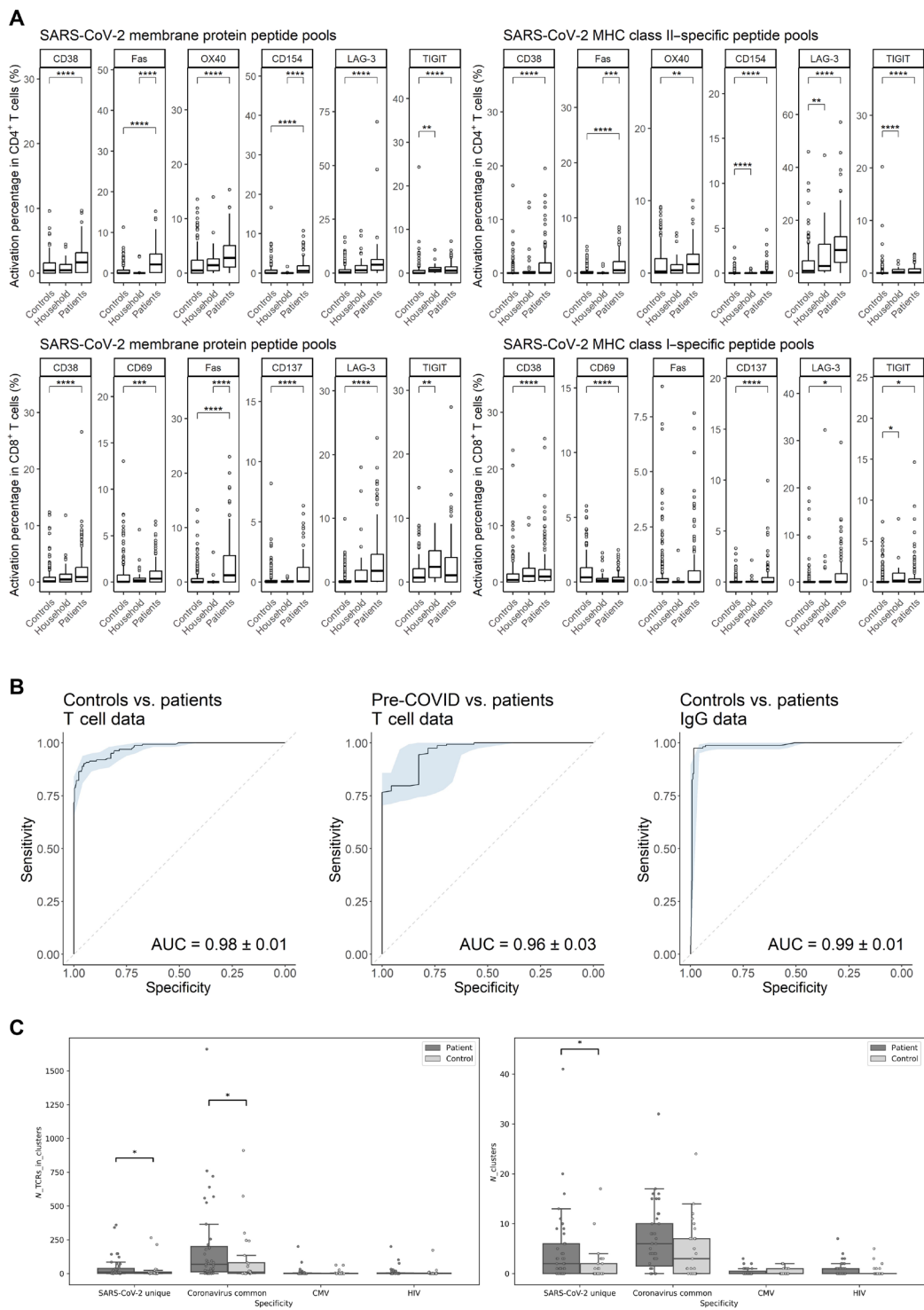


Fig. 1. Comparison of SARS-CoV-2-specific T cell immune response between controls, patients, and household members. (A) Net percentages of T cells with positive expression of functional markers out of the total numbers of CD4⁺ and CD8⁺ T cells poststimulation (normalized with unstimulated controls). Horizontal lines represent median values of each group. Statistical significance was analyzed by Kruskal-Wallis test with Bonferroni correction. **(B)** Receiver operating characteristics (ROCs) and areas under the curve (AUCs) from leave-one-out cross-validation of random forest classifiers using T cell data and IgG data. **(C)** Numbers of clustered TCRs and clusters specific to SARS-CoV-2, coronaviruses, cytomegalovirus (CMV), and HIV in patients and controls. Statistical significance was analyzed by nonparametric *t* test with Benjamini-Hochberg correction. **P* ≤ 0.05; ***P* ≤ 0.01; ****P* ≤ 0.001; *****P* ≤ 0.0001.

used as parameters on which the random forest classifiers were developed. Leave-one-out cross-validation of the first bilateral classifiers confirms that patients could be differentiated from controls and pre-COVID donors using flow cytometry T cell data with area under the curve (AUC) = 0.98 ± 0.01 and 0.96 ± 0.03 , which was comparable to the classifier between controls and patients using IgG data with AUC = 0.99 ± 0.01 (Fig. 1B). On the basis of the results of the first T cell classifiers, a separate group of individuals were defined as “asymptomatic patients” as they initially were recruited to be control cases who did not report COVID-19–related symptoms nor knowingly exposure but were classified as “patients” by the T cell classifier. We note that the celluloeidemiology framework allowed this discovery of asymptomatic patients.

After looking into SARS-CoV-2–specific T cells on a phenotypic level, we took a step further by investigating the ex vivo CD8⁺ TCR repertoires. Because similar TCRs that cluster together generally target the same epitope and, thus, could reveal traces of T cell response, we compared the prevalence of CD8⁺ TCR clusters between controls and patients (Fig. 1C). Regarding the number of unique TCRs in the respective repertoires, recovered patients had significantly more CD8⁺ clusters than controls ($P \leq 0.001$). We further differentiated the CD8⁺ TCRs between those uniquely reactive to SARS-CoV-2 and those that were considered cross-reactive to different coronaviruses. Figure 1C reveals that there was a significantly higher number of SARS-CoV-2 unique clusters and TCRs per cluster in recovered patients compared to healthy controls ($P \leq 0.05$). Patients were also found to have more coronavirus-specific TCRs per cluster than healthy controls ($P \leq 0.05$). Cytomegalovirus and HIV were included as negative controls. Table S8 contains a detailed list of annotated TCR-epitope specificity.

T cell immune profiling is comparable to serological testing in identifying patients who recovered from mild, moderate, and severe COVID-19

Next, we investigated whether celluloeidemiology and/or seroepidemiology could differentiate between previous mild, moderate, and severe COVID-19. Asymptomatic and mild patients were effectively distinguished from moderate and severe patients using both IgG and T cell data (Fig. 2A and fig. S3B). Looking at serological data, we found that the amount of SARS-CoV-2 RBD, NP, and S1-S2 antigen-specific IgG significantly increased as COVID-19 severity increased ($P \leq 0.01$). However, moderate and severe patients could only be distinguished from each other by RBD reactivity ($P \leq 0.05$), not by NP and S1-S2 reactivity. Looking at flow cytometry data, we observed that, upon stimulation by SC2-MP, CD4⁺ T cells displayed an increase in Fas and OX40 expressions as COVID-19 severity increases from asymptomatic to severe ($P \leq 0.05$). A similar pattern could be observed in CD8⁺ T cells upon the same stimulation in the expressions of CD69, Fas, CD137, and TIGIT ($P \leq 0.05$). Particularly, moderate and severe patients differed significantly in their Fas, CD137, and TIGIT expressions ($P \leq 0.05$). These observations are consistent with the increased expression of TIGIT and Fas in severe COVID-19 cases reported by Neidلمان *et al.* (26). The correlation between COVID-19 severity and T cells' expression of CD137, Fas, OX40, and TIGIT reflects an increase in T cell differentiation from naïve to effector/memory phenotypes upon SARS-CoV-2 reencounter as severity increases. This finding adds an important insight, on a population level, to the understanding of pathogen-specific immunity postinfection that could not be obtained from serology. The

multilateral T cell classifiers also performed as effectively as the serology-based approach (Fig. 2B). Using T cell flow cytometry data, asymptomatic patients could be better distinguished from mild, moderate, and severe COVID-19 with AUCs of 0.84 ± 0.10 , 0.88 ± 0.10 , and 0.92 ± 0.10 , respectively, compared to the other pairs (i.e., mild versus moderate, mild versus severe, and moderate versus severe). This is comparable to their equivalents using IgG data, with AUCs being 0.81 ± 0.16 , 0.87 ± 0.13 , and 0.90 ± 0.12 . Overall, these results indicate that a T cell–based infection classification on a population level would yield similar accuracy to the serology-based approaches. However, it should be noted that, in this study, we considerably augmented the capacity of serology to classify and stratify by applying our machine learning framework.

In-depth multiparameter characterization of SARS-CoV-2–specific T cell responses using FlowSOM clustering

One of the potential advantages of the celluloeidemiology approach would be the inherent multiparameter capacity imbedded in T cell assays compared to the limited parameters (at best IgG and IgM titers, and perhaps neutralization assays and avidity testing) in serological assays. To explore this further, CD4⁺ and CD8⁺ T cells in controls, patients who recovered from COVID-19, and household members were gated by their expressions of eight functional markers (i.e., CD154, OX40, CD137, CD69, CD38, Fas, LAG-3, and TIGIT) after stimulation with SARS-CoV-2 peptides. Cells having higher expression of more than two activation markers after stimulation with SARS-CoV-2–derived peptides compared to unstimulated controls were here defined as SARS-CoV-2–specific T cells. To gain more insights into these SARS-CoV-2–specific T cells, we applied FlowSOM clustering in comparing expression of eight activation markers and identified 12 metaclusters for CD4⁺ T cells (Fig. 3) as well as 16 metaclusters for CD8⁺ T cells (Fig. 4). *t*-distributed stochastic neighbor embedding (*t*-SNE) representation of the data highlighted a distinction between controls and patients, but not so much between asymptomatic patients and patients with mild, moderate, and severe COVID-19 (Figs. 3A and 4A).

