

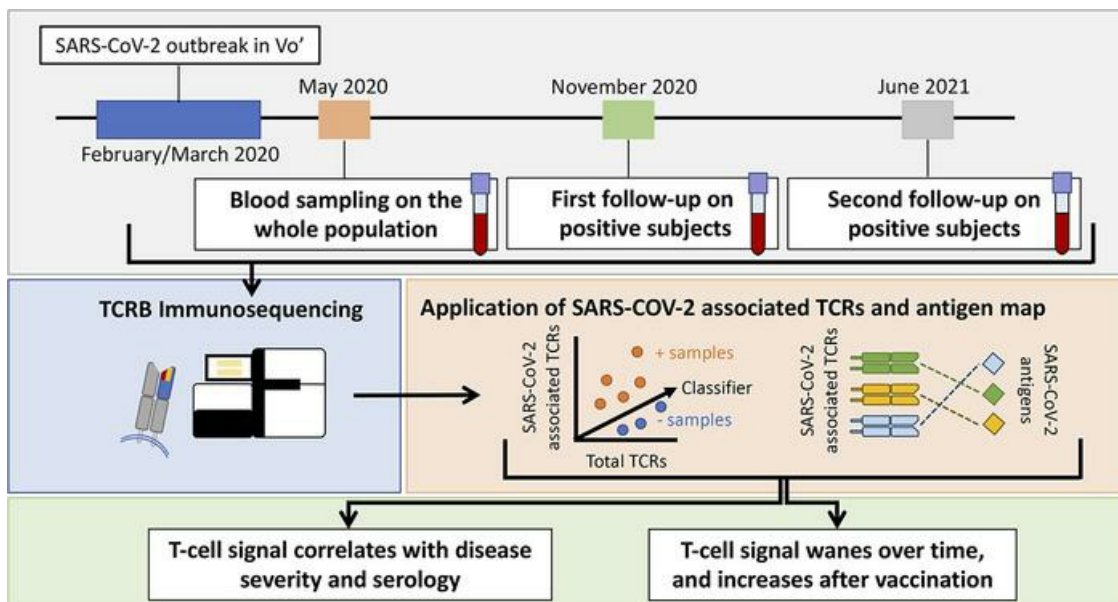
## Longitudinal analysis of T cell receptor repertoires reveals shared patterns of antigen-specific response to SARS-CoV-2 infection

Rachel M. Gittelman, ... , Andrea Crisanti, Giovanni Tonon

*JCI Insight.* 2022;7(10):e151849. <https://doi.org/10.1172/jci.insight.151849>.

Clinical Medicine COVID-19 Immunology

### Graphical abstract



Find the latest version:

<https://jci.me/151849/pdf>



# Longitudinal analysis of T cell receptor repertoires reveals shared patterns of antigen-specific response to SARS-CoV-2 infection

Rachel M. Gittelman,<sup>1</sup> Enrico Lavezzo,<sup>2</sup> Thomas M. Snyder,<sup>1</sup> H. Jabran Zahid,<sup>3</sup> Cara L. Carty,<sup>1</sup> Rebecca Elyanow,<sup>1</sup> Sudeb Dalai,<sup>1,4</sup> Ilan Kirsch,<sup>1</sup> Lance Baldo,<sup>1</sup> Laura Manuto,<sup>2</sup> Elisa Franchin,<sup>2</sup> Claudia Del Vecchio,<sup>2</sup> Monia Pacenti,<sup>5</sup> Caterina Boldrin,<sup>5</sup> Margherita Cattai,<sup>5</sup> Francesca Saluzzo,<sup>6</sup> Andrea Padoan,<sup>7</sup> Mario Plebani,<sup>7</sup> Fabio Simeoni,<sup>8</sup> Jessica Bordini,<sup>9</sup> Nicola I. Lorè,<sup>6</sup> Dejan Lazarević,<sup>8</sup> Daniela M. Cirillo,<sup>6</sup> Paolo Ghia,<sup>9,10</sup> Stefano Toppo,<sup>2,11</sup> Jonathan M. Carlson,<sup>3</sup> Harlan S. Robins,<sup>1</sup> Andrea Crisanti,<sup>2,12</sup> and Giovanni Tonon<sup>8,9,10</sup>

<sup>1</sup>Adaptive Biotechnologies, Seattle, Washington, USA. <sup>2</sup>Department of Molecular Medicine, University of Padova, Padua, Italy. <sup>3</sup>Microsoft Research, Redmond, Washington, USA. <sup>4</sup>Division of Infectious Diseases and Geographic Medicine, Stanford University School of Medicine, Stanford, California, USA. <sup>5</sup>Azienda Ospedale Padova, Microbiology and Virology Unit, Padua, Italy. <sup>6</sup>Division of Immunology, Transplantation and Infectious Disease, IRCCS Ospedale San Raffaele, Milan, Italy. <sup>7</sup>Department of Medicine, University of Padova, Padua, Italy. <sup>8</sup>Center for Omics Sciences and <sup>9</sup>Division of Experimental Oncology, IRCCS Ospedale San Raffaele, Milan, Italy. <sup>10</sup>Vita-Salute San Raffaele University, Milan, Italy. <sup>11</sup>CRIBI Biotech Center, University of Padova, Padua, Italy. <sup>12</sup>Department of Life Sciences, Imperial College London, London, United Kingdom.

**Authorship note:** RMG, EL, and TMS are co-first authors and contributed equally to the manuscript.

**Conflict of interest:** RMG, TMS, RE, CLC, SD, IK, and LB have employment and equity ownership with Adaptive Biotechnologies. HSR has employment, equity ownership, patents, and royalties with Adaptive Biotechnologies. HJZ and JMC have employment and equity ownership with Microsoft. SD, IK, and LB were Adaptive employees at time of research.

**Copyright:** © 2022, Gittelman et al. This is an open access article published under the terms of the Creative Commons Attribution 4.0 International License.

**Submitted:** June 8, 2021

**Accepted:** April 13, 2022

**Published:** April 19, 2022

**Reference information:** *JCI Insight*. 2022;7(10):e151849.  
<https://doi.org/10.1172/jci.insight.151849>.

T cells play a prominent role in orchestrating the immune response to viral diseases, but their role in the clinical presentation and subsequent immunity to SARS-CoV-2 infection remains poorly understood. As part of a population-based survey of the municipality of Vo', Italy, conducted after the initial SARS-CoV-2 outbreak, we sampled the T cell receptor (TCR) repertoires of the population 2 months after the initial PCR survey and followed up positive cases 9 and 15 months later. At 2 months, we found that 97.0% (98 of 101) of cases had elevated levels of TCRs associated with SARS-CoV-2. T cell frequency (depth) was increased in individuals with more severe disease. Both depth and diversity (breadth) of the TCR repertoire were positively associated with neutralizing antibody titers, driven mostly by CD4<sup>+</sup> T cells directed against spike protein. At the later time points, detection of these TCRs remained high, with 90.7% (78 of 96) and 86.2% (25 of 29) of individuals having detectable signal at 9 and 15 months, respectively. Forty-three individuals were vaccinated by month 15 and showed a significant increase in TCRs directed against spike protein. Taken together, these results demonstrate the central role of T cells in mounting an immune defense against SARS-CoV-2 that persists out to 15 months.

