

Original Research

# B-Cell Receptor Features and Database Establishment in Recovered COVID-19 Patients by Combining 5'-RACE with PacBio Sequencing

Zhu Zhu<sup>1,2,3,†</sup>, Pingzhang Wang<sup>1,2,3,†</sup>, Xiaodong Jia<sup>4</sup>, Meng Yu<sup>1,2,3</sup>, Huige Yan<sup>1,2,3</sup>, Lei Liu<sup>5</sup>, Wanbing Liu<sup>5</sup>, Yaqiong Zheng<sup>5</sup>, Guomei Kou<sup>5</sup>, Jie Wang<sup>6</sup>, Weiyan Xu<sup>1,2,3</sup>, Jing Huang<sup>1,2,3</sup>, Fugang Duan<sup>1,2,3</sup>, Fengmin Lu<sup>6</sup>, Ning Fu<sup>7</sup>, Ning Zhang<sup>8</sup>, Yingying Lu<sup>4</sup>, Hui Dai<sup>1,2,3</sup>, Shangen Zheng<sup>5</sup>, Xiaoyan Qiu<sup>1,2,3,\*</sup>

<sup>1</sup>NHC Key Laboratory of Medical Immunology, Peking University, 100191 Beijing, China

<sup>2</sup>Key Laboratory of Molecular Immunology, Chinese Academy of Medical Sciences, 100191 Beijing, China

<sup>3</sup>Department of Immunology, School of Basic Medical Sciences, Peking University, 100191 Beijing, China

<sup>4</sup>Department of Hepatology, Fifth Medical Center of Chinese PLA General Hospital, 100191 Beijing, China

<sup>5</sup>Department of Transfusion Medicine, General Hospital of Central Theater Command of PLA, 430000 Wuhan, Hubei, China

<sup>6</sup>Department of Microbiology and Infectious Disease Center, School of Basic Medical Sciences, Peking University Health Science Center, 100191 Beijing, China

<sup>7</sup>Department of Immunology, School of Basic Medical Science, Southern Medical University, 510515 Guangzhou, Guangdong, China

<sup>8</sup>Peking University International Cancer Institute, Health Science Center, Peking University, 100191 Beijing, China

\*Correspondence: [qiuxy@bjmu.edu.cn](mailto:qiuxy@bjmu.edu.cn) (Xiaoyan Qiu)

†These authors contributed equally.

Academic Editor: Josef Jampilek

Submitted: 21 April 2022 Revised: 14 September 2022 Accepted: 26 September 2022 Published: 28 February 2023

## Abstract

**Background:** Antibodies induced by viral infection can not only prevent subsequent virus infection, but can also mediate pathological injury following infection. Therefore, understanding the B-cell receptor (BCR) repertoire of either specific neutralizing or pathological antibodies from patients convalescing from Coronavirus disease 2019 (COVID-19) infection is of benefit for the preparation of therapeutic or preventive antibodies, and may provide insight into the mechanisms of COVID-19 pathological injury. **Methods:** In this study, we used a molecular approach of combining 5' Rapid Amplification of cDNA Ends (5'-RACE) with PacBio sequencing to analyze the BCR repertoire of all 5 *IgH* and 2 *IgL* genes in B-cells harvested from 35 convalescent patients after severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. **Results:** We observed numerous BCR clonotypes within most COVID-19 patients, but not in healthy controls, which validates the association of the disease with a prototypical immune response. In addition, many clonotypes were found to be frequently shared between different patients or different classes of antibodies. **Conclusions:** These convergent clonotypes provide a resource to identify potential therapeutic/prophylactic antibodies, or identify antibodies associated with pathological effects following infection with SARS-CoV-2.