We looked further into SARS-CoV-2–specific T cell immune responses in controls, patients, and household members by investigating the expressions of PD-1, HLA-DR, TIM-3, and CTLA-4 in SARS-CoV-2–specific CD154^{high}OX40^{high}CD4⁺ and CD137^{high}CD69^{high}CD8⁺ T cells (figs. S4 and S5). In CD154^{high}OX40^{high}CD4⁺ T cells (fig. S4), PD-1 expression was significantly higher in patients than in household members ($P \leq 0.05$). In CD137^{high}CD69^{high}CD8⁺ T cells (fig. S5), we found CTLA-4, HLA-DR, and PD-1 significantly higher in patients than in controls ($P \leq 0.01$), while TIM-3 was highly expressed in patients and household members compared to controls ($P \leq 0.05$). It is evident that CD8⁺ T cells in patients who recovered from COVID-19 and their household members (who were exposed to SARS-CoV-2) became more activated and exhausted than CD4⁺ T cells after reexposure ex vivo.

Celluloeidemiology allows differentiation between health care worker groups

Beside the control, recovered patient, and household member groups, the health care worker cohort could also offer interesting and relevant insights due to their frequent exposure to SARS-CoV-2 in a professional context. To explore this scenario, we recruited different health care worker groups (hospital personnel and general practitioners

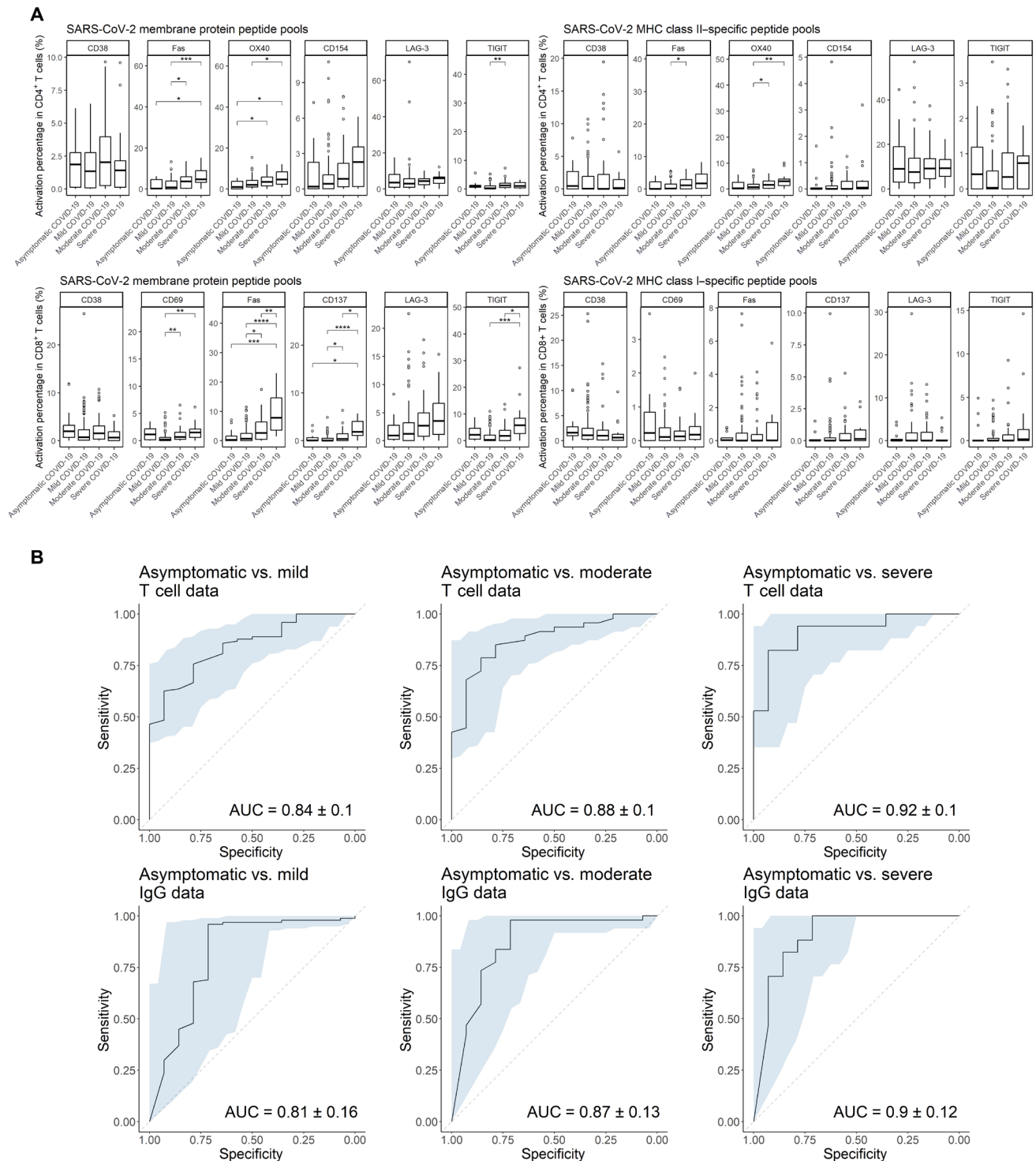


Fig. 2. Comparison between asymptomatic, mild, moderate, and severe patients. (A) Net percentages of T cells with positive expression of functional markers out of the total numbers of CD4⁺ and CD8⁺ T cells poststimulation (normalized with unstimulated controls). Horizontal lines represent median values of each group. Statistical significance was analyzed by Kruskal-Wallis test with Bonferroni correction. * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$. **(B)** ROCs and AUCs from leave-one-out cross-validation of random forest classifiers using flow cytometry T cell data and IgG data.

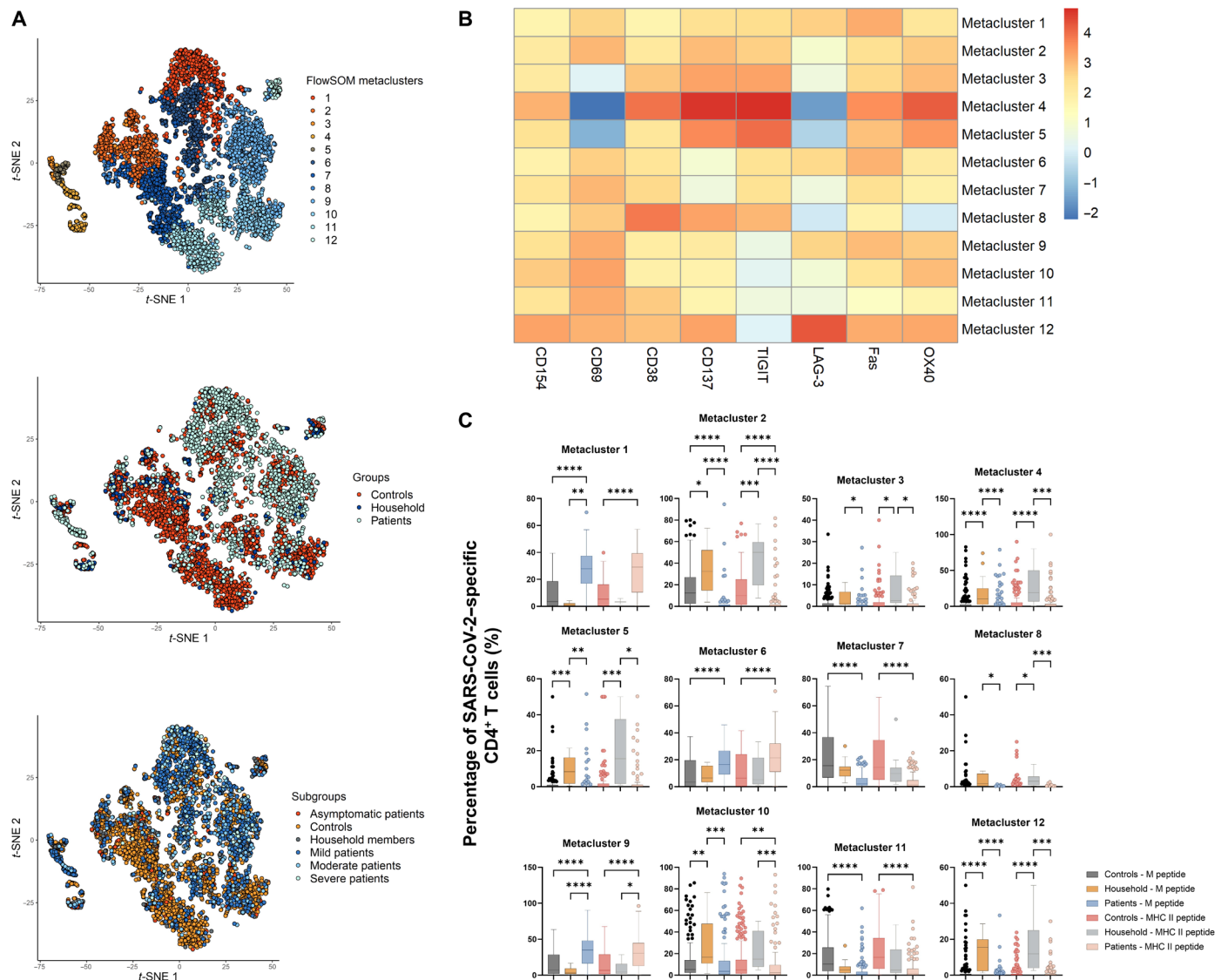


Fig. 3. FlowSOM clustering applied on CD4⁺ T cells of controls, patients, and household members. (A) t-SNE representation of 12 metaclusters projected by cohort groups and subgroups. (B) Mean fluorescence intensity of individual activation marker in each metacluster. (C) Percentage of SARS-CoV-2-specific CD4⁺ T cells in each metacluster. Horizontal lines represent median values of each group. Statistical significance was analyzed by Kruskal-Wallis test with Dunn's correction. * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$. In agreement with the uniparameter approach (Fig. 1A). (B) and (C) revealed that CD4⁺ T cells in patients had high Fas expression (corresponding to metaclusters 1, 6, and 9), while CD4⁺ T cells in household members overexpressed CD154, OX40, CD38, Fas, and TIGIT (corresponding to metaclusters 2, 4, 5, 10, and 12).