## Introduction

The clinical presentation of and subsequent immunity to COVID-19 are diverse, with a wide range of disease severity and susceptibility to future infection. With vaccines widely available, predicting who is still susceptible to severe disease is now a key public health goal across the globe. Antibody response has been shown to correlate with disease severity (1–4) and with protection from future infection (5–9), but it provides incomplete predictive power, making it important to assess additional factors that could influence these outcomes (10).

The T cell response plays a key role in the clearance of viral infections, governing and orchestrating both cellular and humoral immunity (11). Recent evidence demonstrates that the T and B cell responses can be discordant, and, in some individuals, there is a T cell response without antibody production (12–14). It remains uncertain whether, and for how long, prior infection with SARS-CoV-2 provides immunity against future reinfection, nor is it known how the severity of disease might influence long-term immunity (15). In the related SARS-CoV-1 and Middle East respiratory

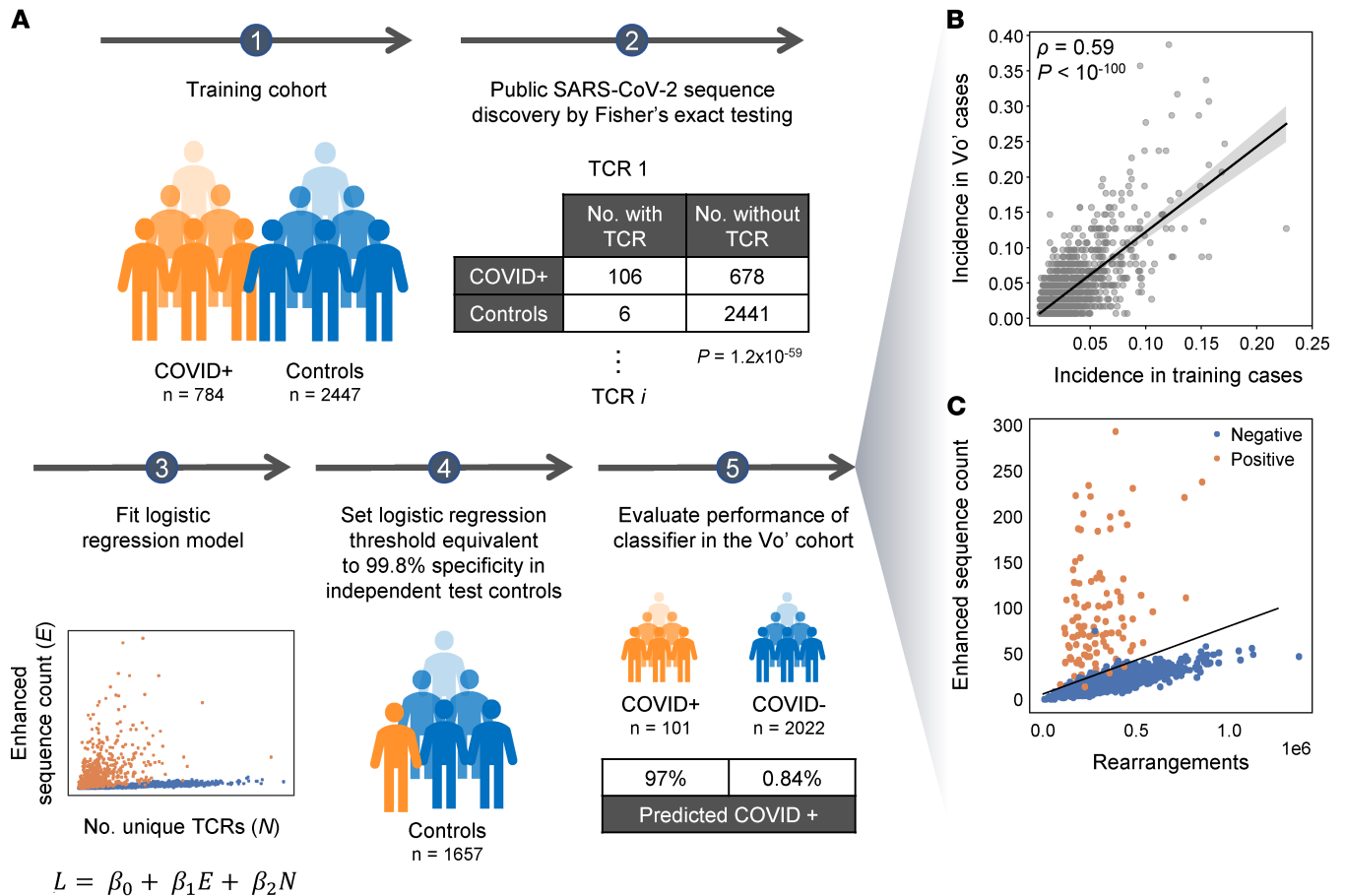
syndrome (MERS) coronaviruses, infections elicit an enduring T cell response with a more fleeting antibody response (16). Several studies have also shown that the T cell response is diverse, targeting regions across the viral genome, while the B cell response is concentrated against spike protein (17–22). These findings likely underlie recent observations that protection against the virus induced by vaccination remains high against recent emerging variants such as Delta and Omicron, despite large reductions in neutralizing antibody (NAb) titers against these variants (23–26). Finally, genetic factors influencing antigen presentation to T cells have also been shown to impact disease severity (27). These data further reinforce the central role of the T cell response in SARS-CoV-2 infection (28–30). Therefore, direct, quantitative measures of the T cell response to SARS-CoV-2 infection, particularly in longitudinal samples following recovery, may offer crucial insights into immunity and more broadly of the mechanisms underlying the immune response to this virus (11, 17, 31).

While various methodologies for measuring the T cell response exist, immunosequencing is a robust and precise tool that could be deployed at population scale (32–34). Following the first reported COVID-19 death in Vo', Italy, and subsequent lockdown of the entire municipality, a large PCR-based study was undertaken to screen the majority of the residents in that area in an unbiased manner. Approximately 60 days after the initial PCR survey, blood samples were collected from the majority of these study participants, and quantitative assessment of both SARS-CoV-2-specific T cells and IgG antibody titers were performed. After this unbiased screen of the municipality, additional longitudinal blood samples were collected from positive individuals at 9 months and 15 months in order to characterize persistence of the adaptive immune response over time. This cohort was also finely characterized in terms of demography, clinical presentation, hospitalization, comorbidities, therapies, and contact network (35). This carefully curated data set provides an invaluable opportunity to characterize the role of T cells over time in infection and vaccination.

## Results

*Characterizing the T cell response to SARS-CoV-2.* T cell receptor (TCR) sequencing was performed on blood samples collected 2 months after infection from 2291 residents of Vo' using ImmunoSEQ. Of those, 76 had a PCR-confirmed diagnosis of COVID-19, and the rest were PCR<sup>-</sup> at both surveys. Serology data were also available from 3 commercial tests, with 2156 samples having results from all 3 assays, as well as NAb titers for 153 samples. An additional 25 samples were determined to be confidently COVID-19<sup>+</sup>, even if they had negative or missing real-time PCR tests, if they were positive to at least 2 SARS-CoV-2 antigens (nucleocapsid protein and spike protein) based on the serology tests or if they had detectable neutralizing antibodies consistent with the baseline ground truth definition in Dorigatti et al. (36). Another set of 2022 samples that were negative by PCR and all 3 serology tests was considered to be confidently COVID-19<sup>-</sup> for comparison. Additional contact tracing and symptom data suggest that the 25 samples came from true cases, compared with lower rates of symptoms and household exposure in the confident negative set (Supplemental Figure 1A; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.151849DS1>). Of the 10 individuals with symptoms but negative PCR results, 5 were symptomatic prior to the PCR survey, while 5 were symptomatic during the PCR survey and may have been false negatives. In total, 33 of these 101 COVID-19 cases were asymptomatic, 58 had symptoms but did not require hospitalization, and 10 were hospitalized. The majority of COVID-19 cases were in patients 50 years of age or older (62%).