**Keywords:** COVID-19; BCR repertoire; 5'RACE; PacBio sequencing; neutralizing antibody

## 1. Introduction

The 2019 coronavirus disease (COVID-19) pandemic was caused by severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2). As of April 2022, SARS-CoV-2 has infected nearly 5 billion people in more than 185 countries and killed more than 6 million (see <https://covid19.who.int/>). Although a total of 112 billion vaccines have been administered globally, no vaccine can induce a 100% protective effect. Moreover, there is, to date, a lack of a long-term clinical evaluation of the effect of various SARS-CoV-2 vaccines [1–4]. Therefore, there is urgent need to develop more effective anti-SARS-CoV-2 therapeutics, such as effective and safe neutralizing antibodies, to be used for COVID-19 prevention and adoptive treatment.

The adaptive immune system is a core protective mechanism against viral infection. Humoral immunity, me-

diated by B-cells, can provide rapid and lasting immune protection against pathogens, such as viruses, by producing antibodies and memory B-cells [5]. When a pathogen, such as SARS-CoV-2, enters the body, a small group of B-cells recognize antigens through the B-cell receptor (BCR) on their surface. With the help of CD4<sup>+</sup> T-cells, the B-cells are activated and undergo clonal expansion, and, following this, some rapidly differentiate into plasma cells that secrete antibodies while others stop differentiation and become memory B-cells which are critical in responding to subsequent exposure to the same antigen [6,7].

During the process of B-cell clonal expansion, two events occur at the gene level. One of these events is a somatic high-frequency mutation within the BCR variable region gene which improves antibody affinity. The second event is antibody class conversion where the ex-



pression of IgM-class immunoglobulin is changed to expression of IgG, IgA, or IgE-class immunoglobulins. This immunoglobulin class switching enhances the immune response by allowing the body to take advantage of various biological effects inherent in different immunoglobulin classes [8,9]. Therefore, antibodies produced in response to viral infection commonly have the characteristics of high efficiency and affinity. At present, although some technologies, such as screening phage libraries, can quickly identify antibodies directed against SARS-CoV-2, using these approaches it is difficult to obtain high affinity antibodies due to the absence of antigen-induced affinity maturation [10–12]. Therefore, from a therapeutic standpoint, the best choice to obtain specific antibodies and gene sequences that encode these antibodies is by the analysis of serum or memory B-cells from individuals infected with SARS-CoV-2.

Antibodies often present a double-edged sword during immune response. One hand, B-cells activated to produce pathogen-specific antibodies following viral infection play an important immuno-protective role. In contrast, some virus-specific antibodies can also enhance viral infection through a process known as antibody-dependent enhancement (ADE). For example, if the titer of virus-specific antibodies in severe COVID-19 patients is higher than that measured in healthy or patients with mild symptoms, it is possible that this may signal that some antibodies also promote harmful effects. Additionally, endogenous antigens released following tissue injury can also activate B-cells to produce autoantibodies. Thus, the collective library of antibodies in COVID-19 patients contains both beneficial and harmful antibodies, and each will have unique variable region sequences arising in B-cell clones.

At present, the antibody sequences expressed by different B-cells can be obtained by analyzing the entire BCR repertoire within a population of B-cells. Prototypical antibody quaternary structure is comprised of 4 peptide chains, specifically, two covalently-linked identical heavy chains and two identical light chains. Both the heavy and light chains are divided into variable and constant regions and the variable region of the heavy and the light chains, together, determine the antigen-binding specificity of the antibody. There is no complete reading frame in the variable region locus of immunoglobulin genes as this region is composed of many gene segments. The human heavy chain locus includes 65 functional variable (V) segments, 27 functional diversity (D) segments, and 6 functional joining (J) segments. In contrast, the human light chains are encoded from two distinct loci, and are termed the  $\kappa$  and  $\lambda$  genes. There are 40 functional V segments, and 5 functional J segments in the  $\kappa$  chain gene, and 30 functional V segments and 5 functional J segments in the  $\lambda$  chain gene. To produce appropriate antigen-binding sequences, the variable region of the heavy chain gene undergoes rearrangement among the V, D and J segments. Following this, the variable region of a light chain gene is randomly rearranged by splicing

one V segment to one J segment. These two events form the complete reading frame of heavy chain and light chain variable region, respectively [13]. Traditionally, the BCR on the surface of each B-cell also has a unique rearrangement pattern such that each B-cell has a unique specificity of antigen recognition, and this leads to vast diversity in antibodies that can be made by an individual. However, after viral infection, some disease-specific B-cells will clonally proliferate resulting in effector or memory B-cells expressing the same BCR variable region which can be detected in peripheral blood.