having worked for minimal 4 weeks on COVID-19 wards more than 3 months before sampling) who were very likely exposed to SARS-CoV-2 before recruitment (group 1: PCR negative or not tested, IgG negative, without fever; group 2: PCR negative or not tested, IgG negative, with fever; group 3: PCR negative or not tested, IgG positive, without fever; group 4: PCR negative or not tested, IgG positive, with fever; and group 5: PCR positive). Groups 1 and 2 (both having negative PCR and serology tests) were not distinguishable from each other by their SARS-CoV-2 RBD, NP, and S1-S2-specific IgG reactivity (fig. S9) but showed substantial differences in their T cell metacluster percentages (Fig. 5). It was evident in Fig. 5A that, among SARS-CoV-2-specific CD4⁺ T cells, group 1 had a higher percentage of

metacluster 3 than group 2 ($P \leq 0.05$), which represented a high expression of TIGIT and OX40. In contrast, group 2 had higher percentages of metaclusters 6 and 9 than group 1 ($P \leq 0.05$), which highlights the difference in Fas expression between these two groups. Figure 5B reveals that, among SARS-CoV-2-specific CD8⁺ T cells, group 1 surpassed group 2 in the percentages of metacluster 1 ($P \leq 0.05$, stimulated with membrane peptide pools) and metacluster 2 ($P \leq 0.001$, stimulated with MHC class I peptide pools), which represented cells with high expression of CD69, CD38, and CD137. In contrast, group 2 surpassed group 1 in the percentages of metaclusters 5 and 7 ($P \leq 0.05$), which contained cells with elevated levels of CD69 and Fas. This highlights that, despite having no confirmed SARS-

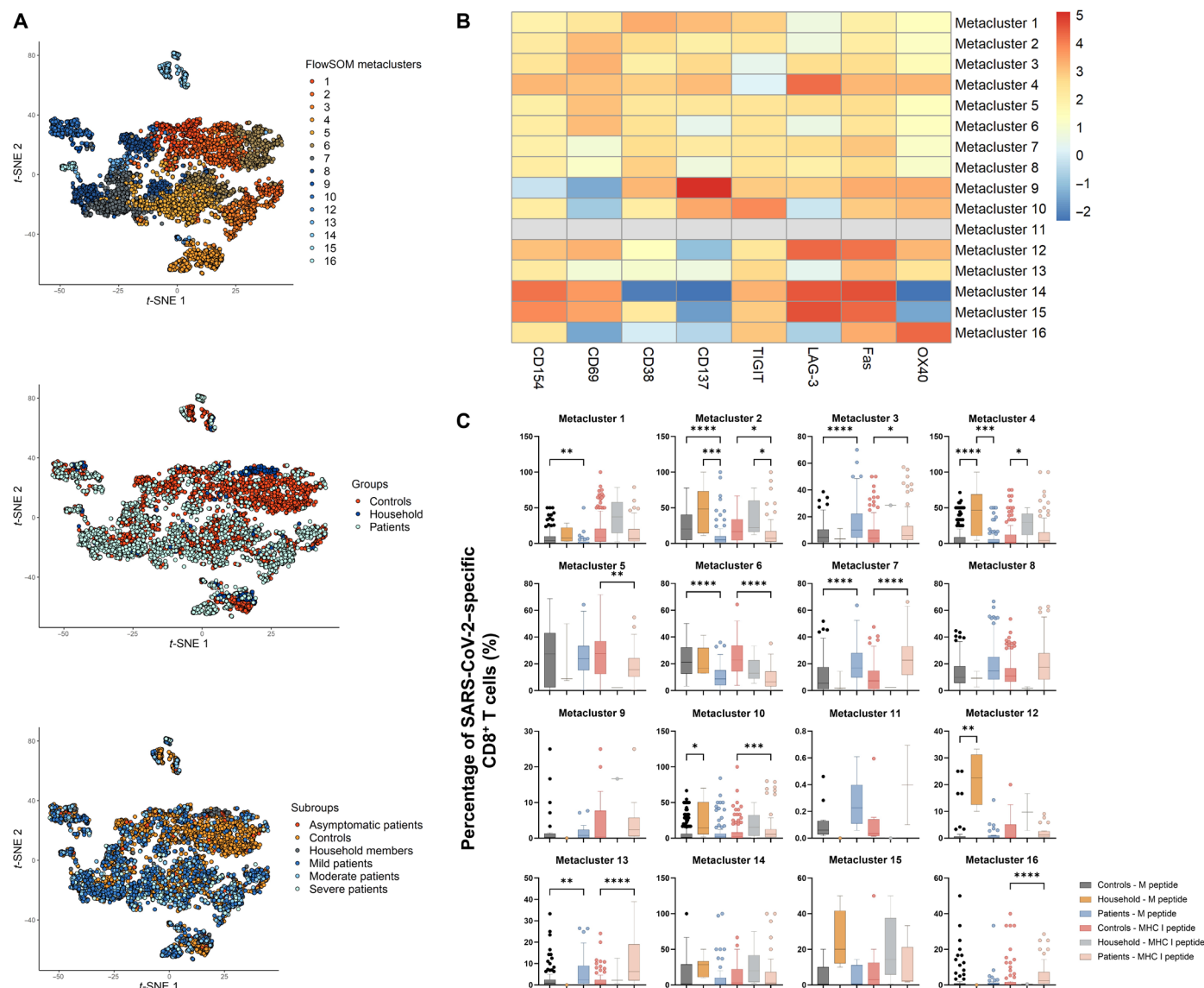


Fig. 4. FlowSOM clustering applied on CD8⁺ T cells of controls, patients, and household members. (A) t-SNE representation of 16 metaclusters projected by cohort groups and subgroups. (B) Mean fluorescence intensity of individual activation marker in each metacluster. (C) Percentage of SARS-CoV-2-specific CD8⁺ T cells in each metacluster. Horizontal lines represent median values of each group. Statistical significance was analyzed by Kruskal-Wallis test with Dunn's correction. * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$. In agreement with the uniparameter approach (Fig. 1A). (B) and (C) highlighted that CD8⁺ T cells in patients had elevated levels of Fas and CD69 (i.e., metaclusters 3, 7, and 13), while CD8⁺ T cells in household members overexpressed CD137, CD69, LAG-3, Fas, and TIGIT (i.e., metaclusters 4, 10, and 12).

CoV-2-positive tests (PCR and/or serology), health care providers in groups 1 and 2 had CD4⁺ and CD8⁺ T cells that showed signs of previous SARS-CoV-2 encounters, which was only reflected in their T cell signatures but not in their IgG reactivity. T cells of those in group 1 displayed SARS-CoV-2-specific activation and exhaustion, whereas T cells of those in group 2 displayed apoptosis susceptibility. Group 5, despite having tested positive by PCR, could not be distinguished from groups 3 and 4 (both having PCR negative or not taken) on the basis of their IgG data, but they were differentiated on the basis of Fas and CD137 expression in CD8⁺ T cells. Cellulepidemiology thus offers a unique approach to distinguish frequently exposed individuals from each other, which would not have been possible if only serology was used.

Patients with long Covid displayed elevated humoral and cellular immune responses against SARS-CoV-2

Following recovery, several patients reported suffering from the post-acute sequelae of COVID-19, also referred to as “long Covid,” which concerns a plethora of persistent symptoms (27, 28). In our current data, there were 79 (of 166) recovered patients with long Covid (potentially indicating a recruitment bias), 40 of which were initially classified as mildly ill patients during the acute phase, 28 were moderately ill, and 11 were severely ill (table S4). Their long Covid symptoms commonly included fatigue, shortness of breath, persistent loss of taste and smell, muscle ache, and chest pressure. Some symptoms were relatively peculiar, such as liver inflammation, diminished eyesight, reduced or altered hearing, hormonal

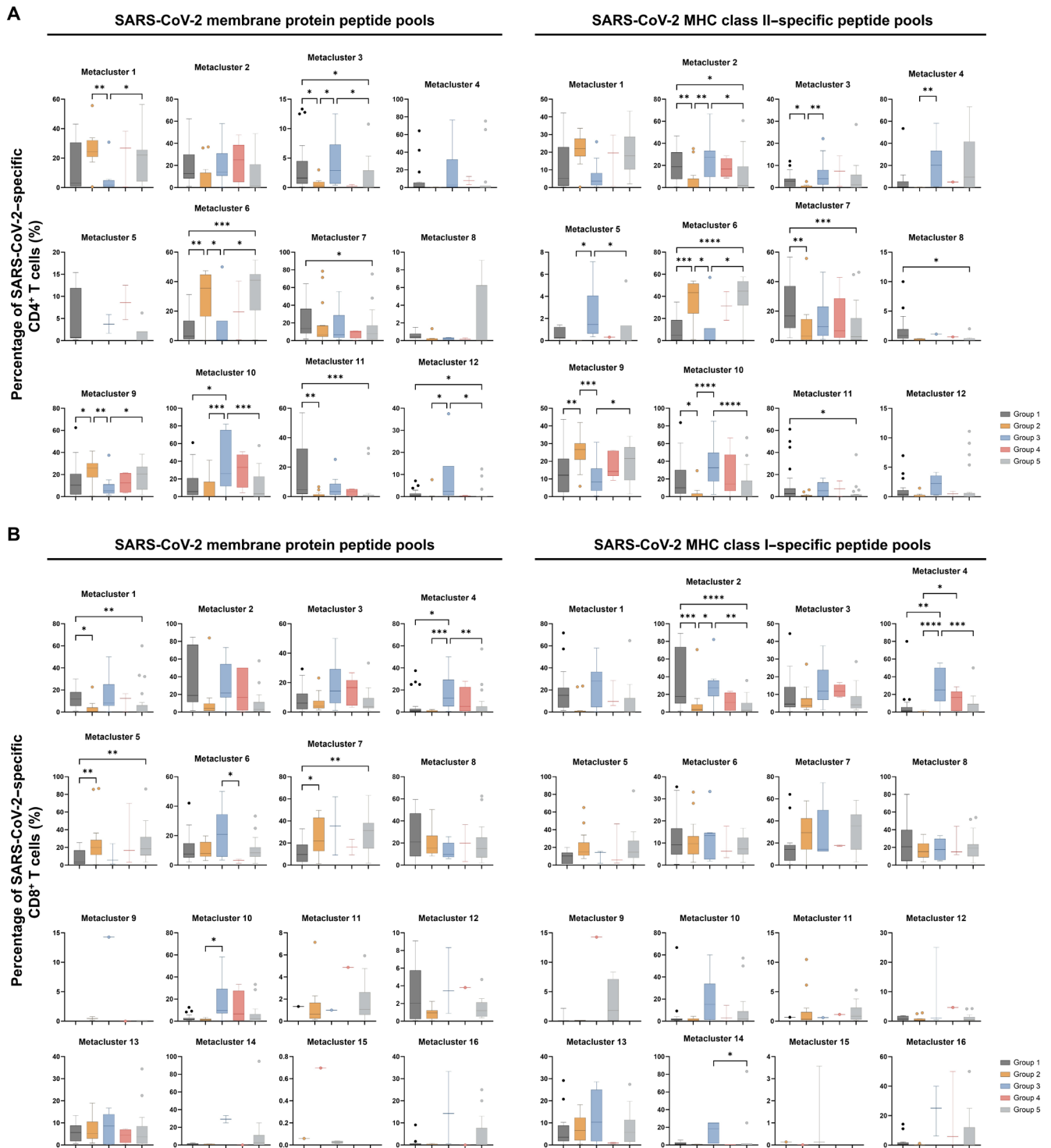


Fig. 5. Percentage of SARS-CoV-2-specific T cells in health care workers in each metacluster of the FlowSOM model developed based on controls', patients', and household members' data. (A) CD4⁺ T cells based on the FlowSOM model in Fig. 3. (B) CD8⁺ T cells based on the FlowSOM model in Fig. 4. Horizontal lines represent median values of each group. Statistical significance was analyzed by Kruskal-Wallis test with Dunn's correction. * $P \leq 0.05$; ** $P \leq 0.01$; * $P \leq 0.001$; **** $P \leq 0.0001$.**

unbalance, cognitive dysfunction (including memory decline and concentration difficulties), palpitations, and limited lung capacity (one severe patient was listed for double lung transplant). These patients with long Covid had higher IgG reactivity to RBD, NP, and S1-S2 antigens and more up-regulated expression of Fas, OX40, CD154, LAG-3, TIGIT in CD4⁺ T cells and CD69, Fas, CD137, LAG-3 in CD8⁺ T cells compared to uninfected controls, asymptomatic patients, and patients with mild and moderate COVID-19 (figs. S6 and S7).