We previously identified 4287 public TCR sequences associated with SARS-CoV-2 using a case/control design that included several cohorts from the United States and Europe (18, 33) and validated their diagnostic accuracy in an independent US-based cohort (Figure 1A). Notably, 58% of these sequences were also present in PCR<sup>+</sup> Vo' cases, and 8.2% were found at an incidence of 5% or higher. Overall, the incidence of these sequences was highly correlated across cohorts, indicating that a substantial subset of public SARS-CoV-2-specific TCRs is common across distinct populations (Figure 1B). Using these public sequences and previously defined classification framework (18, 33), 98 of 101 (97.0%) of the COVID-19<sup>+</sup> individuals had a positive T cell test result, and 74 of 76 (97.4%) of PCR<sup>+</sup> individuals were positive (Supplemental Table 1). There was substantial variation in the number of public sequences across individuals (Figure 1C), and most of the additional 25 samples without PCR support showed equally high signal (Supplemental Figure 1B). Interestingly, 1 sample from an individual who was symptomatic but negative by all 3 serology assays and PCR had 75 SARS-CoV-2-associated sequences, making it strongly positive by the T cell test. The sample

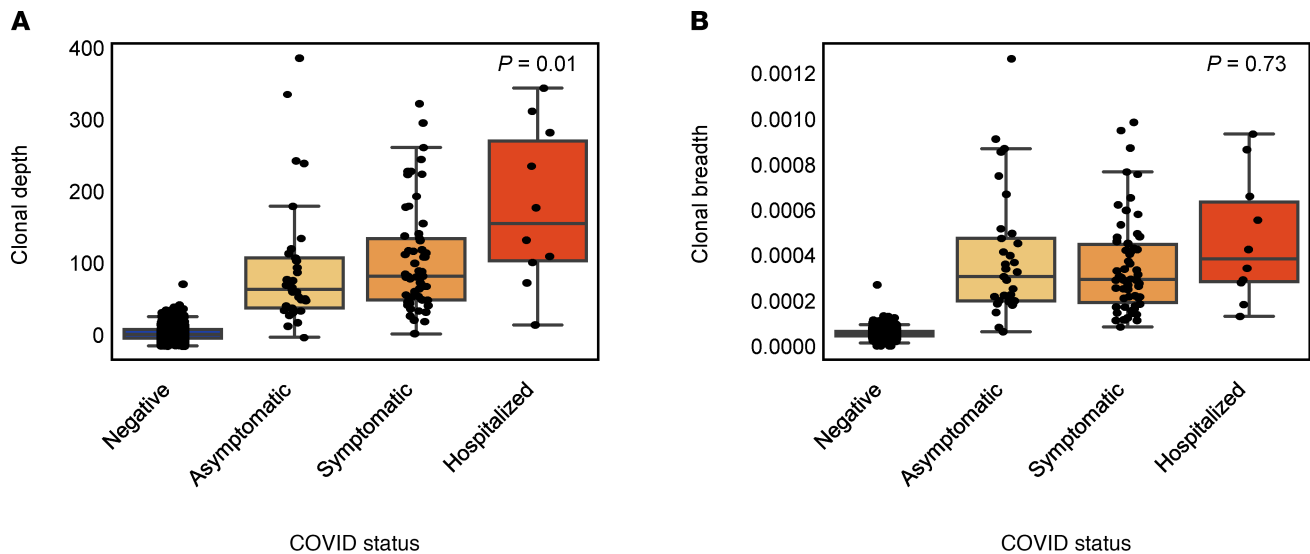


**Figure 1. Identification and use of public T cell receptors to characterize the SARS-CoV-2 immune response.** (A) Schema of the previously generated classification framework, also described in refs. 18 and 33, starting from a case-control design and Fisher's exact testing for each TCR on independent training data, to identify public TCR sequences that are overrepresented in cases versus controls. Following logistic regression to establish the T cell test threshold for determining recent or past infection, the receptors are applied to this Vo' study data set. COVID+ samples include all 101 samples defined by the positive ground truth samples set from Dorigatti et al. (36), while controls included the 2022 samples that were negative by PCR and all 3 serology tests. (B) Incidence of each TCR sequence compared in the training data and in the Vo' PCR+ cases. (C) The count of enhanced sequences is plotted versus the total number of unique TCR rearrangements for individuals in the Vo' study data set that were positive (orange) or negative (blue).

was not evaluated by the neutralization assay but could represent a case in which the T cell response is detectable, even in the absence of a B cell response. Overall, these data indicate that detectable TCR signatures are present 2 months after infection across a range of COVID-19 disease severities.

Two of the 2 COVID-19+ individuals with a negative T cell test (Figure 1C) result were asymptomatic, indicating that the T cell signal may vary with disease severity. To investigate this possibility, we next assessed the clonal depth and breadth of the T cell response to SARS-CoV-2 in comparison with disease severity. Depth and breadth were calculated as defined previously (18), where breadth measures the relative number of distinct SARS-CoV-2-associated T cell clonotypes, and depth measures the extent to which clonotypic T cells have expanded. Clonal depth was significantly lower among individuals who reported an asymptomatic infection, and it increased in symptomatic and hospitalized individuals (Figure 2A), while clonal breadth did not show strong trends (Figure 2B). This association was present both in older (>60 years) and younger (≤60 years) individuals (Supplemental Figure 2). These results suggest that clonal depth of T cells, reflecting the total number of cells that expanded to defend against the infection, may be a more direct measure of response than the clonal breadth. These results are consistent with other recent findings that the magnitude of T cell response is higher in symptomatic individuals, and these differences may persist for at least 6 months (37).

*Comparisons of the T cell and antibody responses across antigens.* B cell measurements correlate with protective immunity at a population level (38, 39), though, to date, no test has been accepted as a definitive measurement of protective immunity against SARS-CoV-2 for any given individual. This is likely due to



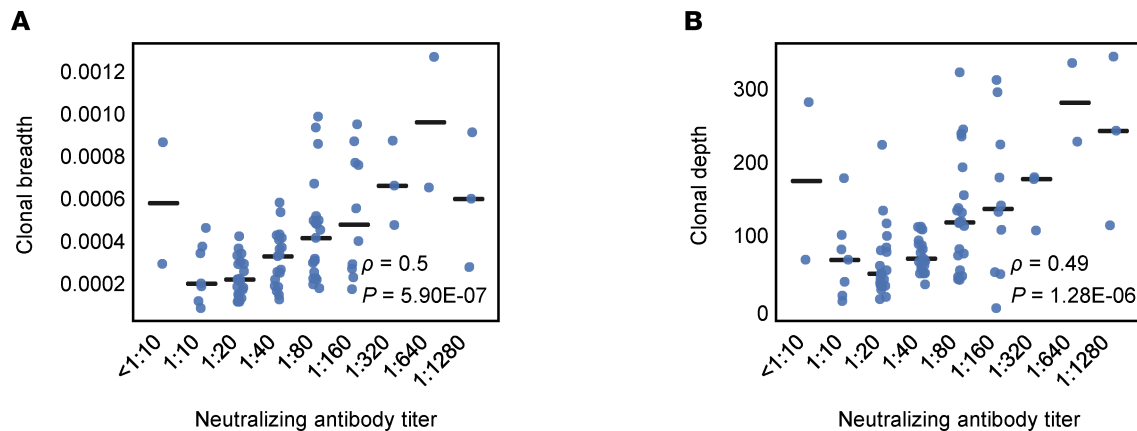
**Figure 2. T cell depth and breadth compared across 2022 confident COVID-19<sup>-</sup> individuals, and in 101 COVID-19<sup>+</sup> individuals faceted by disease severity. (A and B)** Clonal depth indicates the relative proportion of T cells that are SARS-CoV-2 specific, and clonal breadth indicates the fraction of all unique TCR DNA clones that are SARS-CoV-2 specific. *P* values correspond to Jonckheere's 2-sided trend test across the 3 PCR<sup>+</sup> categories. Data are expressed as median ± IQR.

