

## Article

# Analysis of TCR Repertoire by High-Throughput Sequencing Indicates the Feature of T Cell Immune Response after SARS-CoV-2 Infection

Yifan Wang<sup>1,2,3,†</sup>, Fugang Duan<sup>1,2,3,†</sup> , Zhu Zhu<sup>1,2,3,†</sup>, Meng Yu<sup>1,2,3</sup>, Xiaodong Jia<sup>4</sup>, Hui Dai<sup>1,2,3</sup>, Pingzhang Wang<sup>1,2,3</sup>, Xiaoyan Qiu<sup>1,2,3,\*</sup>, Yinying Lu<sup>4,\*</sup> and Jing Huang<sup>1,2,3,\*</sup> 

<sup>1</sup> Department of Immunology, School of Basic Medical Sciences, Peking University, 38 Xueyuan Road, Haidian District, Beijing 100191, China; wangyifanpku@pku.edu.cn (Y.W.); dfg5535@126.com (F.D.); jujuonfire@yeah.net (Z.Z.); 1711110028@pku.edu.cn (M.Y.); daihui@bjmu.edu.cn (H.D.); wangpzh@bjmu.edu.cn (P.W.)

<sup>2</sup> NHC Key Laboratory of Medical Immunology, Peking University, Beijing 100191, China

<sup>3</sup> Key Laboratory of Molecular Immunology, Chinese Academy of Medical Sciences, Beijing 100191, China

<sup>4</sup> Comprehensive Liver Cancer Center, The 5th Medicine Center of PLA General Hospital, Beijing 100039, China; feixiang.5420@163.com

\* Correspondence: qiuxy@bjmu.edu.cn (X.Q.); luyinying1973@163.com (Y.L.); huangjing82@bjmu.edu.cn (J.H.)

† These authors contributed equally to this work.



**Citation:** Wang, Y.; Duan, F.; Zhu, Z.; Yu, M.; Jia, X.; Dai, H.; Wang, P.; Qiu, X.; Lu, Y.; Huang, J. Analysis of TCR Repertoire by High-Throughput Sequencing Indicates the Feature of T Cell Immune Response after SARS-CoV-2 Infection. *Cells* **2022**, *11*, 68. <https://doi.org/10.3390/cells11010068>

Academic Editors: Alessandro Poggi, Isabella Quinti and Franco Pandolfi

Received: 14 November 2021

Accepted: 24 December 2021

Published: 27 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Coronavirus disease 2019 (COVID-19) is a global infectious disease caused by the SARS-CoV-2 coronavirus. T cells play an essential role in the body's fighting against the virus invasion, and the T cell receptor (TCR) is crucial in T cell-mediated virus recognition and clearance. However, little has been known about the features of T cell response in convalescent COVID-19 patients. In this study, using 5'RACE technology and PacBio sequencing, we analyzed the TCR repertoire of COVID-19 patients after recovery for 2 weeks and 6 months compared with the healthy donors. The TCR clustering and CDR3 annotation were exploited to discover groups of patient-specific TCR clonotypes with potential SARS-CoV-2 antigen specificities. We first identified CD4<sup>+</sup> and CD8<sup>+</sup> T cell clones with certain clonal expansion after infection, and then observed the preferential recombination usage of V(D)J gene segments in CD4<sup>+</sup> and CD8<sup>+</sup> T cells of COVID-19 patients with different convalescent stages. More important, the TRBV6-5-TRBD2-TRBJ2-7 combination with high frequency was shared between CD4<sup>+</sup> T and CD8<sup>+</sup> T cells of different COVID-19 patients. Finally, we found the dominant characteristic motifs of the CDR3 sequence between recovered COVID-19 and healthy control. Our study provides novel insights on TCR in COVID-19 with different convalescent phases, contributing to our understanding of the immune response induced by SARS-CoV-2.

**Keywords:** SARS-CoV-2; COVID-19; T cell receptor; CDR3; immune memory

## 1. Introduction

Beginning in December 2019, the coronavirus disease 2019 (COVID-19) outbreak has posed a serious threat to more than 200 countries worldwide and has caused more than 5.2 million deaths (<https://covid19.who.int>, accessed on 13 December 2021). COVID-19 is caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), which usually leads to respiratory infections, and severe cases develop into severe pneumonia or even death [1–3]. The immune response to COVID-19 encompasses both B cell-mediated humoral responses through antibodies as well T cell activity. Many of the early studies on the immune response to SARS-CoV-2 have focused on neutralizing antibodies. A growing body of evidence suggests T cell responses are important for both early viral clearance as well as conferring protection through memory T cells for extended periods in COVID-19 patients [4–8]. However, the specific T cell immune response against SARS-CoV-2, including the underlying mechanisms, remains unclear.

The cellular immune response is mediated by T cells, which play a role in the direct killing of virus-infected cells via CD8<sup>+</sup> cytotoxic T cells as well as helping to direct the overall immune response through CD4<sup>+</sup> helper T cells. The humoral immune response also includes CD4<sup>+</sup> T cells that assist B cells in differentiating into plasma cells and subsequently producing antibodies specific to the viral antigen. CD4<sup>+</sup> T cells also promote the immune response of CD8<sup>+</sup> T cells and the formation of long-term memory CD8<sup>+</sup> T cells to exert antiviral capabilities effectively [9]. Thus, T cells are an important target for assessing the degree of SARS-CoV-2 infection and disease progression [10]. The T cell response to SARS-CoV-2 peaks about one to two weeks after infection and can be detected during the months of recovery [11]. It has been reported that the activated CD38<sup>+</sup>HLA-DR<sup>+</sup>, CD8<sup>+</sup> T cells, and CD38<sup>+</sup>HLA-DR<sup>+</sup> CD4<sup>+</sup> T cells that respond to virus infection are transiently increased in COVID-19 patients [12]. Significant elevation and unusual phenotypes of CD4<sup>+</sup> cells, which exhibit both a proliferative exhausted phenotype and a clonally expanded cytotoxic phenotype, were also observed in mild and moderate COVID-19 cases. However, these phenotypes between these two groups of CD4<sup>+</sup> T cells of mild and moderate COVID-19 cases exhibit distinct functional signatures, distinct TCR sharing patterns, and may represent two divergent destinations for naive CD4<sup>+</sup> T cells [13]. Moreover, a higher proportion of SARS-CoV-2 specific CD8<sup>+</sup> T cells were detected in mild cases, and these CD8<sup>+</sup> T cells have extensive and strong memory after the recovery period of COVID-19 [14]. On the contrary, a significant reduction in effector memory CD4<sup>+</sup> T cells, which were less expanded and skewed toward central memory T cells and TH2-like phenotypes, was detected in COVID-19 patients with severe disease, whereas terminally differentiated CD8<sup>+</sup>GZMK<sup>+</sup> effector cells were clonally expanded both during and after the infection [15]. The CD4<sup>+</sup> T memory response was detected in all recovered patients from COVID-19, and 70% of patients had established CD8<sup>+</sup> T memory response to SARS-CoV-2, lasting for more than 2 months [16,17]. Additionally, in COVID-19-recovered individuals receiving the vaccine, pre-existing SARS-CoV-2 specific memory cells showed both clonal expansion and a phenotypic shift towards more differentiated CCR7<sup>-</sup>CD45RA<sup>+</sup> effector cells. Therefore, memory T cells have always been the focus of attention after SARS-CoV-2 infection.