DISCUSSION

In classical epidemiology, the incidence, prevalence, and distribution of infectious diseases are analyzed at a population level. It offers data to study the effectiveness of intervention strategies such as lockdowns, vaccination, and screening. One of the most common methods to estimate these quantities uses the serological footprint of an infection on the host. Seroepidemiological studies document the presence and kinetics of infection- or vaccine-induced antibodies. Studying of serological profiles in a population as a function of age and time and comparing these profiles between different regions have delivered important insights into the infectivity of pathogens, the identification of risk groups, the transmission routes of infections, the impact of immunization programs, and basic pathophysiological mechanisms. Although seroepidemiology has been a key tool to define successful public health interventions worldwide, it is fundamentally limited due to its use of antibodies as the sole marker of infectious disease immunity.

With the concept of celluloeidemiology, we would like to introduce an approach of studying T cell response against pathogens on a population level, thereby allowing more profound insights when analyzing immune responses from large (potentially not annotated) study groups. We consider celluloeidemiology as an added value to seroepidemiology thanks to the existence of T cell phenotypes and the expression of a diversity of proteins detectable on the T cell surface, thus allowing multidimensional T cell characterization. In addition, the extremely broad repertoire of TCRs, which can simultaneously recognize and memorize distinct pathogenic antigens, has recently been proven to be highly sensitive and specific in the detection of past SARS-CoV-2 infection (29). This implies that for large-scale modeling of infection morbidity, adding data on T cell reactivity will complement serology and augment the value of the models.

Another limitation of conventional seroepidemiology is that antibody titers declined over time after COVID-19 clearance (30). Cellular immunity against SARS-CoV-2, however, was found to decay more slowly over time than neutralizing antibody titers (31). A report on patients who had recovered from SARS-CoV-1 indicated that cellular immune responses were maintained for nearly two decades, whereas memory B cells and antibody responses could not be detected in most individuals at that point (32). Studies have shown that SARS-CoV-2-specific cellular immune responses remained detectable in recovered patients 8 months after infection, although higher in CD4⁺ T cells than in CD8⁺ T cells (33), which agrees with our data. Cohen *et al.* (34) observed that SARS-CoV-2-specific antibody responses and T cell activities might persist for 8 months following infection, which also confirms our findings.

Here, we applied the proposed celluloeidemiology paradigm to characterize SARS-CoV-2-specific T cell responses in patients recovered from COVID-19 for over 3 months and compared them to unexposed controls, household members of the patients who recovered from COVID-19, general practitioners, hospital health care workers,

and pre-COVID controls. Using experimental and computational approaches, we showed that the number of CD4⁺ and CD8⁺ T cells with positive expression of functional markers, after stimulation with SARS-CoV-2 proteins, could be used to distinguish patients who recovered from COVID-19 from controls and pre-COVID donors, as well as differentiate asymptomatic patients from mild, moderate, and severe patients. To achieve this, we used automated T cell clustering algorithms like FlowSOM (35) and random forest (36) to maximize output from our data. Applying machine learning algorithms on three-dimensional Luminex-based serological data allowed previous SARS-CoV-2 infection classification with more than 0.9 AUC scores.

It is evident from the data presented here that T cell assessment on a population level generates insights about an infection during certain conditions that would, otherwise, not have been feasible with conventional serology. For instance, among the health care workers (the cohort representing high exposure to SARS-CoV-2), groups having negative PCR and serological tests were not distinguishable by their SARS-CoV-2 RBD, NP, and S1-S2 antigen-specific IgG reactivity but showed substantial differences in their T cell metacluster percentages. Health care workers having tested positive by PCR could not be distinguished from their PCR-negative IgG-positive counterparts on the basis of SARS-CoV-2-specific IgG data only, but our celluloeidemiology approach could differentiate these groups. Celluloeidemiology showed also to be uniquely capable to identify “unexposed controls” to actually be “asymptomatic” patients.

In our TCR repertoire analysis, we found that SARS-CoV-2 unique and coronavirus-specific TCRs were present in patients who recovered from COVID-19 more than controls. This highlights that a TCR module in the celluloeidemiology paradigm is useful for differentiating between different infections and backward tracing of pathogen-infected patients thanks to the added layer of TCR specificity, which cannot always be obtained from T cell phenotypes alone. Within the population, there remains residual T cell immune response to other coronaviruses, which bear moderate amino acid conservation with SARS-CoV-2 (37). Thus, epitopes recognized by T cells are likely shared between these viruses, leading to this cross-reactivity which has been observed in some studies (17, 38, 39).

Despite the noteworthy findings, our study has several limitations. One of them is that pre-COVID donors were possibly not the best controls to compare with other cohorts due to the different PBMC isolation methods (i.e., Ficoll was used for pre-COVID PBMC, whereas SepMate tubes and Lymphoprep were used for PBMCs of the other cohorts with distinct centrifugation time and speed in the two protocols). This might have an impact on the composition, quantity, and potentially functional activities of the isolated cells (40). In addition, the percentage of individuals with long Covid was notably high compared to other published cohorts (41–43), which might be due to a selection bias that might have occurred in this retrospective recruitment design. Recruitment via general practitioners and hospitals likely excludes patients with no symptoms (although these will likely be among the household members) and patients with long Covid were likely more motivated to participate. Nevertheless, such selection bias has no impact on the feasibility of the paper or its objectives. Another point that is worth discussing is the feasibility and scalability of celluloeidemiology because such intricate T cell information requires substantial (even labor-intensive) data collection and analysis. This is one of the reasons why, despite all the important insights and benefits that celluloeidemiology can offer, seroepidemiology likely remains more feasible for many studies.

In conclusion, we believe that the proposed celluloepidemiology paradigm is complementary to conventional seroepidemiology in offering high dimensionality, sensitivity, and deep insight into the heterogeneity of human immune response against pathogens. Although celluloepidemiology was introduced and evaluated using the SARS-CoV-2 model in this study, the same concept is highly translatable to different pathogens and infectious diseases in the future.

METHODS

Participant recruitment

A total of 582 participants between 18 and 85 years of age were recruited in Belgium from August 2020 until April 2021 via general practitioners, hospitals, and social media. The dominant SARS-CoV-2 strain (in Belgium) was the original Wuhan strain. None of the participants was vaccinated against SARS-CoV-2. Participants in this study were divided into five categories: (i) patients who recovered from COVID-19 ($n = 168$), (ii) household members of the recovered patients ($n = 27$), (iii) controls ($n = 259$), (iv) general practitioners ($n = 37$), and (v) hospital health care workers ($n = 91$). Clinical/demographic details of all participants can be found in table S5. This study was approved by the Antwerp University Hospital IRB (reference number 20/02/003).

Patients who recovered from COVID-19 (referred to as patients in this study) were participants who had recovered from COVID-19 more than 3 months before enrollment and were screened before blood draws to make sure that they were symptom free and in a recovered phase. On their visits, patients were asked to provide proof of positive testing for SARS-CoV-2, either via PCR or serological IgG testing and fever of 38°C or higher without other proven explanations, more than 3 months before recruitment. Patient classification criteria and disease severity (i.e., mild, moderate, and severe infection) were consistent with the COVID-19 case definitions from the US National Health Institute (44), World Health Organization (45, 46), and European Centre for Disease Prevention and Control (47). Mild cases reported various COVID-19–related manifestations (e.g., fatigue, fever, cough with or without sputum production, anorexia, malaise, myalgia, sore throat, dyspnea, nasal congestion, and headache; rarely diarrhea, nausea, and vomiting) that did not require hospitalization. Moderate cases showed evidence of lower respiratory tract disease, required outpatient hospital visits, and had oxygen saturation $\geq 94\%$ on room air at sea level. Severe cases required hospitalization and had oxygen saturation $< 94\%$ on room air at sea level, a ratio of arterial partial pressure of oxygen to fraction of inspired oxygen < 300 mmHg, respiratory frequency > 30 breaths per minute, or lung infiltrates $> 50\%$. Within the patient cohort, there were also “long Covid” cases who reported suffering from the post-acute sequelae of COVID-19 for more than 3 months after recovery (27, 28). Household members were participants who lived in the same household with proven SARS-CoV-2 patients (who tested positive for SARS-CoV-2 by either PCR or IgG and had fever more than 3 months before). Household members did not have a confirmed SARS-CoV-2 infection, either by PCR or IgG.

General practitioners and hospital health care workers (general practitioners and hospital personnel having worked for minimal 4 weeks on COVID-19 wards more than 3 months before sampling) were recruited into five subgroups according to their PCR and IgG test results as well as their symptoms:

- 1) Group 1: PCR negative or not tested, IgG negative, without fever (no indication of SARS-CoV-2 infection since having contact with patients with COVID-19 more than 3 months before blood donation)
- 2) Group 2: PCR negative or not tested, IgG negative, with fever
- 3) Group 3: PCR negative or not tested, IgG positive, without fever
- 4) Group 4: PCR negative or not tested, IgG positive, with fever
- 5) Group 5: PCR positive (certain SARS-CoV-2 infection regardless of IgG test results)

Controls were participants who were considered healthy and had no known history of any systemic diseases, including, but not limited to, autoimmune disease, diabetes, kidney or liver disease, congestive heart failure, malignancy, coagulopathy, hepatitis B or C, or HIV. These participants did not have any known household exposure to SARS-CoV-2 nor positive testing (PCR/IgG). Also, included in this study were 24 pre-COVID donors, whose blood was collected for our previous studies between 2015 and 2018, which would be used as unexposed controls, given that SARS-CoV-2 emerged in late 2019.