other complex factors, including the T cell response that initiates the immune response. However, the extent to which T cell measurements correspond to the B cell response and to protective immunity in SARS-CoV-2 infection is not completely understood. Helper T cells play an important role in initiating the B cell response, while cytotoxic T cells engage independent cellular immune pathways. Additionally, NAbs, which are antibodies that bind to the virus and specifically block infection, are largely directed against the SARS-CoV-2 spike protein (40), while the serology tests compared in this study target either spike (LIAISON SARS-CoV-2 S1/S2 IgG, DiaSorin Molecular) or nucleoprotein (Elecsys Anti-SARS-CoV-2 assay, Roche Diagnostics; ARCHITECT SARS-CoV-2 IgG, Abbott Laboratories), necessitating analysis accounting for these factors. Thus, to investigate this question, we examined NAb titers assayed in 88 individuals who were positive by real-time PCR or by all 4 tests in the original survey (the 3 serology assays and T cell test), enabling comparison across assays without missingness — first comparing the overall T cell signal and then breaking analyses out by cell type and antigen specificity.

We found that both clonal breadth and depth of T cells were correlated with NAb titer (Figure 3, A and B). The correlation to NAb titers was comparable with the range from serology testing (higher than the level seen for Roche and lower than that seen for DiaSorin and Abbott; Supplemental Figure 3), despite being only an indirect measure of antibody response. Two very different molecular measurements (T cells versus antibody titers) being correlated suggests that aspects of the overall adaptive immune response can be inferred using just T cells in the case of SARS-CoV-2 infection. Further research to compare T cell signals to clinical outcomes will be needed to establish their utility as a potential correlate of immunity.

To examine different components of the T cell signature in more detail, we next characterized the cell type and antigen specificity of the public TCR sequences that we identified to diagnose SARS-CoV-2 infection (18). For many of the public TCRs and related sequences, direct observation in independent multiplexed antigen-stimulation experiments allowed us to assign TCR sequences to a specific target (41). Of the 4287 public TCR sequences, we identified 1776 and 1605 as CD8<sup>+</sup> and CD4<sup>+</sup> sequences, respectively, including 769 CD4<sup>+</sup>-associated sequences to spike protein and 836 CD4<sup>+</sup>-associated sequences to all other proteins.

We next compared and contrasted the breadth and depth of the CD4<sup>+</sup> and CD8<sup>+</sup> T cells with the B cell response. We found that the breadth of CD4<sup>+</sup> T cells with antigen assignments was positively correlated with antibody titers (DiaSorin: Spearman's  $\rho = 0.5$ ,  $P = 3 \times 10^{-6}$ ; NAb: Spearman's  $\rho = 0.6$ ,  $P = 2 \times 10^{-8}$ ; Abbott: Spearman's  $\rho = 0.5$ ,  $P = 8 \times 10^{-6}$ ; Roche: Spearman's  $\rho = 0.2$ ,  $P = 0.03$ ; Supplemental Figure 4). Similar correlations were observed between antibody titers and the depth of the CD4<sup>+</sup> T cell response (Supplemental Figure 5). Conversely, there was no correlation between CD8<sup>+</sup> T cell breadth and antibody levels (DiaSorin: Spearman's  $\rho = 0.01$ ,  $P = 0.92$ ; NAb: Spearman's  $\rho = -0.03$ ,  $P = 0.8$ ;



**Figure 3. Clonal breadth and depth compared across cases.** (A and B) Clonal breadth and depth compared across 88 cases that were positive by real-time PCR and/or all 4 additional tests and faceted by neutralizing antibody titer. Spearman's correlations are indicated by  $\rho$  and corresponding  $P$  values by  $P$ .

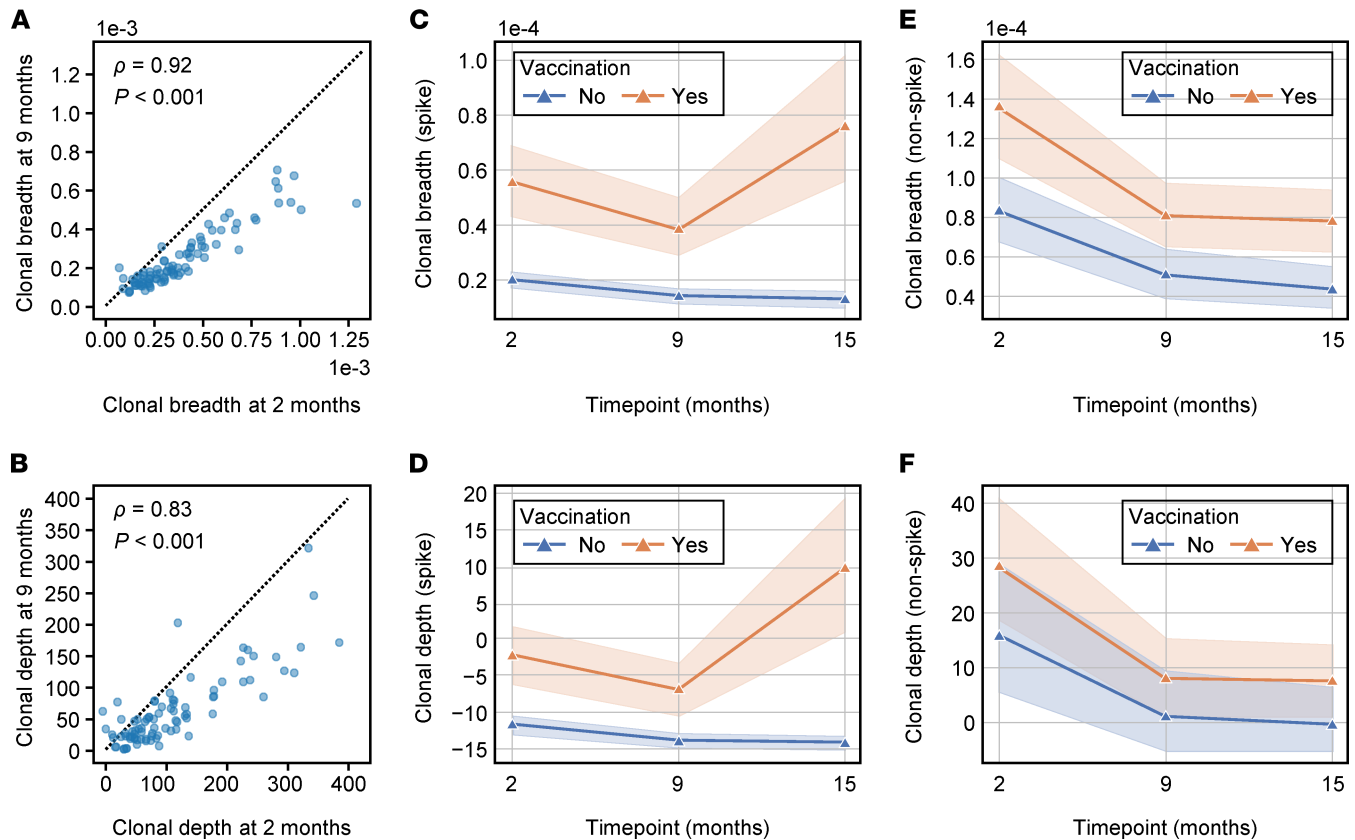
Abbott: Spearman's  $\rho = -0.06$ ,  $P = 0.59$ ; Roche: Spearman's  $\rho = -0.09$ ,  $P = 0.38$ ). These results underscore the role of helper T cells in supporting the generation of antibodies.