T cells recognize pathogen-derived peptides presented with the major histocompatibility complex (MHC) on the cell surface using hypervariable T cell receptor (TCR). TCR diversity is widely recognized as a direct measure of immune competence, as it quantifies the variety of foreign antigens and hence acts upon them. At the early stage of T cell development, TCR is generated by somatic rearrangement of variable (V), diversity (D), and joining (J) gene segments, known as V(D)J recombination [18]. Thus, the TCR repertoire is a critical factor in viral clearance. After activation by antigen recognition, T cells undergo clonal expansion, during which activated T cells rapidly proliferate to generate large numbers with identical TCRs to eliminate virus-infected cells. Most of the TCRs on circulating T cells are alpha and beta subunit heterodimers, and the specificity for an antigen is shaped by VDJ recombination [19].

Recently, sequencing technology has been widely used to reveal activation-induced phenotypic profiles of antigen-reactive T cells [20,21]. Some studies have revealed that clonality and skewing of TCR repertoires from COVID-19 patients by next-generation sequencing were associated with severity of diseases, such as early CD4<sup>+</sup> and CD8<sup>+</sup> T cell activation and interferon type I and III responses. The emergence of shared T cell clusters occurs in the rehabilitation of COVID-19 patients by next-generation sequencing technology [22]. It has also been reported that the patients with moderate COVID-19 have highly clonally expanded CD8<sup>+</sup> T cells [23]. Notably, the memory phenotype of T cell response is positively correlated with the severity of the disease, and there exists the specific memory CD8<sup>+</sup> T cell in convalescent COVID-19 patients by single-cell sequencing [24]. Additionally, preferential usage of V and J gene segments of TCR was also observed in COVID-19 patients [25]. However, studies on the TCR repertoire of COVID-19 patients with different convalescent stages are lacking. Further exploration is essential to understand the

mechanism of T cell response to SARS-CoV-2 and the formation of SARS-CoV-2-specific memory T cells.

In this study, we used the 5'RACE (Rapid Amplification of cDNA Ends) method combined with third-generation high-throughput sequencing technology to investigate the TCR $\alpha$  and TCR $\beta$  library sequences on peripheral blood CD4<sup>+</sup> and CD8<sup>+</sup> T cells of convalescent patients after SARS-CoV-2 infection, and then further analyzed the characteristics of their TCR repertoire of convalescent patients with different convalescent stages.

## 2. Materials and Methods

### 2.1. Donors and Blood Samples

Nine confirmed COVID-19 patients of 2 weeks convalescence and five healthy controls were obtained from Fifth Medical Center of Chinese PLA General Hospital, Beijing, China. The clinicopathologic characteristics of the study group participants with 2 weeks convalescence and healthy controls are listed in Table 1. Additionally, twenty confirmed COVID-19 patients of 6 months convalescence were obtained from the General Hospital of Central Theater Command of PLA, Wuhan, China. The study protocol was approved by the Institutional Review Board of Peking University (PUIRB). We complied with all relevant ethical regulations, and informed consent was obtained from all human participants (No:2020072D). No statistically significant difference in age or sex was observed among the three groups of patients.

**Table 1.** Clinicopathologic characteristics of the study population.

	HC (n = 5)		CP (2 Weeks) (n = 9)	
	Healthy (n = 5)	Moderate (n = 3)	Severe (n = 3)	Critical (n = 3)
<b>Severity</b>				
Age (years)	65.4 ± 8.76	58 ± 12.78	67 ± 9.9	73 ± 5.57
<b>Sex</b>				
Male	2	1	1	2
Female	3	2	2	1
<b>White blood cells (10<sup>9</sup>/L)</b>	5.83 ± 1.2	4.97 ± 1.35	7.76 ± 4.87	7.2 ± 3.68
<b>Percentage of neutrophil</b>	60.48 ± 6	59.3 ± 11.47	79.3 ± 2.4	90 ± 6.9
<b>Number of neutrophils (10<sup>9</sup>/L)</b>	3.22 ± 0.6	3.1 ± 1.27	6.11 ± 3.68	7.29 ± 3.7
<b>Percentage of lymphocytes</b>	20.5 ± 10.2	32.7 ± 11.46	16.4 ± 0.7	5.97 ± 4.5
<b>Number of lymphocytes (10<sup>9</sup>/L)</b>	1.6 ± 0.92	1.49 ± 0.23	1.26 ± 0.74	0.36 ± 0.04
<b>Percentage of monocytes</b>	6.27 ± 0.84	6.48 ± 0.45	4 ± 3.39	2.85 ± 2.69
<b>Number of monocytes (10<sup>9</sup>/L)</b>	0.28 ± 0.03	0.32 ± 0.08	0.4 ± 0.46	0.24 ± 0.03
<b>CRP (mg/L)</b>	1.91 ± 2.4	2.89 ± 1.1	8.49 ± 10.8	44.89 ± 39
<b>CK (U/L)</b>	25 ± 5.4	39.8 ± 16.46	35.5 ± 14.8	133.7 ± 100
<b>Alkalosis</b>	N	N	Y (2)	Y (3)
<b>Complication</b>				
ARDS	N	N	Y (2)	Y (3)
Secondary infections	N	N	N	N
<b>History of smoking</b>	N	N	N	N

HC, healthy controls; CP (2 weeks), COVID-19 patients with 2-week convalescence phase; CP (6 months), COVID-19 patients with 6-month convalescence phase; CRP, C-reactive protein; CK, creatinine kinase; ARDS, acute respiratory distress syndrome. N, none; Y, yes. The parentheses indicate the number of patients.

### 2.2. Isolation of PBMCs and T Cell Subpopulations

PBMC were isolated with the Ficoll-Paque density gradient centrifugation protocol. The CD4<sup>+</sup> T and CD8<sup>+</sup> T cells were purified from PBMCs with the Easysep<sup>®</sup> human CD4<sup>+</sup> or CD8<sup>+</sup> positive selection kits from Stem Cell technologies (San Diego, CA, USA). Purified cells were eluted and washed with PBS containing 2% (*v/v*) fetal bovine serum (FBS) and 1 mM/L EDTA.

### 2.3. TCR $\alpha$ and TCR $\beta$ Library Preparation and PCR Amplification

For the sorted CD4<sup>+</sup> T and CD8<sup>+</sup> T cells, total RNA was extracted using RaPure Total RNA Micro Kit (Magen, China) according to the manufacturer's instructions. A universal

primer binding site, sample barcode, and unique molecular identifier (UMI) sequences were introduced using the SMARTer<sup>®</sup> RACE 5'/3' kit (Takara, Japan) with TCR $\alpha$  and TCR $\beta$  constant segment-specific primers (Forward: CTAATACGACTCACTATAGGGC; Reverse1: GATTACGCCAAGCTTCACGGCAGGGTCAGGGTTCTGGATAT; Reverse 2: GATTACGCCAAGCTTCTCGGGTGGGAACACGTTTTTCAGGTCCTC) for cDNA synthesis. cDNA libraries were amplified in two PCR steps, introducing the second sample barcode and Illumina TruSeq adapter sequences at the second PCR step. The PCR program for both rounds was: 5 cycles at 94 °C for 30 s, 5 cycles at 68 °C for 30 s, and 25 cycles at 72 °C for 3 min (first-round PCR) and 40 cycles at 94 °C for 30 s, 68 °C for 30 s, and 72 °C for 2 min (second-round PCR). The PCR products were separated on 1.5% agarose gel by electrophoresis.