Blood collection, processing, and storage

Whole blood from all participants was collected in lithium-heparin and serum separator Vacutainer tubes (BD, catalog no. 367526, 366444) and processed within 6 hours. All blood samples were pseudonymized. To isolate the PBMCs, whole blood in lithium-heparin tubes was diluted 1:2 with phosphate-buffered saline (PBS) (Thermo Fisher Scientific, catalog no. 14190250) and layered into SepMate-50 tubes (STEMCELL Technologies Inc., catalog no. 85460) preloaded with Lymphoprep (STEMCELL Technologies Inc., catalog no. 07861). The SepMate tubes with layered blood and Lymphoprep were spun for 10 min at 1200g at room temperature, and the PBMCs were then harvested by collecting the supernatant and washing with PBS. The harvested PBMCs were aliquoted into cryovials in fetal bovine serum (FBS) (Thermo Fisher Scientific, catalog no. 10270-106) with 10% dimethyl sulfoxide (Merck Life Science, catalog no. D2650) and stored in liquid nitrogen until use. PBMCs from pre-COVID controls were processed like described previously (48).

Blood in serum separator tubes was allowed to clot for 30 to 60 min and then centrifuged for 10 min at 1200g at room temperature. After centrifugation, serum located above the polymer barrier was transferred to cryovials and stored at -80°C for Luminex assays.

Marker selection with mass cytometry

Mass cytometry, or CyTOF, was performed using an extensive panel of isotope-tagged antibodies for T cells to select markers that could display the most distinction in T cell activation between controls and patients for the next flow cytometry measurements. Only PBMCs from 45 donors were used in the first CyTOF runs including 30 patients and 15 pre-COVID donors (table S6).

Cryopreserved PBMCs were thawed at 37°C in a water bath and then diluted in pre-warmed AIM-V medium (Thermo Fisher Scientific, catalog no. 12055091) with 5% FBS and subsequently centrifuged for 5 min at 300g at room temperature. After the supernatant was discarded, ~ 1 to 2 million thawed PBMCs were used for each stimulating condition: (i) stimulated with SARS-CoV-2 membrane protein (Prot_M) PepTivator (Miltenyi Biotec, catalog no. 130-126-702) at 2 $\mu\text{g}/\text{ml}$, (ii) stimulated with MHC class I–specific pool, and (iii) MHC class II–specific pool (JPT Peptide Technologies, custom

synthesized) also at 2 µg/ml. Unstimulated samples were used as negative controls. Samples stimulated with phytohaemagglutinin (PHA) at 5 µg/ml were used as positive controls. The custom-synthesized peptides were based on a curated list of known SARS-CoV-2 epitopes, extracted from various existing studies (49–53) using the Enzyme-Linked Immunosorbent assay (ELISpot), in silico prioritization and TCR-epitope simulation. Those epitopes that were most consistently reported as immunogenic across different studies were retained. In addition, these epitopes were compared to a list of 119 Nidovirales genomes from the Corona OMA Orthology Database (54) to identify which were SARS-CoV-2 specific. A list of these peptides and their amino acid sequences is displayed in table S7. CD40 (Miltenyi Biotec, catalog no. 130-094-133) and CD28 (BD, catalog no. 555725) were both added at 1 µg/ml to each culture, and PBMCs in all four stimulating conditions were incubated for 16 hours at 37°C. After incubation, staining started with CD45 barcodes (i.e., CD45 antibodies conjugated with ¹⁰⁶Cd, ¹¹⁰Cd, ¹¹¹Cd, ¹¹³Cd, ¹¹⁴Cd, and ¹¹⁶Cd) for 30 min, then isotope-tagged antibodies (table S1) for another 30 min, and ¹⁰³Rh-intercalator (Standard BioTools Inc., catalog no. 201103A) for 15 min. Cells were subsequently fixed for 10 min in 4% paraformaldehyde (Thermo Fisher Scientific, catalog no. 28908) before being stained with Ir-intercalator (Standard BioTools Inc., catalog no. 201192A) overnight at 4°C. Measurements were performed on the Helios instrument (Standard BioTools Inc.) at the KU Leuven Flow and Mass Cytometry Facility.

CyTOF output data were normalized and debarcoded in the CyTOF software (Standard BioTools Inc.). Then, FlowJo (FlowJo LLC.) was used for data gating (gating strategies are displayed in fig. S10). Total numbers of CD4⁺ and CD8⁺ T cells as well as numbers of T cells with high expression of activation markers (i.e., CD38, OX40, TIGIT, PD-1, CTLA-4, CD69, Fas, CD154, TIM-3, CD57, HLA-DR, LAG-3, CD127, and CD137) in stimulated samples relative to T cells' expression in unstimulated controls were then exported from FlowJo, and percentages of antigen-specific T cells were analyzed in R. Boruta package (25), a feature selection algorithm, was used to find the optimal markers that could distinguish patients with COVID-19 from pre-COVID participants and differentiate patient subgroups (i.e., mild, moderate, and severe).

Flow cytometry measurements

In the first set of experiments, PBMCs from controls, patients, household members, hospital health care workers, general practitioners, and pre-COVID donors were used for flow cytometry measurements. Cells were thawed and stimulated as previously described for CyTOF. After stimulation, cells were stained for 30 min with a panel of 17 fluorescent antibodies (table S2) in Brilliant Stain Buffer (BD, catalog no. 563794) before being analyzed on the NovoCyte Quanteon flow cytometer (Agilent Technologies Inc.). In the second set of experiments, only PBMCs from controls, patients, and household members were analyzed. Cells were stimulated with a combination of MHC class I- and class II-specific pool (JPT Peptide Technologies, custom synthesized) at 3 µg/ml. Unstimulated samples were used as negative controls. Samples stimulated with PHA at 5 µg/ml were used as positive controls. After stimulation, cells were prepared following the overnight staining protocol (55). Briefly, cells were first stained with the fixable viability stain for 15 min. Then, they were fixed for 10 min in 0.4% paraformaldehyde (BioLegend, catalog no. 420801) at room temperature. Afterward, they were stained with a panel of 12

fluorescent antibodies (table S3) overnight at 4°C. Measurements were also performed on the NovoCyte Quanteon flow cytometer.

Data compensation, transformation, and automated gating were performed using R packages flowCore, flowWorkspace, and openCyto. FlowSOM was performed following the authors' instructions (35) to cluster cells for immunophenotyping. Like previous CyTOF data analysis, total numbers of CD4⁺ and CD8⁺ T cells as well as numbers of T cells with high expression of functional markers (i.e., CD38, CD69, Fas, OX40, CD137, CD154, LAG-3, and TIGIT) in stimulated samples relative to T cells' expression in unstimulated controls were exported (figs. S11 and S12). The percentages of antigen-specific T cells were used to develop classification models. The data were split into training data and held-out testing data. Random forest classifiers (36) were trained on the basis of the training data using the R package randomForest. Leave-one-out cross-validation was used to evaluate classifiers' performance. Receiver operating characteristic (ROC) curves and the AUC were used using the R package pROC (56) to evaluate the classifier performance. AUC values were displayed as means ± confidence interval (95%).

Luminex assay

Luminex assays were performed on serum isolated from blood in BD Vacutainer serum separator tubes following the same procedure as in a previous publication (6). The lyophilized recombinant NP, RBD, and spike (S1-S2) antigens (Sino Biological Inc.) were resuspended in a buffer according to the manufacturer's instructions and stored until use. One microgram of RBD and S1-S2 and 2 µg of NP were coupled to 1.25 × 10⁶ paramagnetic MAGPLEX COOH-microsphere beads (Luminex Corporation) for IgG detection. Bead working solution was prepared in a mixture of PBS and bovine serum albumin (BSA) with a concentration of 2000 beads per antigen per well. Serum samples were diluted 1:300 and added to each well together with the bead working solution in a final volume of 150 µl. Plates were incubated at room temperature for 2 hours in the dark and then washed with 200 µl per well of PBS-BSA buffer. After incubation with biotin-labeled anti-human IgG secondary antibody and streptavidin-R-phycoerythrin conjugate, reactions were read on a Luminex 100/200 analyzer (Luminex Corporation). Results were expressed as median fluorescent intensities and normalized for further statistical analyses. When considering combinations of (RBD-NP and S1S2 antigens), the test has the following sensitivities at the 99% specificity cutoff: severe-recent (95%), recent-mild (99%), and old-mild (96%).

T cell receptor sequencing

TCR-seq was performed on enriched CD8⁺ T cells. Cryopreserved PBMCs were thawed at 37°C in a water bath, then diluted in pre-warmed AIM-V medium supplemented with 5% FBS, and subsequently centrifuged for 5 min at 300g at room temperature. After the supernatant was discarded, cells were washed with precooled MACS Separation Buffer (Miltenyi Biotec, catalog no. 130-091-221) and centrifuged for 5 min at 300g, after which the supernatant was discarded. Magnetic cell separation was then performed using MS columns, OctoMACS Separator, and CD8 anti-human MicroBeads (Miltenyi Biotec, catalog nos. 130-042-201, 130-042-108, and 130-045-201, respectively) following the supplier's instructions. After enrichment, ~200,000 CD8⁺ T cells were used for RNA extraction via the Quick-RNA Microprep Kit (Zymo Research, catalog no.

R1050). TCR libraries were prepared from the extracted RNA using the QIAseq Immune Repertoire RNA Library Kit (QIAGEN, catalog no. 333705) following the supplier's instructions. The amplified TCR libraries were equimolarly pooled and sequenced on NextSeq 500 (Illumina).

TCR data analysis

TCR clonotype annotation was performed using MiXCR v3.0.13 with the default input parameters (57). Epitope specificity of the TCRs was predicted with TCRex (58). SARS-CoV-2 epitopes were considered coronavirus common if they were found in protein sequence data of at least 2 of the 119 *Nidovirales* species as had been previously reported (59). Clustering of TCRs with similar CDR3 sequences was performed for CD8⁺ TCR repertoire of every individual separately using ClusTCR v1.0.2 with default parameters (60). Clusters that contained TCRs, for which specificity had been identified with TCRex, were annotated with specificity of those TCRs. Only clusters containing TCRs with the same predicted (or unknown) specificity were retained for further analysis. For the remaining clusters, all TCRs within one cluster were considered to have the same specificity. After clustering, counts of annotated TCRs and clusters were compared between different groups.