We then explored the potential association of spike-specific signal between T cells and antibodies. DiaSorin IgG spike and NAb titers were more significantly correlated with spike-specific CD4<sup>+</sup> T cell breadth as compared with non-spike-specific CD4<sup>+</sup> T cell breadth (Supplemental Figure 6). The breadth of spike-specific CD4<sup>+</sup> T cells had a partial Spearman's correlation of 0.5 ( $P = 7 \times 10^{-6}$ ) and 0.3 ( $P = 4 \times 10^{-3}$ ) to the DiaSorin IgG spike and NAb titers, respectively. Similar results were observed for clonal depth of spike-specific CD4<sup>+</sup> T cells (Supplemental Figure 7). On the contrary, when we examined the association with the nucleocapsid phosphoprotein (NP), no significant correlation between spike-specific T cell signal and the Abbott and Roche anti-NP titers was observed. Conversely, the breadth of non-spike-specific CD4<sup>+</sup> T cells had a partial Spearman correlation of 0.3 ( $P = 0.006$ ) and 0.2 ( $P = 0.04$ ) to the Abbott and Roche anti-NP titers. Overall, these results indicate that robust antibody-mediated immunity is associated with increased diversity of the antigen-specific CD4<sup>+</sup> T cell repertoire. These results also suggest that the information obtained from the TCR repertoire data may extend beyond the direct measure of the cellular immune response and help simultaneously dissect the concomitant humoral immune response.

*Longitudinal T cell response to SARS-CoV-2.* We next sought to characterize the persistence of the T cell response over time in the 101 COVID-19<sup>+</sup> samples from the initial sample collection. We performed additional T cell sequencing on 86 samples at 9 months and 72 samples collected at 15 months from these same individuals. At 9 months, 90.7% still had a positive T cell test. T cell signal was highly correlated between the 2 time points (Spearman's  $\rho = 0.83$  for clonal depth and Spearman's  $\rho = 0.92$  for clonal breadth), indicating that the magnitude of the initial memory response can be predictive of lasting T cell signal (Figure 4, A and B). In the subset of cases that were asymptomatic, a slightly lower proportion had a positive T cell test at 9 months: 84.6% (22 of 26).

Importantly, vaccines were available in Vo' by June 2021, when the 15-month samples were collected. Of the 72 samples collected at 15 months, 29 came from individuals who had not received any vaccine dose; 86.2% of these individuals still had a positive T cell test. In the subset of cases that were asymptomatic, the proportion positive by the T cell test decreased to 75% (6 of 8) at 15 months. The high proportion of cases, including asymptomatic ones, with a detectable T cell response at 15 months after infection indicates the strong durability of the memory T cell compartment and persistence of the T cell signal. NAb titers were available longitudinally on a large subset of the cases. Clonal depth and breadth continued to correlate with neutralization at each time point (Supplemental Figure 8).

The vaccinated individuals all had positive T cell tests at 15 months, which provided us the opportunity to investigate the T cell response to vaccination. All of the vaccines administered in the Vo' participants elicit a response against the SARS-CoV-2 spike protein but not against the viral proteins coded by the rest of the viral genome. In vaccinated individuals, we reasoned that the T cells targeting spike should rebound after the vaccination, while T cells directed against non-spike proteins should continue their slow declining trend. We again took advantage of our prior data (18) to characterize the antigen specificity of T cells in COVID-19 cases facing



**Figure 4. Longitudinal assessment of clonal depth and breadth.** (A and B) Clonal breadth (A) and depth (B) for COVID cases ( $n = 86$ ) at 2 months and 9 months. The dotted line indicates the line  $y = x$ . Spearman's correlations are indicated by  $\rho$  and corresponding  $P$  values by  $P$ . (C–F) Mean clonal breadth and mean clonal depth of T cells specific to spike protein are shown across time, while E depicts mean clonal breadth and F depicts mean clonal depth of T cells specific to non-spike proteins across time. The shaded areas represent the 95% CI of the mean. Samples were grouped by vaccination status at month 15.

samples by their vaccination status at 15 months (Figure 4, C–F). Indeed, though still largely detectable, both spike and non-spike T cells waned over time in individuals who had a prior COVID-19 infection and were not vaccinated (tests for negative trend:  $P < 0.001$  and  $P < 0.001$ , respectively). In vaccinated COVID-19 cases, however, spike T cells significantly increased both in breadth and depth after vaccination (comparison of month 9 to month 15:  $P < 0.001$  and  $P < 0.001$ , respectively), whereas non-spike T cell clonal depth and breadth continued to decline (test for trend:  $P < 0.001$ ). We did observe significant differences in the clonal depth and breadth of spike T cells between vaccinated and unvaccinated cases, even at prevaccine time points (all  $P < 0.02$ ), likely due to biases in age, disease severity, and possibly other factors in the subset of individuals who elected early vaccination (Supplemental Table 2). However it is unlikely that these differences could explain the significant increases in clonal depth and breadth observed after vaccination. A similar trend was observed in antibody levels (24), and this further supports the vaccine-related changes observed in T cells.

### Discussion

In this study, we present evidence — gathered in an unbiased and carefully annotated population — that the T cell response, as assessed by immunosequencing, is a highly sensitive and specific indicator of prior SARS-CoV-2 infection. We found that T cell responses were associated with clinical severity and with NAb levels and that they were boosted by vaccination. Importantly, they were present at least 2 months after infection, even in cases that were asymptomatic, and they persisted 15 months after the initial infection in a majority of individuals.

The findings that T cell immunity has similar sensitivity to the antibody response in detecting SARS-CoV-2 infection have important conceptual and diagnostic implications; however, they are not entirely unexpected. Previous experiences with the related MERS and SARS-CoV-1 infections demonstrated that

coronavirus-specific T cells have long-term persistence and contribute to protection even in individuals without seroconversion (28–30). Recent evidence suggests that a similar pattern is present during SARS-CoV-2 infection (14, 42, 43).

The structure of the T cell repertoire, in terms of diversity and clonal expansion, also provides important information to promote understanding some of the differences that characterize symptomatic and asymptomatic individuals. Our findings demonstrate that the T cell response correlated with prior disease symptoms and severity at a convalescent time point over 2 months after infection, suggesting — as potential hypotheses — that differences in viral load or viral persistence during acute infection correlate both with symptoms and the depth of the T cell response. This elevated signal persisted over time and was detected at 15 months, though at slightly lower levels. In addition, information gathered from the TCR repertoire analysis also correlates with the humoral responses to the virus, providing insight on the overall immune system response and possible defense against SARS-CoV-2. The T cell response from helper T cells strongly correlated with NAb titer, a potential measure of protection. Helper T cells recognizing spike protein antigens correlated more strongly with the overall antibody levels to the same protein but not other proteins, suggesting that antigen-specific CD4<sup>+</sup> T cell help may be required for robust development of antibodies during viral infections (44). These correlations between helper T cells and NAb may provide evidence relevant to disease pathology, as it has not been established whether the T cell response is exclusively beneficial or whether it might also contribute to immunopathology (11).