#### 2.4. TCR $\alpha$ and TCR $\beta$ Sequencing and Data Preprocessing

We measured the concentration of PCR products with a barcode and mixed the same amount of PCR products from all COVID-19 patients and healthy controls. The mixed PCR products were separated with 2% agarose gel, and then the extracted DNA was sent to Novogene Institute (Beijing, China). The amplicons were sequenced on a PacBio Sequel system using V3 chemistry (Courtesy of Pacific Biosciences of California Inc., Menlo Park, CA, USA). The raw data (raw reads) in FASTQ format were first processed using Python and custom Perl scripts. Clean data (clean reads) were obtained after removing various reads, including poly-N reads, reads without a 3' adapter or an insert tag, reads with 5' adapter contamination, reads with poly-A, T, G, or C low-quality reads from the raw data. Moreover, the Q20, Q30, and GC contents of the raw data were calculated. All sequences shorter than 200 bps, having homopolymers of 6 bps barcodes, containing primer mismatches, and with a quality score lower than 19, were removed.

#### 2.5. TCR Repertoire Analysis

The processing stage of TCR repertoire analysis started with the mapping of V, D, and J segments. Filtered reads were subsequently aligned to V, D, J, and C gene segments of TCR alpha (TRA) and TCR beta (TRB) locus for clonotype assembly of complementarity-determining region 3 (CDR3) nucleic acid sequences by IMGT/HighV-QUEST web server. Furthermore, the rarefaction curve of the TCR repertoire was generated using VDJtools for each subject. We obtained the core tabular format for VDJ tools by the conversion supported in the program, which contained count, frequency, CDR3 nucleotide sequence, V end, D start, D end, and J start. The frequency of various V–J junctions was calculated and displayed using basic analysis of VDJtools, making circus-style V–J usage plots. Additionally, the diversity estimation was utilized to compute a set of diversity statistics as well as to visualize the repertoire clonality and the comparison of diversity estimates. The V(D)J patterns of each sample were combined to analyze the dominant rearrangement pattern in different groups. The CDR3nt and CDR3aa in advantageous V(D)J usage were merged to further search for the dominant CDR3. All statistical analysis was implemented with R software (Version 4.1.0, downloaded from <http://www.r-project.org>, accessed on 10 November 2021).

### 3. Results

#### 3.1. Clinical Characteristics of COVID-19 Patients with Different Convalescent Stages for TCR Repertoire Profiling Analysis

To evaluate the blood immune characteristics of COVID-19 patients with different convalescent stages, thirty-four samples were included in the current study, including nine COVID-19 patients with two-week convalescent stage, twenty COVID-19 patients at the six-month-convalescent stage, and five healthy controls (Table 1). Moreover, these COVID-19 patients were classified into three groups, including moderate, severe, and critical cases. As shown in Table 1, the analysis of the blood routine test showed the number of white blood cells, mainly neutrophil, increased in the severe and critical cases of

COVID-19 patients at 2 weeks convalescence compared with healthy controls. However, the percentage and absolute number of lymphocytes and monocytes obviously reduced in the severe and critical cases compared with healthy controls. In addition, C-reactive protein (CRP) and creatine kinase (CK) were also elevated in the severe and critical cases, which had complications, including ARDS but not secondary infection.

We next collected their peripheral blood samples and isolated peripheral blood mononuclear cells (PBMCs). For each blood sample, we purified CD4<sup>+</sup> T cells and CD8<sup>+</sup> T cells with magnetic beads, respectively. From all samples, we isolated RNA and performed 5'RACE coupled with long-read high-throughput sequencing to amplify the complete DNA sequence of TCR $\alpha$  (TRA) and TCR $\beta$  (TRB) chains of antigen-specific T cells (Figure 1). A total of 32,796 and 15,086 TCR $\alpha$  and  $\beta$  chain sequences were obtained from COVID-19 patients and healthy controls, respectively. Among them, there are 11,324 effective TRA sequences and 6933 TRB sequences detected in CD4<sup>+</sup> T cells. There are 21,472 effective TRA sequences and 8153 TRB sequences detected in CD8<sup>+</sup> T cells (Table 2).

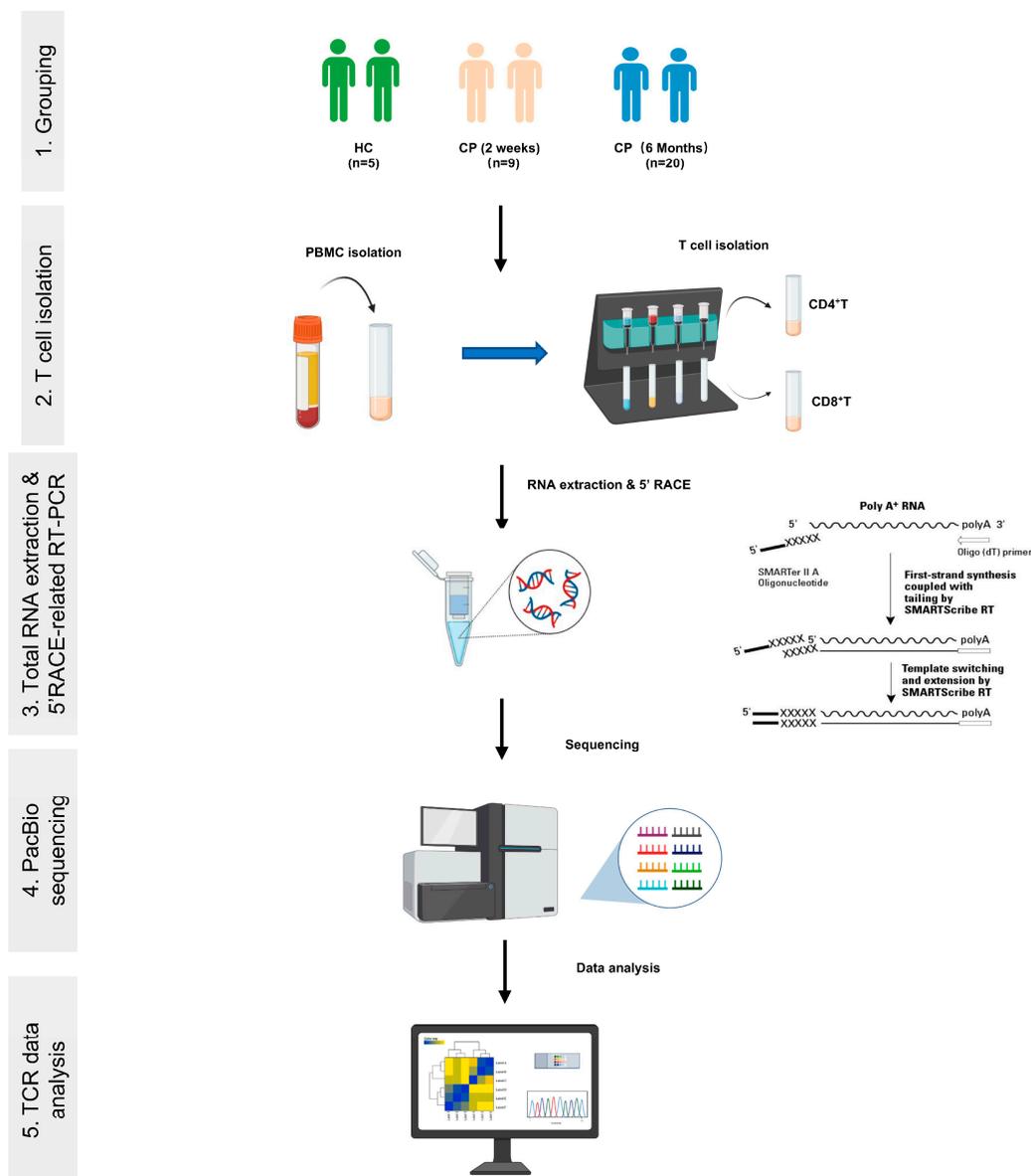
### *3.2. Significance of T Cell Receptor Bias during the Different Convalescent Phases of COVID-19 Patients Compared with Healthy Controls*