Statistical analyses

The two-sided Fisher's exact test was performed in R to evaluate the significance of T cells' response to membrane (M) and MHC class I- and class II-specific peptides compared to unstimulated T cells. Statistical significance analyses and multiple comparison correction were done in R or GraphPad Prism (GraphPad Software Inc.). Two-group comparison was done with Mann-Whitney test. Multiple groups were compared by Kruskal-Wallis test. These symbols indicate *P* values of statistical significance: **P* ≤ 0.05; ***P* ≤ 0.01; ****P* ≤ 0.001; *****P* ≤ 0.0001.

Supplementary Materials

The PDF file includes:

Figs. S1 to S13
Tables S1 to S3
Legends for tables S4 to S9

Other Supplementary Material for this manuscript includes the following:

Tables S4 to S9

REFERENCES AND NOTES

- N. Hens, M. Aerts, C. Faes, Z. Shkedy, O. Lejeune, P. Van Damme, P. Beutels, Seventy-five years of estimating the force of infection from current status data. *Epidemiol. Infect.* **138**, 802–812 (2010).
- F. A. Murphy, "Epidemiology of Human and Animal Viral Diseases" in *Encyclopedia of Virology* (Academic Press, 2008), pp. 140–148.
- Q. X. Long, B. Z. Liu, H. J. Deng, G. C. Wu, K. Deng, Y. K. Chen, P. Liao, J. F. Qiu, Y. Lin, X. F. Cai, D. Q. Wang, Y. Hu, J. H. Ren, N. Tang, Y. Y. Xu, L. H. Yu, Z. Mo, F. Gong, X. L. Zhang, W. G. Tian, L. Hu, X. X. Zhang, J. L. Xiang, H. X. Du, H. W. Liu, C. H. Lang, X. H. Luo, S. B. Wu, X. P. Cui, Z. Zhou, M. M. Zhu, J. Wang, C. J. Xue, X. F. Li, L. Wang, Z. J. Li, K. Wang, C. C. Niu, Q. J. Yang, X. J. Tang, Y. Zhang, X. M. Liu, J. J. Li, D. C. Zhang, F. Zhang, P. Liu, J. Yuan, Q. Li, J. L. Hu, J. Chen, A. L. Huang, Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat. Med.* **26**, 845–848 (2020).
- Q. X. Long, X. J. Tang, Q. L. Shi, Q. Li, H. J. Deng, J. Yuan, J. L. Hu, W. Xu, Y. Zhang, F. J. Lv, K. Su, F. Zhang, J. Gong, B. Wu, X. M. Liu, J. J. Li, J. F. Qiu, J. Chen, A. L. Huang, Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* **26**, 1200–1204 (2020).
- A. S. Iyer, F. K. Jones, A. Nodoushani, M. Kelly, M. Becker, D. Slater, R. Mills, E. Teng, M. Kamruzzaman, W. F. Garcia-Beltran, M. Astudillo, D. Yang, T. E. Miller, E. Oliver, S. Fischinger, C. Atyeo, A. J. Iafate, S. B. Calderwood, S. A. Lauer, J. Yu, Z. Li, J. Feldman, B. M. Hauser, T. M. Caradonna, J. A. Branda, S. E. Turbett, R. C. LaRocque, G. Mellon, D. H. Barouch, A. G. Schmidt, A. S. Azman, G. Alter, E. T. Ryan, J. B. Harris, R. C. Charles, Persistence and decay of human antibody responses to the receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. *Sci. Immunol.* **5**, eabe0367 (2020).
- J. Mariën, A. Ceulemans, J. Michiels, L. Heyndrickx, K. Kerkhof, N. Fogue, M.-A. Widdowson, L. Mortgat, E. Duysburgh, I. Desombere, H. Jansens, M. Van Esbroeck, K. K. Ariën, Evaluating SARS-CoV-2 spike and nucleocapsid proteins as targets for antibody detection in severe and mild COVID-19 cases using a Luminex bead-based assay. *J. Virol. Methods* **288**, 114025 (2021).
- P. Kellam, W. Barclay, The dynamics of humoral immune responses following SARS-CoV-2 infection and the potential for reinfection. *J. Gen. Virol.* **101**, 791–797 (2020).
- L. Kuri-Cervantes, M. B. Pampena, W. Meng, A. M. Rosenfeld, C. A. G. Ittner, A. R. Weisman, R. S. Agyekum, D. Mathew, A. E. Baxter, L. A. Vella, O. Kuthuru, S. A. Apostolidis, L. Bershaw, J. Dougherty, A. R. Greenplate, A. Pattekar, J. Kim, N. Han, S. Gouma, M. E. Weirick, C. P. Arevalo, M. J. Bolton, E. C. Goodwin, E. M. Anderson, S. E. Hensley, T. K. Jones, N. S. Mangalmurti, E. T. Luning Prak, E. J. Wherry, N. J. Meyer, M. R. Betts, Comprehensive mapping of immune perturbations associated with severe COVID-19. *Sci. Immunol.* **5**, eabd7114 (2020).
- D. Mathew, J. R. Giles, A. E. Baxter, D. A. Oldridge, A. R. Greenplate, J. E. Wu, C. Alanio, L. Kuri-Cervantes, M. B. Pampena, K. D'Andrea, S. Manne, Z. Chen, Y. J. Huang, J. P. Reilly, A. R. Weisman, C. A. G. Ittner, O. Kuthuru, J. Dougherty, K. Nzingha, N. Han, J. Kim, A. Pattekar, E. C. Goodwin, E. M. Anderson, M. E. Weirick, S. Gouma, C. P. Arevalo, M. J. Bolton, F. Chen, S. F. Lacey, H. Ramage, S. Cherry, S. E. Hensley, S. A. Apostolidis, A. C. Huang, L. A. Vella, UPenn COVID Processing Unit, M. R. Betts, N. J. Meyer, E. J. Wherry, Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* **369**, eabc8511 (2020).
- B. Diao, C. Wang, Y. Tan, X. Chen, Y. Liu, L. Ning, L. Chen, M. Li, Y. Liu, G. Wang, Z. Yuan, Z. Feng, Y. Zhang, Y. Wu, Y. Chen, Reduction and functional exhaustion of T cells in patients with coronavirus disease 2019 (COVID-19). *Front. Immunol.* **11**, 827 (2020).
- M. Zheng, Y. Gao, G. Wang, G. Song, S. Liu, D. Sun, Y. Xu, Z. Tian, Functional exhaustion of antiviral lymphocytes in COVID-19 patients. *Cell. Mol. Immunol.* **17**, 533–535 (2020).
- H. Y. Zheng, M. Zhang, C. X. Yang, N. Zhang, X. C. Wang, X. P. Yang, X. Q. Dong, Y. T. Zheng, Elevated exhaustion levels and reduced functional diversity of T cells in peripheral blood may predict severe progression in COVID-19 patients. *Cell. Mol. Immunol.* **17**, 541–543 (2020).
- I. Thevarajan, T. H. O. Nguyen, M. Koutsakos, J. Druce, L. Caly, C. E. van de Sandt, X. Jia, S. Nicholson, M. Catton, B. Cowie, S. Y. C. Tong, S. R. Lewin, K. Kedzierska, Breadth of concomitant immune responses prior to patient recovery: A case report of non-severe COVID-19. *Nat. Med.* **26**, 453–455 (2020).
- S. De Biasi, M. Meschiari, L. Gibellini, C. Bellinazzi, B. Borella, L. Fidanza, L. Gozzi, A. Iannone, D. Lo Tartaro, M. Mattioli, A. Paolini, M. Menozzi, J. Milić, G. Franceschi, R. Fantini, R. Tonelli, M. Sita, M. Sarti, T. Trenti, L. Brugioni, L. Cicchetti, F. Facchinetti, A. Pietrangelo, E. Cini, M. Girardis, G. Guaraldi, C. Mussini, A. Cossarizza, Marked T cell activation, senescence, exhaustion and skewing towards TH17 in patients with COVID-19 pneumonia. *Nat. Commun.* **11**, 434 (2020).
- W. Wang, B. Su, L. Pang, L. Qiao, Y. Feng, Y. Ouyang, X. Guo, H. Shi, F. Wei, X. Su, J. Yin, R. Jin, D. Chen, High-dimensional immune profiling by mass cytometry revealed immunosuppression and dysfunction of immunity in COVID-19 patients. *Cell. Mol. Immunol.* **17**, 650–652 (2020).
- J. Neidleman, X. Luo, J. Frouard, G. Xie, G. Gill, E. S. Stein, M. McGregor, T. Ma, A. F. George, A. Kusters, W. C. Greene, J. Vasquez, E. Ghosn, S. Lee, N. R. Roan, 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).
- A. Grifoni, D. Weiskopf, S. I. Ramirez, J. Mateus, J. M. Dan, C. R. Moderbacher, S. A. Rawlings, A. Sutherland, L. Premkumar, R. S. Jodi, D. Marrama, A. M. de Silva, A. Frazier, A. F. Carlin, J. A. Greenbaum, B. Peters, F. Krammer, D. M. Smith, S. Crotty, A. Sette, Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489–1501.e15 (2020).
- L. Ni, F. Ye, M. L. Cheng, Y. Feng, Y. Q. Deng, H. Zhao, P. Wei, J. Ge, M. Gou, X. Li, L. Sun, T. Cao, P. Wang, C. Zhou, R. Zhang, P. Liang, H. Guo, X. Wang, C. F. Qin, F. Chen, C. Dong, Detection of SARS-CoV-2-specific humoral and cellular immunity in COVID-19 convalescent individuals. *Immunity* **52**, 971–977.e3 (2020).
- T. Sekine, A. Perez-Potti, O. Rivera-Ballesteros, K. Strålin, J. B. Gorin, A. Olsson, S. Llewellyn-Lacey, H. Kamal, G. Bogdanovic, S. Muschiol, D. J. Wullmann, T. Kammann, J. Enggård, T. Parrot, E. Folkesson, M. Akber, L. Berglin, H. Bergsten, S. Brighenti, D. Brownlie, M. Butrym, B. Chambers, P. Chen, M. C. Jeannin, J. Grip, A. C. Gomez, L. Dillner, I. D. Lozano, M. Dzidic, M. F. Tullberg, A. Färnert, H. Glans, A. Haroun-Izquierdo, E. Henriksson, L. Hertwig, S. Kalsum, E. Kokkinou, E. Kvedaraitė, M. Loreti, M. Lourda, K. Maleki, K. J. Malmberg, N. Marquardt, C. Maucourant, J. Michaelsson, J. Mjösberg, K. Moll, J. Muva, J. Mårtensson, P. Naucclér, A. Norrby-Teglund, L. P. Medina, B. Persson, L. Radler, E. Ringqvist, J. T. Sandberg, E. Söhlberg, T. Soini, M. Svensson, J. Tynnell, R. Varnaite, A. Von Kries, C. Unge, O. Rooyackers, L. I. Eriksson, J. I. Henter, A. Sönnernberg,