Our findings also have important implications for measuring response to vaccines, as well as to natural infection, since the broad set of TCRs measured in this assay allows for separating responses to spike, which is the primary target of most current vaccines, from other viral proteins. Indeed, the T cell response was sensitive to changes elicited by the current generation vaccines targeting the SARS-CoV-2 spike protein, in line with previous observations using T cell functional assays (45). Clonal breadth and depth for sequences inferred to target the spike protein significantly increased after vaccination at the 15-month time point in cases receiving 1 or more vaccines, but they did not increase for sequences outside the spike. Additionally, no enhancement in depth and breadth was evident in individuals who were unvaccinated.

Correlations of T cell responses with disease severity, antibody measurements, and vaccination are promising, but additional validation studies with orthogonal T cell assays are needed to confirm the extent to which the depth and breadth of the public T cell response, as measured by immunosequencing, capture the overall T cell response. T cell phenotype and function represent another important dimension that has been shown to impact disease severity (46). Since bulk immunosequencing does not capture functional information, this limitation will need to be addressed in future studies. Finally, additional studies on novel SARS-CoV-2 variants are needed to assess how robust the current set of SARS-CoV-2-associated TCRs are to changes in SARS-CoV-2 as the virus evolves over time, though an early study examining the overlap of novel variants and previously generated T cell/epitope maps is promising (47).

Given the continued emergence of new viral variants capable of at least partial evasion of vaccine-induced immunity (23), there is a strong need for correlates of protection against severe disease. While antibody titers correlate with disease severity and with the probability of breakthrough infection after vaccination, they are not sufficiently predictive to inform individuals of their COVID-19 risk (38, 39). Despite the limitations discussed above, our data suggest that T cells represent another key component of the adaptive immune response that correlates with important clinical factors, including disease severity. Mounting evidence also suggests that, while antibody-derived immunity is severely impacted by SARS-CoV-2 evolution (23, 24, 48), the T cell response may be more robust against viral variants (25, 47, 49). Though tools exist to study T cells in a research setting, they remain difficult to quantify accurately at population scale (50). T cell immunosequencing, which requires only genomic DNA as an input, is thus an attractive and scalable tool for use in clinical applications.

## Methods

*Clinical cohort and sample collection.* This report extends results for the Vo', Italy, cohort initially described in Lavezzo et al. (35). Upon the detection of SARS-CoV-2 in a deceased resident of Vo' on February 21, 2020, an epidemiological study was conducted to investigate the prevalence of SARS-CoV-2 infection in the municipality. Sampling for viral PCR testing was performed on the majority of the population immediately after the detection of the first cases (February 21–29, 2020) and again at the end of a 2-week lockdown (March 7, 2020). Follow-up serum and whole blood samples were collected 56 days



later (month 2 time point) in early May, at month 9, and again at month 15 for antibody serology and T cell testing. Antibody response was measured using 3 commercial serology testing kits: LIAISON SARS-CoV-2 S1/S2 IgG (DiaSorin Molecular), Elecsys Anti-SARS-CoV-2 assay (Roche Diagnostics), and ARCHITECT SARS-CoV-2 IgG (Abbott Laboratories). Manufacturer cutoffs were used for determining if samples were positive by each test. Serum samples were also used to perform microneutralization assays; more detail on the antibody testing methods is in Dorigatti et al. (36).

In addition to biospecimen collection, clinical data were collected for each study participant, including the results of SARS-CoV-2 testing, demographics, health records, and residence and contact network information. The definition of symptomatic used in this study is a participant who required hospitalization and/or reported fever (yes/no or a temperature above 37°C), and/or cough, and/or at least 2 of the following symptoms: sore throat, headache, diarrhea, vomit, asthenia, muscle pain, joint pain, loss of taste or smell, or shortness of breath. Symptomatic cases that reported hospitalization are split out separately as “hospitalized” in the disease severity analyses.

*Immunosequencing of TCR repertoires.* Genomic DNA was extracted from frozen, plasma-depleted blood samples using the Qiagen DNeasy Blood Extraction Kit (Qiagen). As much as 18 µg of input DNA was then used to perform immunosequencing of the third complementarity determining (CDR3) regions of TCR-β chains using the ImmunoSEQ Assay (Adaptive Biotechnologies). Briefly, input DNA was amplified in a bias-controlled multiplex PCR, followed by high-throughput sequencing. Sequences were collapsed and filtered to identify and quantitate the absolute abundance of each unique TCR-β CDR3 region for further analysis, as previously described (32, 51, 52). In order to quantify the proportion of T cells out of total nucleated cells input for sequencing, or T cell fraction, a panel of reference genes present in all nucleated cells was amplified simultaneously (53).

*Characterization of the T cell response.* Classification of prior infection with SARS-CoV-2, as well as the clonal depth and breadth of T cell response, were calculated using a method similar to prior work (18). Briefly, TCR repertoires from 784 unique cases of real-time PCR-confirmed SARS-CoV-2 infection and 2447 healthy controls collected before 2020 were compared by 1-tailed Fisher’s exact tests to identify 4469 public TCR-β sequences (enhanced sequences) significantly enriched in SARS-CoV-2<sup>+</sup> samples (all training data to identify the enhanced sequences for SARS-CoV-2 infection came from multiple other study cohorts and not the population being analyzed here). The enhanced sequences were used to develop a classifier predicting current or past infection with SARS-CoV-2 using a simple 2-feature logistic regression with independent variables E and N, where E is the number of unique TCR-β DNA sequences that encode an enhanced sequence and N is the total number of unique TCR-β DNA sequences in that subject. Application of this initial clinical classifier to this study demonstrated the high sensitivity (97%) reported above.

We have since developed a method to improve specificity near the decision boundary of the logistic regression by filtering enhanced sequences that may be potential false positives. Specifically, TCRs that are likely associated with CMV or with multiple antigens in different HLA backgrounds — and, thus, not truly diagnostic of SARS-CoV-2 infection — are identified by Fisher’s exact testing on TCR-β repertoires of ~2000 healthy controls with available HLA genotyping and CMV serotyping data. From this list of ~1.8 million sequences, the 182 sequences that were also identified as SARS-CoV-2-enhanced sequences were removed, leaving 4287 enhanced sequences. The 2-feature logistic regression classifier was refitted to the original training data using this pruned enhanced sequence list, and a decision boundary representing 99.8% specificity on 1657 controls was used to define the test-positive threshold used in the present study. The pruned list of enhanced sequences was also used to calculate the clonal depth and breadth using the same formulae as in Snyder et al. (18).

*Antigen-specific assignment of TCRs.* We assigned public TCRs to antigens and CD4<sup>+</sup> or CD8<sup>+</sup> cellular phenotypes by cross-referencing enhanced sequences identified via our case/control design with TCRs observed in multiplexed antigen-stimulation experiments, both described in prior work (18). Briefly, in these experiments, viral peptide panels were designed separately for class I and class II HLA binding. Peptides in each panel were pooled in a combinatorial fashion as described previously (41). T cells from COVID-19<sup>+</sup> donor PBMCs were expanded, stimulated with the peptide panel pools, and then sorted into activated CD4<sup>+</sup> or CD8<sup>+</sup> subsets for sequencing.