To investigate the changes in TCR repertoire after COVID-19 infection, we performed a comparative analysis of the T cell repertoire between the COVID-19 patients at different convalescent phases and healthy controls. Profiling of TRA and TRB repertoire showed that there is more diverse TCR clonality in CD8<sup>+</sup> T cells compared with CD4<sup>+</sup> T cells. At the same time, no significant difference was observed in overall frequencies of abundant TRA and TRB clonotypes in convalescent COVID-19 patients, relative to healthy controls (Figure 2A–D), possibly because the clonality frequency of TCR was not caused directly by SARS-CoV-2 infection. Moreover, certain TCRs were shared between the convalescent COVID-19 patients and healthy controls, while more TCRs were shared between the COVID-19 patients at the two-week convalescent stage and six-month convalescent stage (Figure 2E–H).

T cell receptors are generated by rearrangement of V and J gene segments for the TRA and by V, D, and J gene segments for the TRB [22]. Here, we explored the usage bias of V, D, and J gene segments for convalescent COVID-19 patients. First, we found that some V, D, and J gene segments on the TRA and TRB were significantly more frequent than healthy controls. Among convalescent COVID-19 patients, for TRA, the most frequently used gene segments were TRAV12-3 and TRAJ42 in either CD4<sup>+</sup> T or CD8<sup>+</sup> T cells (Figure 3A), while the frequencies of TRBV23-1 and TRBJ2-7 were significantly higher for TRB (Figure 3B).

To further explore whether a unique V(D)J recombination pattern is specific for convalescent COVID patients, we next compared the V(D)J pairing of TRA and TRB in each individual separately. In the CD4<sup>+</sup> T cells of COVID-19 patients at the two-week convalescent stage, 37 unique pairs of VJ rearrangement of TRA were found, such as TRAV8-4/TRAJ54 and TRAV17/TRAJ54 (Figure 3C), and 23 specific pairs of TRB VDJ patterns were also found in CD4<sup>+</sup> T cells (Figure 3D). Similar results were obtained in CD8<sup>+</sup> T cells, which showed there were 16 individual pairs of TRA VJ patterns and 18 specific pairs of TRB VDJ patterns in CD8<sup>+</sup> T cells of convalescent COVID-19 patients (Figure 3E,F). Of note, these unique gene pairs of TRB, such as TRBV5-4/D2/TRBJ2-5 and TRBV5-4/D1/TRBJ2-2, appeared at the 2-week convalescent stage and continued to the 6-month convalescent stage (Figure 3D,F), suggesting the clones with these TRB VDJ patterns might represent memory T cells phenotype. They may have expanded and participated in the elimination of residual SARS-CoV-2 in the convalescence stage. More important, we found some unique VDJ recombination of TRB was shared in both CD4<sup>+</sup> T cells and CD8<sup>+</sup> T cells of two-week convalescent COVID-19 patients, including TRBV6-5/D2/TRBJ2-7 and TRBV2/D1/TRBJ2-1, suggesting that these T cell clones with the above VDJ recombination were specific for SARS-CoV-2 antigen (Figure 3D,F). Additionally, these specific pairs of TRA and TRB VDJ patterns in CD4<sup>+</sup> T or CD8<sup>+</sup> T cells existed among the three groups, including moderate,

severe, and critical cases, suggesting that these sequences were not significantly correlated with disease severity (data not shown).