- T. Allander, J. Albert, M. Nielsen, J. Klingström, S. Gredmark-Russ, N. K. Björkström, J. K. Sandberg, D. A. Price, H. G. Ljunggren, S. Aleman, M. Buggert, Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *Cell* **183**, 158–168.e14 (2020).
20. C. Thieme, M. Anft, K. Paniskaki, A. Blázquez Navarro, A. Doevelaar, F. S. Seibert, B. Hölzer, M. J. Konik, T. Brenner, C. Tempfer, C. Watzl, S. Dölff, U. Dittmer, O. Witzke, T. H. Westhoff, U. Stervbo, T. Roch, N. Babel, The SARS-CoV-2 T-cell immunity is directed against the spike, membrane, and nucleocapsid protein and associated with COVID 19 severity. *SSRN Electron. J.* 10.2139/ssrn.3606763 (2020).
 21. T. M. Snyder, R. M. Gittelman, M. Klinger, D. H. May, E. J. Osborne, R. Taniguchi, H. J. Zahid, I. M. Kaplan, J. N. Dines, M. N. Noakes, R. Pandya, X. Chen, S. Elasady, E. Svejnoha, P. Ebert, M. W. Pesesky, P. De Almeida, H. O'Donnell, G. DeGottardi, G. Keitany, J. Lu, A. Vong, R. Elyanow, P. Fields, J. Greissl, L. Baldo, S. Semprini, C. Cerchione, F. Nicolini, M. Mazza, O. M. Delmonte, K. Dobbs, R. Laguna-Goya, G. Carreño-Tarragona, S. Barrio, L. Imberti, A. Sottini, E. Quiros-Roldan, C. Rossi, A. Biondi, L. R. Bettini, M. D'Angio, P. Bonfanti, M. F. Tompkins, C. Alba, C. Dalgard, V. Sambri, G. Martinelli, J. D. Goldman, J. R. Heath, H. C. Su, L. D. Notarangelo, E. Paz-Artal, J. Martinez-Lopez, J. M. Carlson, H. S. Robins, Magnitude and dynamics of the T-cell response to SARS-CoV-2 infection at both individual and population levels. medRxiv 20165647 [Preprint] (2020).
 22. A. Vujkovic, M. Ha, T. de Block, L. van Petersen, I. Brosius, C. Theunissen, S. H. van Ierssel, E. Bartholomeus, W. Adriaensens, G. Vanham, G. Elias, P. Van Damme, V. Van Tendeloo, P. Beutels, M. van Frankenhuysen, E. Vlieghe, B. Ogunjimi, K. Laukens, P. Meysman, K. Vercauteren, Diagnosing viral infections through t-cell receptor sequencing of activated CD8⁺ T Cells. *J. Infect. Dis.* **229**, 507–516 (2024).
 23. L. H. Kuller, M. B. Bracken, S. Ogino, R. L. Prentice, R. P. Tracy, The role of epidemiology in the era of molecular epidemiology and genomics: Summary of the 2013 AJE-sponsored society of epidemiologic research symposium. *Am. J. Epidemiol.* **178**, 1350–1354 (2013).
 24. K. Yin, M. J. Peluso, X. Luo, R. Thomas, M. G. Shin, J. Neidleman, A. Andrew, K. C. Young, T. Ma, R. Hoh, K. Anglin, B. Huang, U. Argueta, M. Lopez, D. Valdivieso, K. Asare, T. M. Deveau, S. E. Munter, R. Ibrahim, L. Ständker, S. Lu, S. A. Goldberg, S. A. Lee, K. L. Lynch, J. D. Kelly, J. N. Martin, J. Münch, S. G. Deeks, T. J. Henrich, N. R. Roan, Long COVID manifests with T cell dysregulation, inflammation and an uncoordinated adaptive immune response to SARS-CoV-2. *Nat. Immunol.* **25**, 218–225 (2024).
 25. M. B. Kursa, W. R. Rudnicki, Feature selection with the Boruta package. *J. Stat. Softw.* **36**, 1–13 (2010).
 26. J. Neidleman, X. Luo, A. F. George, M. McGregor, J. Yang, C. Yun, V. Murray, G. Gill, W. C. Greene, J. Vasquez, S. A. Lee, E. Ghosn, K. L. Lynch, N. R. Roan, Distinctive features of SARS-CoV-2-specific T cells predict recovery from severe COVID-19. *Clin. Rep.* **36**, 109414 (2021).
 27. A. Roth, P. S. Chan, W. Jonas, Addressing the long COVID crisis: Integrative health and long COVID. *Glob. Adv. Integr. Med. Health.* **10**, 21649561211056596 (2021).
 28. F. J. Ryan, C. M. Hope, M. G. Masavuli, M. A. Lynn, Z. A. Mekonnen, A. E. L. Yeow, P. Garcia-Valtanen, Z. Al-Delfi, J. Gummow, C. Ferguson, S. O'Connor, B. A. J. Reddi, P. Hissaria, D. Shaw, C. Kok-Lim, J. M. Gleadle, M. R. Beard, S. C. Barry, B. Grubor-Bauk, D. J. Lynn, Long-term perturbation of the peripheral immune system months after SARS-CoV-2 infection. *BMC Med.* **20**, 26 (2022).
 29. S. C. Dalai, J. N. Dines, T. M. Snyder, R. M. Gittelman, T. Eerkes, P. Vaney, S. Howard, K. Akers, L. Skewis, A. Monteforte, P. R. Witte, C. Wolf, H. Nesse, M. Herndon, J. Qadeer, S. Duffy, E. Svejnoha, C. Taromino, I. M. Kaplan, J. Alsobrook, T. Manley, L. Baldo, Clinical validation of a novel T-cell receptor sequencing assay for identification of recent or prior severe acute respiratory syndrome coronavirus 2 infection. *Clin. Infect. Dis.* **75**, 2079–2087 (2022).
 30. Z. He, L. Ren, J. Yang, L. Guo, L. Feng, C. Ma, X. Wang, Z. Leng, X. Tong, W. Zhou, G. Wang, T. Zhang, Y. Guo, C. Wu, Q. Wang, M. Liu, C. Wang, M. Jia, X. Hu, Y. Wang, X. Zhang, R. Hu, J. Zhong, J. Yang, J. Dai, L. Chen, X. Zhou, J. Wang, W. Yang, C. Wang, Seroprevalence and humoral immune durability of anti-SARS-CoV-2 antibodies in Wuhan, China: A longitudinal, population-level, cross-sectional study. *Lancet* **397**, 1075–1084 (2021).
 31. L. Guo, G. Wang, Q. Zhang, L. Ren, X. Gu, T. Huang, J. Zhong, Y. Wang, X. Wang, L. Huang, L. Xu, C. Wang, L. Chen, X. Xiao, Y. Peng, J. C. Knight, T. Dong, B. Cao, J. Wang, SARS-CoV-2-specific antibody and T-cell responses 1 year after infection in people recovered from COVID-19: A longitudinal cohort study. *Lancet Microbe.* **3**, e348–e356 (2022).
 32. N. N. Jajour, D. Masopust, S. C. Jameson, T cell memory: Understanding COVID-19. *Immunity* **54**, 14–18 (2021).
 33. J. M. Dan, J. Mateus, Y. Kato, K. M. Hastie, E. D. Yu, C. E. Falitti, A. Grifoni, S. I. Ramirez, S. Haupt, A. Frazier, C. Nakao, V. Rayaprolu, S. A. Rawlings, B. Peters, F. Krammer, V. Simon, E. O. Saphire, D. M. Smith, D. Weiskopf, A. Sette, S. Crotty, Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021).
 34. K. W. Cohen, S. L. Linderman, Z. Moodie, J. Czartoski, L. Lai, G. Mantus, C. Norwood, L. E. Nyhoff, V. V. Edara, K. Floyd, S. C. De Rosa, H. Ahmed, R. Whaley, S. N. Patel, B. Prigmore, M. P. Lemos, M. P. Davis, S. Furth, J. B. O'Keefe, M. P. Gharpure, S. Gunisetty, K. Stephens, R. Antia, V. I. Zarnitsyna, D. S. Stephens, S. Edupuganti, N. Roupheal, E. J. Anderson, A. K. Mehta, J. Wrammert, M. S. Suthar, R. Ahmed, M. J. McElrath, Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Clin. Rep. Med.* **2**, 100354 (2021).
 35. K. Quintelier, A. Couckuyt, A. Emmaneel, J. Aerts, Y. Saeys, S. Van Gassen, Analyzing high-dimensional cytometry data using FlowSOM. *Nat. Protoc.* **16**, 3775–3801 (2021).
 36. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).
 37. P. Moss, The T cell immune response against SARS-CoV-2. *Nat. Immunol.* **23**, 186–193 (2022).
 38. N. le Bert, A. T. Tan, K. Kunasegaran, C. Y. L. Tham, M. Hafezi, A. Chia, M. H. Y. Chng, M. Lin, N. Tan, M. Linster, W. N. Chia, M. I.-C. Chen, L.-F. Wang, E. E. Ooi, S. Kalimuddin, P. A. Tambyah, J. G.-H. Low, Y.-J. Tan, A. Bertoletti, SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **584**, 457–462 (2020).
 39. D. Weiskopf, K. S. Schmitz, M. P. Raadsen, A. Grifoni, N. M. A. Okba, H. Endeman, J. P. C. van den Akker, R. Molenkamp, M. P. G. Koopmans, E. C. M. van Gorp, B. L. Haagmans, R. L. de Swart, A. Sette, R. D. de Vries, Phenotype and kinetics of SARS-CoV-2-specific T cells in COVID-19 patients with acute respiratory distress syndrome. *Sci. Immunol.* **5**, eabd2071 (2020).
 40. F. H. Seeger, T. Tonn, N. Krzossok, A. M. Zeiher, S. Dimmeler, Cell isolation procedures matter: A comparison of different isolation protocols of bone marrow mononuclear cells used for cell therapy in patients with acute myocardial infarction. *Eur. Heart J.* **28**, 766–772 (2007).
 41. S. A. M. van Kessel, T. C. Olde Hartman, P. L. B. J. Lucassen, C. H. M. van Jaarsveld, Post-acute and long-COVID-19 symptoms in patients with mild diseases: A systematic review. *Fam. Pract.* **39**, 159–167 (2022).
 42. C. E. Hastie, D. J. Lowe, A. McAuley, A. J. Winter, N. L. Mills, C. Black, J. T. Scott, C. A. O'Donnell, D. N. Blane, S. Browne, T. R. Ibbotson, J. P. Pell, Outcomes among confirmed cases and a matched comparison group in the Long-COVID in Scotland study. *Nat. Commun.* **13**, 5663 (2022).
 43. E. Espín, C. Yang, C. P. Shannon, S. Assadian, D. He, S. J. Tebbutt, Cellular and molecular biomarkers of long COVID: A scoping review. *EBioMedicine* **91**, 104552 (2023).
 44. R. M. Gulick, A. K. Pau, E. Daar, L. Evans, R. T. Gandhi, P. Tebas, R. Ridzon, H. Masur, H. C. Lane, J. Aberg, A. Adimora, J. Baker, L. B. Kreuziger, R. Bedimo, P. S. Belperio, A. Bhalla, T. Burgess, D. M. Campbell, S. V. Cantrill, K. Chew, K. Chiotos, C. M. Coopersmith, R. T. Davey, A. L. Dzierba, D. Eisnor, G. Eschenauer, J. Francis, J. J. Gallagher, D. V. Glidden, N. Goldenberg, B. Grund, A. Han, E. J. Hardy, C. Harrison, L. Henderson, E. S. Higgs, C. Hinkson, B. L. Hughes, S. Johnson, M. J. Keller, A. Y. Kim, R. Knight, S. Kuriakose, J. L. Lennox, A. M. Lerner, M. M. Levy, J. Z. Li, C. MacBrayne, G. Martin, N. R. Nadig, M. C. Nason, P. Patel, M. Proschian, A. T. Pavia, G. Schuler, N. Seam, V. Sheikh, S. Q. Simpson, K. Singh, S. Swindells, P. Tien, T. M. Uyeki, A. A. Waghmare, C. R. Wolfe, J. Yazdany, National Institutes of Health COVID-19 treatment guidelines panel: Perspectives and lessons learned. *Ann. Intern. Med.* **177**, 1547–1557 (2024).
 45. World Health Organization (WHO), "COVID-19 case definition" (WHO, 2020).
 46. World Health Organization (WHO), "Clinical management of severe acute respiratory infection when novel coronavirus (2019-nCoV) infection is suspected: Interim guidance" (WHO, 2020).
 47. European Centre for Disease Prevention and Control (ECDC), "Case definition for coronavirus disease 2019 (COVID-19)" (ECDC, 2022).
 48. G. Elias, P. Meysman, E. Bartholomeus, N. De Neuter, N. Keersmaekers, A. Suls, H. Jansens, A. Souquette, H. De Reu, M. P. Emonds, E. Smits, E. Lion, P. G. Thomas, G. Mortier, P. Van Damme, P. Beutels, K. Laukens, V. Van Tendeloo, B. Ogunjimi, Preexisting memory CD4 T cells in naïve individuals confer robust immunity upon hepatitis B vaccination. *eLife* **11**, e68388 (2022).
 49. R. Vita, S. Mahajan, J. A. Overton, S. K. Dhanda, S. Martini, J. R. Cantrell, D. K. Wheeler, A. Sette, B. Peters, The Immune Epitope Database (IEDB): 2018 Update. *Nucleic Acids Res.* **47**, D339–D343 (2019).
 50. J. N. Dines, T. J. Manley, E. Svejnoha, H. M. Simmons, R. Taniguchi, M. Klinger, L. Baldo, H. Robins, The ImmuneRACE Study: A prospective multicohort study of immune response action to COVID-19 events with the ImmuneCODE open access database. medRxiv 20175158 [Preprint] (2020).
 51. A. Poran, D. Harjanto, M. Malloy, C. M. Arieta, D. A. Rothenberg, D. Lenkala, M. M. Van Buuren, T. A. Addona, M. S. Rooney, L. Srinivasan, R. B. Gaynor, Sequence-based prediction of SARS-CoV-2 vaccine targets using a mass spectrometry-based bioinformatics predictor identifies immunogenic T cell epitopes. *Genome Med.* **12**, 70 (2020).
 52. A. Grifoni, J. Sidney, R. Vita, B. Peters, S. Crotty, D. Weiskopf, A. Sette, SARS-CoV-2 human T cell epitopes: Adaptive immune response against COVID-19. *Cell Host Microbe* **29**, 1076–1092 (2021).
 53. A. Nelde, T. Bilich, J. S. Heitmann, Y. Maringer, H. R. Salih, M. Roerden, M. Lübke, J. Bauer, J. Rieth, M. Wacker, A. Peter, S. Hörber, B. Traenkle, P. D. Kaiser, U. Rothbauer, M. Becker, D. Junker, G. Krause, M. Strengert, N. Schneiderhan-Marra, M. F. Templin, T. O. Joos, D. J. Kowalewski, V. Stos-Zweifel, M. Fehr, A. Rabsteyn, V. Mirakaj, J. Karbach, E. Jäger, M. Graf, L.-C. Gruber, D. Rachfalski, B. Preuß, I. Hagelstein, M. Märklin, T. Bakchoul, C. Gouttefangeas, O. Kohlbacher, R. Klein, S. Stevanović, H.-G. Rammensee, J. S. Walz, SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. *Nat. Immunol.* **22**, 74–85 (2021).