To maximize the number of TCR antigen assignments, we identified a set of public TCRs from an augmented sample of repertoire data. We combined the training and validation repertoires with an additional 1143 COVID<sup>+</sup> samples accrued since this model was developed and included samples from this

study that were identified as COVID-19<sup>-</sup> by our T cell test as controls. Our final sample of repertoires consisted of 1927 cases and 4135 controls. We identified ~500,000 public TCRs with a Fisher's Exact Test (FET)  $P < 0.05$ . At this level of significance, we expected a significant fraction of the public TCRs to be false positives; however, we cross-referenced this list of TCRs with a set of ~400,000 TCRs that was independently derived from our antigen-stimulation experiments, yielding 3381 overlapping TCRs. The fraction of false-positive TCRs in this overlapping set was significantly smaller. These 3381 overlapping TCRs had antigen protein and CD4<sup>+</sup>/CD8<sup>+</sup> assignments determined from our antigen-stimulation experiments.

*Data and materials availability.* Clinical data and T cell repertoire profiles are available as part of the ImmuneCODE data resource (54) and can be downloaded from the Adaptive Biotechnologies immune-ACCESS site under the immuneACCESS Terms of Use at <https://clients.adaptivebiotech.com/pub/gittelman-2022-jci>.

*Statistics.* Statistical methods used for identification of SARS-CoV-2-associated TCRs and for assignment of public TCRs to antigens are described above. We used Jonckheere's 2-sided trend test to evaluate ordered differences in clonal depth or breadth among COVID-19<sup>+</sup> individuals grouped by disease severity.  $P$  values less than 0.05 were considered significant.

We calculated the Spearman's rank correlations between antibody titers and CD4<sup>+</sup> and CD8<sup>+</sup> T cell response using the Pingouin package in Python (55) and reported the 2-sided significance. We noted that the spike-specific CD4<sup>+</sup> T cell signal correlates with the non-spike-specific signal (Spearman's  $\rho = 0.5$ ,  $P = 2 \times 10^{-6}$ ; Supplemental Figure 5). To disentangle the confounding correlations, we calculated partial Spearman's rank correlations between spike and non-spike specific T cell response and antibody titers and reported the 2-sided significance. The partial correlation coefficients and  $P$  values are denoted with tildes, appearing above the significance symbol, when reporting them in the figures. We examined the CD4<sup>+</sup> T cell response specific to spike and all other assayed proteins. When calculating the partial correlation between the antibody titers of 1 test and spike-specific T cell response, we took the non-spike protein-specific responses as a covariate, and vice versa. The partial correlations we calculated characterize the correlation between 2 variables that cannot be explained by the covariates and, thus, is conservative.

For longitudinal analyses of vaccinated or unvaccinated individuals, we used generalized estimating equation models with exchangeable covariance structure to account for correlation over time and adjusted for age, sex, and COVID-19 severity. To assess differences between the 9- and 15-month time points, we used linear regression models adjusted for age, sex, and COVID-19 severity and clustering on study participant in vaccinated or in unvaccinated individuals.

*Study approval.* The first, second, and third surveys of the Vo' population were approved by the Ethics Committee for Clinical Research of the province of Padova. Study participation was approved by written informed consent. For participants younger than 18 years, consent was provided by a parent or legal guardian.

## Author contributions

Order of co-first authors was assigned alphabetically. TMS, EL, IK, LB, DMC, PG, ST, JMC, HSR, GT, and AC conceptualized the study. Data curation was performed by EL, LM, and ST (clinical and serology); by IK and HSR (T cell repertoire); and by EF, CDV, M Pacenti, CB, MC, F Saluzzo, AP, M Plebani, F Simeoni, JB, NIL, and DL (laboratory testing). RMG, TMS, HJZ, RE, and CLC conducted data analyses, and SD and LB contributed to clinical interpretation. TMS, EL, GT, HSR, and AC supervised the study. Data visualizations were created by RMG, HJZ, and RE. TMS, HMC, HSR, GT, and RMG wrote the original draft of the manuscript, and RMG, SD, IK, EL, ST, DMC, PG, AC, and CLC reviewed and edited the manuscript.

## Acknowledgments

We thank the population of Vo' for volunteering en masse to participate in this study. We would like to thank Damon May for helpful discussions, as well as Beryl Crossley and Mitch Pesesky for assistance reviewing the data. This work was supported by funding from Adaptive Biotechnologies; the Veneto region; the Fondazione Umberto Veronesi, Misura Ricerca COVID-19, year 2020; the EDCTP2 programme supported by the European Union; European Union's Horizon 2020 research and innovation program, under grant agreement no. 874735 (VEO); and the University of Padova and the Department of Molecular Medicine (STARS-CoG ISS-MYTH and PRID/SID 2020 [EL], and TOPP\_PRIV20\_01 and TOPP\_SID19\_01 [ST]). Medical writing support was funded by Adaptive Biotechnologies and provided by Leslie Mitchell, Rachel Salmon, and Melanie Styers, of BluPrint Oncology Concepts LLC.

Address correspondence to: Harlan S. Robins, 1551 Eastlake Ave E, Seattle, Washington 98102, USA. Phone: 206.659.0067; Email: hrobins@adaptivebiotech.com. Or to: Andrea Crisanti, Via VIII Febbraio, 2, 35122 Padova, Italy. Phone: 39.049.827.5111; Email: andrea.crisanti@unipd.it. Or to: Giovanni Tonon, Via Olgettina, 60, 20132 Milano, Italy. Phone: 39.02.26431; Email: tonon.giovanni@hsr.it.

SD's present address is: Palo Alto Medical Foundation, Palo Alto, California, USA. IK's present address is: Merck Research Labs, South San Francisco, California, USA. LB's present address is: Freenome Inc., South San Francisco, California, USA.