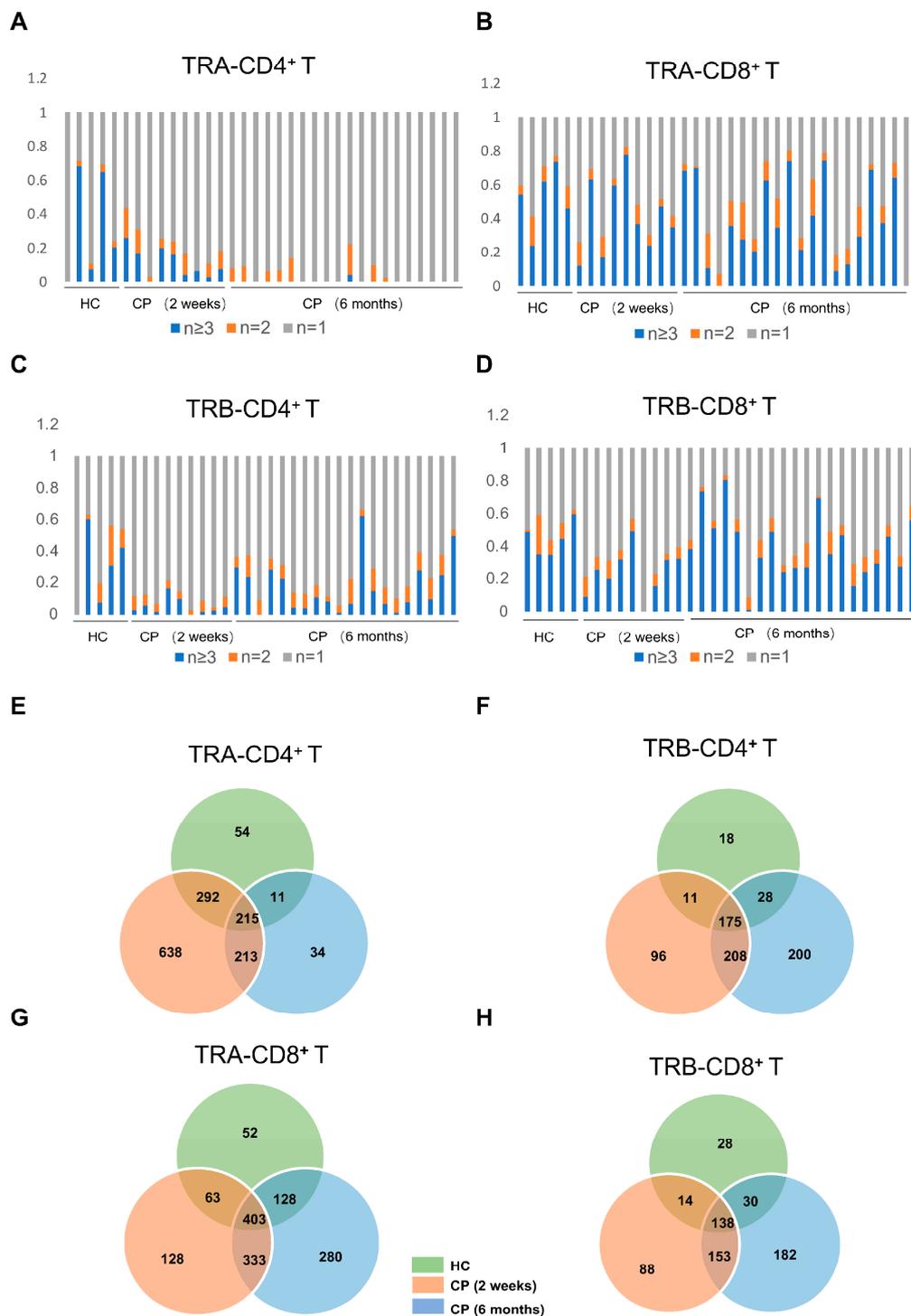


**Figure 1.** Schematic diagram of TCR repertoire analysis in the convalescent COVID-19 patients. Peripheral blood mononuclear cells were collected from COVID-19 patients with 2-week and 6-month convalescence phase and healthy controls, and CD4<sup>+</sup> T and CD8<sup>+</sup> T cells were sorted to perform 5' RACE-related RT-PCR combined with PacBio sequencing.

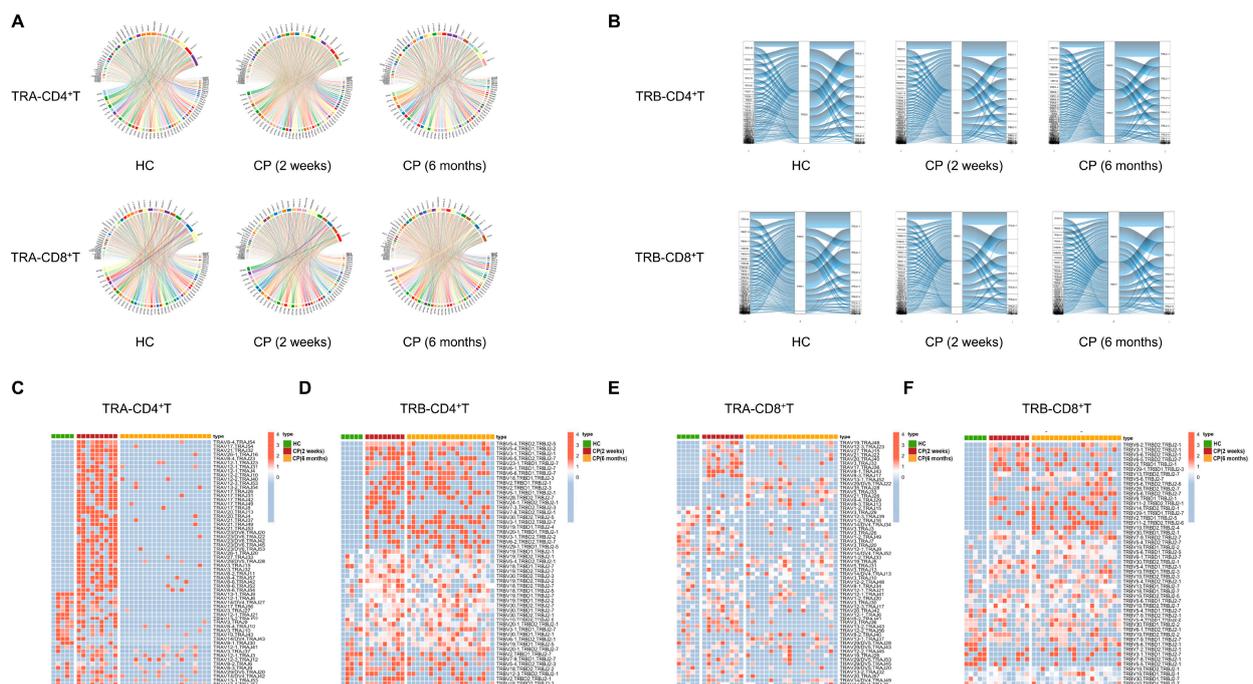
**Table 2.** The total number of TCR  $\alpha$  and  $\beta$  chain sequences were obtained.

	CD4 <sup>+</sup> T			CD8 <sup>+</sup> T			Total
	HC	CP (2 Weeks)	CP (6 Months)	HC	CP (2 Weeks)	CP (6 Months)	
<b>TRA</b>	3372	7298	654	4125	6670	10,677	32,796
<b>TRB</b>	951	2066	3916	805	1820	5528	15,086

HC, healthy controls; CP (2 weeks), COVID-19 patients with 2-week convalescence phase; CP (6 months), COVID-19 patients with 6-month convalescence phase; TRA, T cell receptor  $\alpha$  chain; TRB, T cell receptor  $\beta$  chain.



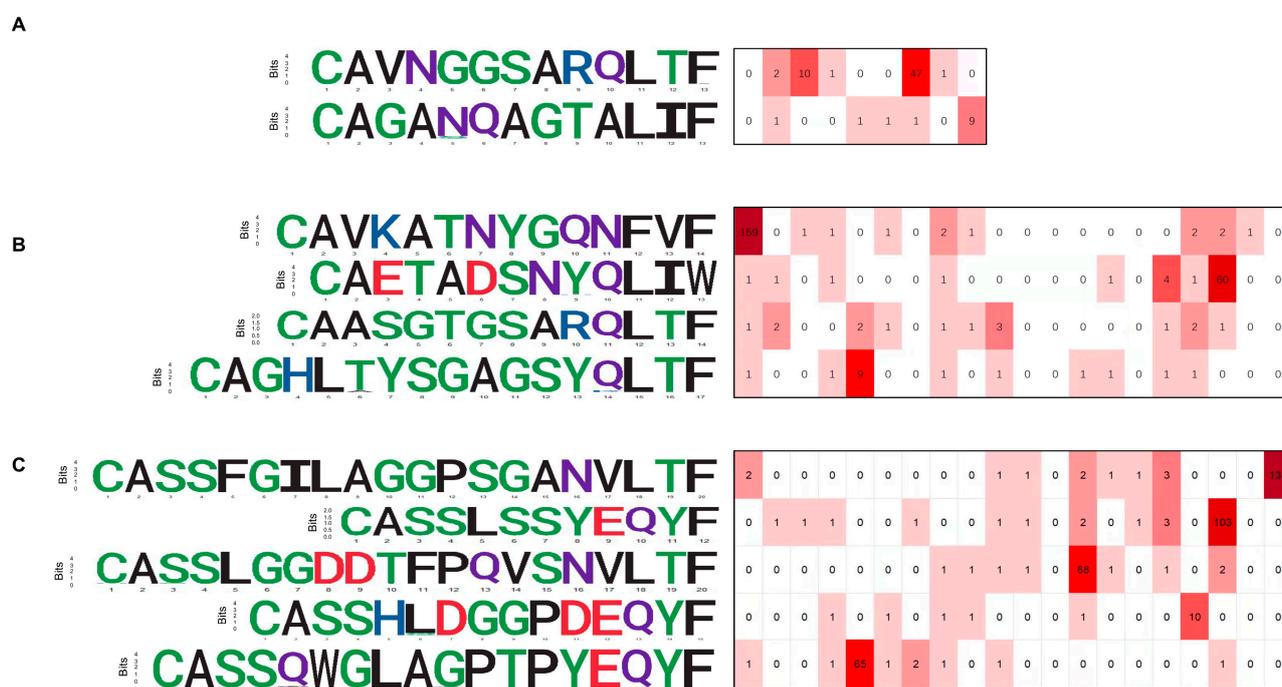
**Figure 2.** The types of TCR $\alpha$  (TRA) and TCR $\beta$  (TRB) clones from CD4<sup>+</sup> T and CD8<sup>+</sup> T cells of COVID-19 patients with 2-week and 6-month convalescence phase and healthy controls are shown (A–D). The TCR distribution TRA (A,B) and TRB (C,D) in CD4<sup>+</sup> T and CD8<sup>+</sup> T cells of convalescence COVID-19 patients (CP) and healthy controls (HC) are shown. Unique (n = 1), duplicated (n = 2), and clonal (n  $\geq$  3) TCRs are marked with different colors (E–H). The Venn diagram shows that the common and specific TRA (E,G) and TRB (F,H) numbers in CD4<sup>+</sup> T and CD8<sup>+</sup> T cells of convalescence COVID-19 patients (CP) and healthy controls (HC).