54. A. M. Altenhoff, N. M. Glover, C.-M. Train, K. Kaleb, A. Warwick Vesztrocy, D. Dylus, T. M. de Farias, K. Zile, C. Stevenson, J. Long, H. Redestig, G. H. Gonnet, C. Dessimoz, The OMA orthology database in 2018: Retrieving evolutionary relationships among all domains of life through richer web and programmatic interfaces. *Nucleic Acids Res.* **46**, D477–D485 (2018).
55. C. E. Whyte, D. J. Tumes, A. Liston, O. T. Burton, Do more with less: Improving high parameter cytometry through overnight staining. *Curr. Protoc.* **2**, e589 (2022).
56. X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J. C. Sanchez, M. Müller, pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77 (2011).
57. D. A. Bolotin, S. Poslavsky, I. Mitrophanov, M. Shugay, I. Z. Mamedov, E. V. Putintseva, D. M. Chudakov, MiXCR: Software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380–381 (2015).
58. S. Gielis, P. Moris, W. Bittremieux, N. De Neuter, B. Ogunjimi, K. Laukens, P. Meysman, Detection of enriched T cell epitope specificity in full T cell receptor sequence repertoires. *Front. Immunol.* **10**, 2820 (2019).
59. A. Postovskaya, A. Vujkovic, T. de Block, L. van Petersen, M. van Frankenhuijsen, I. Brosius, E. Bottieau, C. Van Dijk, C. Theunissen, S. H. van Ierssel, E. Vlieghe, E. Bartholomeus, K. Mullan, W. Adriaensen, G. Vanham, B. Ogunjimi, K. Laukens, K. Vercauteren, P. Meysman, Leveraging T-cell receptor – epitope recognition models to disentangle unique and cross-reactive T-cell response to SARS-CoV-2 during COVID-19 progression/ resolution. *Front. Immunol.* **14**, 1130876 (2023).
60. S. Valkiers, M. Van Houcke, K. Laukens, P. Meysman, ClusTCR: A python interface for rapid clustering of large sets of CDR3 sequences with unknown antigen specificity. *Bioinformatics* **37**, 4865–4867 (2021).

Acknowledgments: We appreciate the participation of all patients and patients' families in this study. We are grateful to all unmentioned clinicians, nurses, and lab colleagues. We thank S. Schlenner, P. Andrée Penttilä, R. Chinnaraj, and other staff at KU Leuven Flow and Mass Cytometry Facility for help with data acquisition on the Helios instrument. We thank S. Van Gassen and Y. Saeys for help with FlowSOM. We would also like to thank H. Goossens, J. Vereecken, M. Mattheijssens, and S. Van Goethem for help in participant recruitment. Data analysis was

performed using the resources and services provided by the Flemish Supercomputer Center, funded by the Research Foundation Flanders and the Flemish Government. This manuscript was published with support from the Universitaire Stichting van België. **Funding:** This work was supported by the European Union's Horizon 2020 research and innovation programme, grant agreement 851752-CELLULO-EPI (B.O.); Methusalem financing program of the Flemish Government to the University of Antwerp (E.L.); University of Antwerp special research fund BOF VAX-IDEA and VAXINAID-C2P Methusalem/Centre of Excellence (P.B.); Special Research, Core Facility funding, University of Antwerp (H.D.R.); and Research Foundation Flanders, grants G0H4520N (B.O., P.B., S.C., R.N., P.V.D., K.L., P.M., E.L., E.V., K.K.A., and K.V.), 1S38723N (A.P.), 1SH6624N (V.V.D.), G0G4220N (K.K.A.), and 1861219N (B.O.). **Author contributions:** Conceptualization: H.D.R., P.V.D., and B.O. Data curation: M.K.H. Formal analysis: M.K.H., A.P., H.B., J.M., and K.K.A. Funding acquisition: A.P., V.V.D., P.M., K.V., P.B., P.V.D., E.L., E.V., K.L., S.C., R.N., K.K.A., and B.O. Investigation: M.K.H., A.P., M.K., V.V.D., H.D.R., J.S., K.P., L.H., B.W., P.V.D., and S.C. Methodology: M.K.H., P.M., H.D.R., B.W., J.M., E.B., K.K.A., and B.O. Validation: M.K.H., H.D.R., and K.K.A. Visualization: M.K.H. and A.P. Project administration: K.K.A. and B.O. Resources: H.D.R., K.P., L.H., P.V.D., E.V., S.C., and K.K.A. Software: M.K.H., A.P., and H.B. Supervision: P.M., K.V., P.V.D., E.L., K.L., K.K.A., and B.O. Writing—original draft: M.K.H., P.V.D., K.K.A., and B.O. Writing—review and editing: all authors. **Competing interests:** P.M., K.L., and B.O. are shareholders and board members of ImmuneWatch BV. The other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The sequencing data that support the findings of this study have been deposited on Zenodo (DOI: 10.5281/zenodo.7785755). Flow Cytometry Standard data files are deposited at FlowRepository: CELLULO-EPI mass cytometry SARS-CoV-2-specific T cells (http://flowrepository.org/public_experiment_representations/6316), CELLULO-EPI flow cytometry SARS-CoV-2-specific T cell activation (http://flowrepository.org/public_experiment_representations/6313), and CELLULO-EPI flow cytometry SARS-CoV-2-specific T cell exhaustion (http://flowrepository.org/public_experiment_representations/8877).

Submitted 20 September 2024

Accepted 15 April 2025

Published 16 May 2025

10.1126/sciadv.adt2926