1. Iyer AS, et al. Persistence and decay of human antibody responses to the receptor binding domain of SARS-CoV-2 spike protein in COVID-19 patients. *Sci Immunol.* 2020;5(52):eabe0367.
2. Isho B, et al. Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. *Sci Immunol.* 2020;5(52):eabe5511.
3. Seow J, et al. Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans. *Nat Microbiol.* 2020;5(12):1598–1607.
4. Boonyaratanakornkit J, et al. Clinical, laboratory, and temporal predictors of neutralizing antibodies against SARS-CoV-2 among COVID-19 convalescent plasma donor candidates. *J Clin Invest.* 2021;131(3):e144930.
5. Gilbert PB, et al. Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy clinical trial. *Science.* 2022;375(6576):43–50.
6. Frenck RW, et al. Safety, immunogenicity, and efficacy of the BNT162b2 Covid-19 vaccine in adolescents. *N Engl J Med.* 2021;385(3):239–250.
7. Gundlapalli AV, et al. SARS-CoV-2 serologic assay needs for the next phase of the US COVID-19 pandemic response. *Open Forum Infect Dis.* 2020;8(1):ofaa555.
8. Khoury DS, et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat Med.* 2021;27(7):1205–1211.
9. Earle KA, et al. Evidence for antibody as a protective correlate for COVID-19 vaccines. *Vaccine.* 2021;39(32):4423–4428.
10. Rydzynski Moderbacher C, et al. Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell.* 2020;183(4):996–1012.
11. Altmann DM, Boyton RJ. SARS-CoV-2 T cell immunity: specificity, function, durability, and role in protection. *Sci Immunol.* 2020;5(49):eabd6160.
12. Oja AE, et al. Divergent SARS-CoV-2-specific T- and B-cell responses in severe but not mild COVID-19 patients. *Eur J Immunol.* 2020;50(12):1998–2012.
13. Sekine T, et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *Cell.* 2020;183(1):158–168.
14. Gallais F, et al. Intrafamilial exposure to SARS-CoV-2 associated with cellular immune response without seroconversion, France. *Emerg Infect Dis.* 2021;27(1):113–121.
15. Poland GA, et al. SARS-CoV-2 immunity: review and applications to phase 3 vaccine candidates. *Lancet.* 2020;396(10262):1595–1606.
16. Mok CKP, et al. T-cell responses to MERS coronavirus infection in people with occupational exposure to dromedary camels in Nigeria: an observational cohort study. *Lancet Infect Dis.* 2021;21(3):385–395.
17. Peng Y, et al. Broad and strong memory CD4<sup>+</sup> and CD8<sup>+</sup> T cells induced by SARS-CoV-2 in UK convalescent COVID-19 patients. *Nat Immunol.* 2020;21(11):1336–1345.
18. Snyder MT, et al. Magnitude and dynamics of the T-cell response to SARS-CoV-2 infection at both individual and population levels [preprint]. <https://doi.org/10.1101/2020.07.31.2016564>. Posted on medRxiv September 17, 2020.
19. Barnes CO, et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature.* 2020;588(7839):682–687.
20. Barnes CO, et al. Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell.* 2020;182(4):828–842.
21. Tarke A, et al. Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. *Cell Rep Med.* 2021;2(2):100204.
22. Sette A, Crotty S. Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell.* 2021;184(4):861.
23. Harvey WT, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol.* 2021;19(7):409–424.
24. Lavezzo E, et al. Neutralising reactivity against SARS-CoV-2 B.1.617.2 (Delta) variant by vaccination status and pre-exposure [preprint]. <https://doi.org/10.21203/rs.3.rs-1131280/v1>. Posted on Research Square on December 8, 2021.
25. Tarke A, et al. Impact of SARS-CoV-2 variants on the total CD4<sup>+</sup> and CD8<sup>+</sup> T cell reactivity in infected or vaccinated individuals. *Cell Rep Med.* 2021;2(7):100355.
26. Wang Z, et al. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature.* 2021;592(7855):616–622.
27. Peng Y, et al. An immunodominant NP<sub>108-113</sub>-B\*07:02 cytotoxic T cell response controls viral replication and is associated with less severe COVID-19 disease. *Nat Immunol.* 2022;23(1):50–61.
28. Tang F, et al. Lack of peripheral memory B cell responses in recovered patients with severe acute respiratory syndrome: a six-year follow-up study. *J Immunol.* 2011;186(12):7264–7268.
29. Channappanavar R, et al. Virus-specific memory CD8 T cells provide substantial protection from lethal severe acute respiratory syndrome coronavirus infection. *J Virol.* 2014;88(19):11034–11044.
30. Zhao J, et al. Recovery from the Middle East respiratory syndrome is associated with antibody and T-cell responses. *Sci Immunol.* 2017;2(14):eaan5393.

31. 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.* 2020;5(48):eabd2071.
32. Carlson CS, et al. Using synthetic templates to design an unbiased multiplex PCR assay. *Nat Commun.* 2013;4:2680.
33. Dalai SC, et al. Clinical validation of a novel T-cell receptor sequencing assay for identification of recent or prior SARS-CoV-2 infection [preprint]. <https://doi.org/10.1101/2021.01.06.21249345>. Posted on medRxiv on January 8, 2021.
34. Emerson RO, et al. Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire. *Nat Genet.* 2017;49(5):659–665.
35. Lavezzo E, et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature.* 2020;584(7821):425–429.
36. Dorigatti I, et al. SARS-CoV-2 antibody dynamics and transmission from community-wide serological testing in the Italian municipality of Vo'. *Nat Commun.* 2021;12(1):4383.
37. Zuo J, et al. Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection. *Nat Immunol.* 2021;22(5):620–626.
38. Post N, et al. Antibody response to SARS-CoV-2 infection in humans: a systematic review. *PLoS One.* 2020;15(12):e0244126.
39. Wu J, et al. SARS-CoV-2 infection induces sustained humoral immune responses in convalescent patients following symptomatic COVID-19. *Nat Commun.* 2021;12(1):1813.
40. Yang Y, Du L. SARS-CoV-2 spike protein: a key target for eliciting persistent neutralizing antibodies. *Signal Transduct Target Ther.* 2021;6(1):95.
41. Klinger M, et al. Multiplex identification of antigen-specific T cell receptors using a combination of immune assays and immune receptor sequencing. *PLoS One.* 2015;10(10):e0141561.
42. Thieme CJ, et al. Robust T cell response toward spike, membrane, and nucleocapsid SARS-CoV-2 proteins is not associated with recovery in critical COVID-19 patients. *Cell Rep Med.* 2020;1(6):100092.
43. Schultheiß C, et al. Next-generation sequencing of T and B cell receptor repertoires from COVID-19 patients showed signatures associated with severity of disease. *Immunity.* 2020;53(2):442–455.
44. Sette A, et al. Selective CD4+ T cell help for antibody responses to a large viral pathogen: deterministic linkage of specificities. *Immunity.* 2008;28(6):847–858.
45. Sahin U, et al. COVID-19 vaccine BNT162b1 elicits human antibody and T<sub>H</sub>1 T cell responses. *Nature.* 2020;586(7830):594–599.
46. Le Bert N, et al. Highly functional virus-specific cellular immune response in asymptomatic SARS-CoV-2 infection. *J Exp Med.* 2021;218(5):e20202617.
47. May DH, et al. Immunosequencing and epitope mapping reveal substantial preservation of the T cell immune response to omicron generated by SARS-CoV-2 vaccines [preprint]. <https://doi.org/10.1101/2021.12.20.21267877>. Posted on medRxiv on December 27, 2021.
48. Edara VV, et al. mRNA-1273 and BNT162b2 mRNA vaccines have reduced neutralizing activity against the SARS-CoV-2 omicron variant. *Cell Reports Med.* 2022;3(2):100529.
49. Keeton R, et al. T-cell responses to SARS-CoV-2 spike cross-recognize Omicron. *Nature.* 2022;603(7901):488–492.
50. Ameratunga R, et al. Perspective: diagnostic laboratories should urgently develop T cell assays for SARS-CoV-2 infection. *Expert Rev Clin Immunol.* 2021;17(5):421–430.
51. Robins H, et al. Ultra-sensitive detection of rare T cell clones. *J Immunol Methods.* 2012;375(1–2):14–19.
52. Robins HS, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood.* 2009;114(19):4099–4107.
53. Pruessmann W, et al. Molecular analysis of primary melanoma T cells identifies patients at risk for metastatic recurrence. *Nat Cancer.* 2020;1(2):197–209.
54. Nolan S, et al. A large-scale database of T-cell receptor beta (TCRβ) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2 [preprint]. <http://doi.org/10.21203/rs.3.rs-51964/v1>. Posted on Research Square on August 4, 2020.
55. Vallat R. Pingouin: statistics in Python. *J Open Source Softw.* 2018;3(31):1026.