**Figure 3.** The V, D, and J gene usage comparison of TCR $\alpha$  (TRA) and TCR $\beta$  (TRB) chain in the convalescent COVID-19 patients. **(A)** The Circos plots show the difference between CD4<sup>+</sup> T cell TRA-VJ (top) and CD8<sup>+</sup> T cell TRA-VJ (bottom) among different COVID-19 patients with 2-week and 6-month convalescence phase (CP) and healthy controls (HC). **(B)** Sankey diagram shows the recombinant pattern of TRB VDJ in the CD4<sup>+</sup> T (top) and CD8<sup>+</sup> T cell (bottom) among different convalescence COVID-19 patients (CP) and healthy controls (HC). **(C–F)** A histogram of V(D)J gene recombinant pattern of TRA and TRB in CD4<sup>+</sup> T and CD8<sup>+</sup> T cells of COVID-19 patients (CP) with 2-week and 6-month-convalescence phase compared with healthy controls (HC). **(C)** TRA in CD4<sup>+</sup> T cells. **(D)** TRB in CD4<sup>+</sup> T cell. **(E)** TRA in CD8<sup>+</sup> T cells. **(F)** TRB in CD8<sup>+</sup> T cell.

### 3.3. CDR3 Sequence Motifs of Responding Clones

It was previously shown that TCRs that recognize the same antigen usually have highly similar TCR sequences [26], and the complementarity determining region 3 (CDR3) motif in the TCR component sequence plays an important role [27]. Moreover, the above data revealed potential SARS-CoV-2-specific and highly shared TCR recombination in CD4<sup>+</sup> T cells and CD8<sup>+</sup> T cells of convalescent COVID-19 patients. Therefore, we identified the dominant CDR3 sequences in each group. Due to the highly diverse nature of CDR3, only a few consensus sequences were identified, of which two CDR3 sequences of TRA were shared in CD8<sup>+</sup> T cells of two-week convalescent COVID-19 patients (Figure 4A), as well as four CDR3 sequences of TRA and five CDR3 sequences of TRB that were shared in CD8<sup>+</sup> T cells of six-month convalescent COVID-19 patients (Figure 4B,C). In contrast, no identical CDR3 sequences were found in all CD4<sup>+</sup> T cells of different convalescent COVID-19 patients. Compared with the published SARS-CoV-2-related TCR clusters, similar CDR3 motifs have not been reported, suggesting that these might be potential SARS-CoV-2-specific CDR3 motifs. Taken together, our results indicate that the convalescent COVID-19 patients had undergone distinct T cell responses during SARS-CoV-2 viral infection.



**Figure 4.** The unique TCR sequence motif and epitope specificity of CD8<sup>+</sup> T cell clones in COVID-19 patients with 2-week and 6-month convalescence phase. (A) The dominant CDR3 motif of TRA in COVID-19 patients with 2-week convalescence phase. (B) The dominant CDR3 motif of TRA in COVID-19 patients with 6-month convalescence phase. (C) The dominant CDR3 motif of TRB in COVID-19 patients with 6-month convalescence phase.

#### 4. Discussion

To date, the T cell immune regulation mechanism of COVID-19 is still unclear. In this study, we used 5' RACE combined with the third-generation high-throughput sequencing technology to confirm (a) the changes in TCR clones during the different convalescent phases of COVID-19 and the priority use of V and J gene fragments in COVID-19 patients compared with healthy controls; (b) there is persistent TCR dominant V(D)J recombination during the convalescent phase of COVID-19 from 2 weeks to 6 months; (c) the unique CDR3 motif specific for SARS-CoV-2 in CD8<sup>+</sup> T cells at the convalescent phase in COVID-19 patients. This finding is in line with other recent studies, where preferential usage of V and J gene segments in convalescent COVID-19 patients was found to be specific to SARS-CoV-2 antigens, suggesting a T cell-mediated immune response to achieve virus clearance. More important, we identified the dynamics of both CD4<sup>+</sup> T and CD8<sup>+</sup> T cell responses at different convalescent phases, showing the potential immune memory phenotype, which helps us understand the immune response induced by SARS-CoV-2 infection.

In recent years, TCR repertoire in various diseases was characterized using TCR-seq. Multiplex PCR-, target enrichment-, and 5' RACE (RNA only)-based approaches are the most widely used TCR-seq strategies. When using multiplex PCR, preferential amplification of highly abundant gene products can occur, contributing to inaccuracies in reported clone frequencies. Additionally, intronic sequences and incomplete VDJ recombination products may contribute to data noise when using gDNA samples. Target enrichment strategies reduce the extent of amplification bias seen in standard multiplex PCR but are more labor/time-intensive and do not resolve the issues associated with gDNA input [28]. However, the 5' RACE approach avoids the PCR bias seen in the above two methods. It achieves equally efficient amplification of all TCR transcripts regardless of abundance when combined with sequencing platforms capable of long reads, such as the illumine Miseq, as fragment size can exceed 600 bp [29]. Currently, single-cell TCR-seq in tandem with single-cell flow cytometry sorting is an approach for precisely identifying SARS-CoV-2-

specific T cells, while it can be limited by their depth and their ability to capture minor clones. Therefore, in our study, we applied 5' RACE technology combined with PacBio Sequel system capable of long reads (10,000 bp) as third-generation sequencing technology to obtain a wealth of TCR sequences, which effectively avoided the sequence tendency caused by traditional multiple PCR and retained the accuracy of sequencing to the greatest extent. Therefore, the method we adopted more truly and comprehensively demonstrated the characteristics of TCR.

The TCR is composed of TCR $\alpha$  and TCR $\beta$  chains, which containing VJ and VDJ gene segments, separately [30]. The results of single-cell sequencing by Wang et al. have reported that the highest frequency of TCR sequence recombination during the convalescent phase of COVID-19 is TRAV12-2-J27-TRBV7-9-J2-3 [25]. Alina et al. had found that TRAV12-1 and TRBV7-9 were used by 71% and 16% of SARS-CoV-2 YLQ-epitope-specific TCRs, and TRAV13-2 and TRBV6-5 were used by 15% and 25% of SARS-CoV-2 RLQ-epitope-specific TCRs, compared with just 3–4% gene usage in control TCRs [4]. Similarly, our results also show the preferential V(D)J sequences only existing in the convalescent COVID-19 patients, including TRAV8-4/TRAJ54 and TRAV17/TRAJ54, TRBV5-4/D2/TRBJ2-5, and TRBV5-4/D1/TRBJ2-2, which were different from public SARS-CoV-2-specific TCRs. Combinations of TCR $\alpha$  and TCR $\beta$  genes were highly sample-specific between different COVID-19 patients and healthy controls. A larger population cohort is urgently needed to provide more statistical T cell receptor preference. Interestingly, our results also revealed that identical VDJ sequences, such as TRBV6-5/D2/TRBJ2-7 and TRBV2/D1/TRBJ2-1, were shared by CD4<sup>+</sup> T cells and CD8<sup>+</sup> T cells. Although CD4<sup>+</sup> T and CD8<sup>+</sup> T cells exhibit their roles differently, CD8<sup>+</sup> T cells recognize and kill virus-infected cells through TCR-mediated viral antigens [31], while CD4<sup>+</sup> T cells have multiple roles in coordinating and mediating immune responses against viruses [32]. Therefore, our data suggest that cross-reactive CD4<sup>+</sup> T and CD8<sup>+</sup> T cells can participate in the joint resistance to SARS-CoV-2 infection via the production of a virus-specific TCR recombination pattern.

In the process of V(D)J recombination, random nucleotide deletions and insertions often occur in the CDR3, which is essential for antigen binding [33]. Therefore, this process produces a large number of recombinant TCRs, and the differences in CDR3 sequences are usually used to characterize essential indicators of the immune repertoire. We also identified a few consensus CDR3 sequences shared in CD8<sup>+</sup> T but not CD4<sup>+</sup> T cells of convalescent COVID-19 patients. Moreover, no identical CDR3 motifs have not been reported relative to the published SARS-CoV-2-related TCR clusters. Our results suggest that these CD8<sup>+</sup> T cells with unique SARS-CoV-2-associated CDR3 motifs might play a crucial role in the immune response to eliminate the SARS-CoV-2 virus efficiently. However, these unique T cell sequence motifs require prospective validation to be used in COVID-19 patients, and specific T cells reactive to SARS-CoV-2 epitopes also need to be identified in the future. Therefore, our TCR data might help to fingerprint a shared clonal expression T cell phenotype that existed within a population of COVID-19 patients and that defined anti-SARS-CoV-2 immunity in individuals.

Furthermore, T cells with long-lasting memory have been detected in patients recovered from SARS-CoV-2 infection [7], and SARS-CoV-2-specific T cells exhibit a multifunctional memory phenotype during the convalescent phase [34]. In our study, we also found that some particular TRB VDJ recombination patterns in CD4<sup>+</sup> T and CD8<sup>+</sup> T cells, such as TRBV5-4/D2/TRBJ2-5 and TRBV5-4/D1/TRBJ2-2, appear at the 2-week convalescent stage and continue to the 6-month convalescent stage, supporting that these T cell clones might have long-term memory characteristics after SARS-CoV-2 infection to participate in the elimination of residual SARS-CoV-2 in the convalescence stage.

## 5. Conclusions

Our findings reveal that COVID-19 patients have unique cellular immune characteristics during the different convalescent phases, which contributes to our understanding of

the immune response induced by SARS-CoV-2 and might be used as a basis for prognostic evaluation and targeted therapy for COVID-19 patients.

**Author Contributions:** J.H., Y.L. and X.Q. designed the project. Y.W., F.D. and Z.Z. performed experiments and analyzed data. M.Y., X.J., H.D. and P.W. assisted experiments and data interpretation. J.H. and F.D. wrote the manuscript. X.Q. provided funding support. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the China Ministry of Science and Technology (2020YFA0707801) and the Discipline Construction Project of Peking University Health Science Center (BMU2020HKYZX004).

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and the study was approved by the Institutional Review Board of Peking University (PUIRB). We have complied with all relevant ethical regulations, and informed consent was obtained from all human participants (No:2020072D).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Wu, Z.; McGoogan, J.M. Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases from the Chinese Center for Disease Control and Prevention. *JAMA* **2020**, *323*, 1239–1242. [[CrossRef](#)] [[PubMed](#)]
2. Guan, W.J.; Ni, Z.Y.; Hu, Y.; Liang, W.H.; Ou, C.Q.; He, J.X.; Liu, L.; Shan, H.; Lei, C.L.; Hui, D.S.C.; et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* **2020**, *382*, 1708–1720. [[CrossRef](#)]
3. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497–506. [[CrossRef](#)]
4. Shomuradova, A.S.; Vagida, M.S.; Sheetikov, S.A.; Zornikova, K.V.; Kiryukhin, D.; Titov, A.; Peshkova, I.O.; Khmelevskaya, A.; Dianov, D.V.; Malasheva, M.; et al. SARS-CoV-2 Epitopes Are Recognized by a Public and Diverse Repertoire of Human T Cell Receptors. *Immunity* **2020**, *53*, 1245–1257.e5. [[CrossRef](#)] [[PubMed](#)]
5. Nelde, A.; Bilich, T.; Heitmann, J.S.; Maringer, Y.; Salih, H.R.; Roerden, M.; Lübke, M.; Bauer, J.; Rieth, J.; Wacker, M.; et al. SARS-CoV-2-derived peptides define heterologous and COVID-19-induced T cell recognition. *Nat. Immunol.* **2021**, *22*, 74–85. [[CrossRef](#)]
6. Braun, J.; Loyal, L.; Frensch, M.; Wendisch, D.; Georg, P.; Kurth, F.; Hippenstiel, S.; Dingeldey, M.; Kruse, B.; Fauchere, F.; et al. SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* **2020**, *587*, 270–274. [[CrossRef](#)]
7. Le Bert, N.; Tan, A.T.; Kunasegaran, K.; Tham, C.Y.L.; Hafezi, M.; Chia, A.; Chng, M.H.Y.; Lin, M.; Tan, N.; Linster, M.; et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **2020**, *584*, 457–462. [[CrossRef](#)] [[PubMed](#)]
8. Rosendahl Huber, S.; van Beek, J.; de Jonge, J.; Luytjes, W.; van Baarle, D. T cell responses to viral infections—opportunities for Peptide vaccination. *Front. Immunol.* **2014**, *5*, 171. [[CrossRef](#)] [[PubMed](#)]
9. Rouse, B.T.; Sehrawat, S. Immunity and immunopathology to viruses: What decides the outcome? *Nat. Rev. Immunol.* **2010**, *10*, 514–526. [[CrossRef](#)]
10. Altmann, D.M.; Boyton, R.J. SARS-CoV-2 T cell immunity: Specificity, function, durability, and role in protection. *Sci. Immunol.* **2020**, *5*, eabd6160. [[CrossRef](#)] [[PubMed](#)]
11. Snyder, T.M.; Gittelman, R.M.; Klinger, M.; May, D.H.; Osborne, E.J.; Taniguchi, R.; Zahid, H.J.; Kaplan, I.M.; Dines, J.N.; Noakes, M.N.; et al. Magnitude and Dynamics of the T-Cell Response to SARS-CoV-2 Infection at Both Individual and Population Levels. *medRxiv* **2020**, 20165647. [[CrossRef](#)]
12. Thevarajan, I.; Nguyen, T.H.O.; Koutsakos, M.; Druce, J.; Caly, L.; van de Sandt, C.E.; Jia, X.; Nicholson, S.; Catton, M.; Cowie, B.; et al. Breadth of concomitant immune responses prior to patient recovery: A case report of non-severe COVID-19. *Nat. Med.* **2020**, *26*, 453–455. [[CrossRef](#)] [[PubMed](#)]
13. Su, Y.; Chen, D.; Yuan, D.; Lausted, C.; Choi, J.; Dai, C.L.; Voillet, V.; Duvvuri, V.R.; Scherler, K.; Troisch, P.; et al. Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19. *Cell* **2020**, *183*, 1479–1495.e20. [[CrossRef](#)] [[PubMed](#)]
14. Peng, Y.; Mentzer, A.J.; Liu, G.; Yao, X.; Yin, Z.; Dong, D.; Dejnirattisai, W.; Rostron, T.; Supasa, P.; Liu, C.; et al. Broad and strong memory CD4(+) and CD8(+) T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat. Immunol.* **2020**, *21*, 1336–1345. [[CrossRef](#)]

15. Notarbartolo, S.; Ranzani, V.; Bandera, A.; Gruarin, P.; Bevilacqua, V.; Putignano, A.R.; Gobbin, A.; Galeota, E.; Manara, C.; Bombaci, M.; et al. Integrated longitudinal immunophenotypic, transcriptional and repertoire analyses delineate immune responses in COVID-19 patients. *Sci. Immunol.* **2021**, *6*, eabg5021. [[CrossRef](#)] [[PubMed](#)]
16. Neidleman, J.; Luo, X.; Frouard, J.; Xie, G.; Gill, G.; Stein, E.S.; McGregor, M.; Ma, T.; George, A.F.; Kosters, A.; et al. SARS-CoV-2-Specific T Cells Exhibit Phenotypic Features of Helper Function, Lack of Terminal Differentiation, and High Proliferation Potential. *Cell Rep. Med.* **2020**, *1*, 100081. [[CrossRef](#)]
17. Grifoni, A.; Weiskopf, D.; Ramirez, S.I.; Mateus, J.; Dan, J.M.; Moderbacher, C.R.; Rawlings, S.A.; Sutherland, A.; Premkumar, L.; Jadi, R.S.; et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **2020**, *181*, 1489–1501.e15. [[CrossRef](#)]
18. Zhang, S.; Mao, Y.; Huang, J.; Ma, T.; Zhang, L.; Zhu, X.; Zheng, J.; Wu, L.; Yin, C.C.; Qiu, X. Immunoglobulin gene locus events in epithelial cells of lactating mouse mammary glands. *Cell. Mol. Life Sci.* **2010**, *67*, 985–994. [[CrossRef](#)]
19. Dash, P.; Fiore-Gartland, A.J.; Hertz, T.; Wang, G.C.; Sharma, S.; Souquette, A.; Crawford, J.C.; Clemens, E.B.; Nguyen, T.H.O.; Kedzierska, K.; et al. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature* **2017**, *547*, 89–93. [[CrossRef](#)] [[PubMed](#)]
20. Furman, D.; Davis, M.M. New approaches to understanding the immune response to vaccination and infection. *Vaccine* **2015**, *33*, 5271–5281. [[CrossRef](#)] [[PubMed](#)]
21. Wang, C.; Sanders, C.M.; Yang, Q.; Schroeder, H.W., Jr.; Wang, E.; Babrzadeh, F.; Gharizadeh, B.; Myers, R.M.; Hudson, J.R., Jr.; Davis, R.W.; et al. High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1518–1523. [[CrossRef](#)] [[PubMed](#)]
22. Schultheiß, C.; Paschold, L.; Simnica, D.; Mohme, M.; Willscher, E.; von Wenserski, L.; Scholz, R.; Wieters, I.; Dahlke, C.; Tolosa, E.; 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*, 442–455.e4. [[CrossRef](#)] [[PubMed](#)]
23. Liao, M.; Liu, Y.; Yuan, J.; Wen, Y.; Xu, G.; Zhao, J.; Cheng, L.; Li, J.; Wang, X.; Wang, F.; et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **2020**, *26*, 842–844. [[CrossRef](#)] [[PubMed](#)]
24. Zhang, F.; Gan, R.; Zhen, Z.; Hu, X.; Li, X.; Zhou, F.; Liu, Y.; Chen, C.; Xie, S.; Zhang, B.; et al. Adaptive immune responses to SARS-CoV-2 infection in severe versus mild individuals. *Signal Transduct. Target. Ther.* **2020**, *5*, 156. [[CrossRef](#)]
25. Wang, P.; Jin, X.; Zhou, W.; Luo, M.; Xu, Z.; Xu, C.; Li, Y.; Ma, K.; Cao, H.; Huang, Y.; et al. Comprehensive analysis of TCR repertoire in COVID-19 using single cell sequencing. *Genomics* **2020**, *113*, 456–462. [[CrossRef](#)] [[PubMed](#)]
26. Minervina, A.; Pogorelyy, M.; Mamedov, I. T-cell receptor and B-cell receptor repertoire profiling in adaptive immunity. *Transpl. Int.* **2019**, *32*, 1111–1123. [[CrossRef](#)]
27. Rosati, E.; Dowds, C.M.; Liaskou, E.; Henriksen, E.K.K.; Karlsen, T.H.; Franke, A. Overview of methodologies for T-cell receptor repertoire analysis. *BMC Biotechnol.* **2017**, *17*, 61. [[CrossRef](#)]
28. Lin, Y.H.; Hung, S.J.; Chen, Y.L.; Lin, C.H.; Kung, T.F.; Yeh, Y.C.; Tseng, J.T.; Liu, T. Dissecting efficiency of a 5' rapid amplification of cDNA ends (5'-RACE) approach for profiling T-cell receptor beta repertoire. *PLoS ONE* **2020**, *15*, e0236366. [[CrossRef](#)]
29. Yeku, O.; Frohman, M.A. Rapid amplification of cDNA ends (RACE). *Methods Mol. Biol.* **2011**, *703*, 107–122. [[PubMed](#)]
30. Nishana, M.; Raghavan, S.C. Role of recombination activating genes in the generation of antigen receptor diversity and beyond. *Immunology* **2012**, *137*, 271–281. [[CrossRef](#)]
31. Andersen, M.H.; Schrama, D.; Thor Straten, P.; Becker, J.C. Cytotoxic T Cells. *J. Investig. Dermatol.* **2006**, *126*, 32–41. [[CrossRef](#)] [[PubMed](#)]
32. Sant, A.J.; McMichael, A. Revealing the role of CD4(+) T cells in viral immunity. *J. Exp. Med.* **2012**, *209*, 1391–1395. [[CrossRef](#)] [[PubMed](#)]
33. La Gruta, N.L.; Gras, S.; Daley, S.R.; Thomas, P.G.; Rossjohn, J. Understanding the drivers of MHC restriction of T cell receptors. *Nat. Rev. Immunol.* **2018**, *18*, 467–478. [[CrossRef](#)] [[PubMed](#)]
34. Sekine, T.; Perez-Potti, A.; Rivera-Ballesteros, O.; Strålin, K.; Gorin, J.B.; Olsson, A.; Llewellyn-Lacey, S.; Kamal, H.; Bogdanovic, G.; Muschiol, S.; et al. Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19. *Cell* **2020**, *183*, 158–168.e14. [[CrossRef](#)] [[PubMed](#)]