Recently, numerous SARS-CoV-2-specific BCR sequences were obtained by single cell sequencing and multiple PCR-related next generation sequencing approaches [14–17]. The analysis of the SARS-CoV-2-related BCR repertoire was mainly focused on total Ig, or IgG and IgA [18]. However, this approach overlooks the fact that all classes of antibodies have immune defense effects and, moreover, different types of BCR/Ig variable regions have unique rearrangement patterns and antigen specificity [19]. This latter point further reduces the specific antibody pool that can be produced. These recent studies also mainly focused on the V region but not the entire variable region which includes V as well as D and J regions. However, V(D)J recombination, somatic hypermutation, and diversity of junctions between V, D and J gene segments drive antibody diversity that recognize different antigens.

In this study, we used the 5' Rapid Amplification of cDNA Ends (5'-RACE) combined with high-throughput PacBio sequencing to investigate the BCR repertoire of 4 classes of heavy chains and the 2 types of light chains in peripheral blood CD19<sup>+</sup> B-cells harvested from convalescent patients after SARS-CoV-2 infection, and also compare the characteristics of the BCR repertoire in patients at different stages of convalescence.

## 2. Materials and Methods

### 2.1 COVID-19 Convalescent Patients and Healthy Donors

15 confirmed COVID-19 patients of 2 weeks convalescence were obtained from Fifth Medical Center of Chinese PLA General Hospital, Beijing, China. The clinical information of these patients was listed in **Supplementary Table 1**. 20 confirmed COVID-19 patients of 6 months convalescence were obtained from General Hospital of Central Theater Command of PLA, Wuhan, China.

### 2.2 ELISA Quantification

ELISA plates were coated with S1 protein at 0.01 mg/mL in coating buffer (0.1 mol/L Na<sub>2</sub>CO<sub>3</sub>: 0.1 mol/L NaHCO<sub>3</sub> = 3:7) at 4 °C overnight. After standard washing and blocking, diluted sera (1:100) were applied to each well. After a 1 h incubation at 37 °C, plates were washed and incubated with 0.25 μg/mL goat anti-human Ig antibodies conjugated with HRP (SouthernBiotech, Birmingham, AL, USA) for 1 h at 37 °C. TMB was used as the substrate,

and the reaction was ceased by 2 mol/L H<sub>2</sub>SO<sub>4</sub>. The absorbance at 450 nm was measured by a microplate reader.

### 2.3 Pseudovirus Neutralization Assay

The pseudovirus neutralization assays were performed using HEK293-hACE2 cell lines, which express human ACE2.  $5 \times 10^4$ /well cells were added to the well and cultured at 37 °C for 8 h. Serum diluted at 1:150 with PBS was incubated with TCID<sub>50</sub> of  $5 \times 10^4$  TU SARS-CoV-2 pseudovirus (purchased from Future Biotherapeutics, Suzhou, Jiangsu, China) at 37 °C for 1 h. Then the mixed solution of virus and serum was added to the cell and cultured for 18 h at 37 °C supplied with 5% CO<sub>2</sub>. The cells were harvest and washed with PBS and detected by flow cytometry.

### 2.4 Enrichment of B Cells from PBMC

B cells were isolated from fresh or previously frozen PBMCs by immunomagnetic positive selection according to the manufacturer's protocol (EasySep™ Human CD19 Positive Selection Kit II, STEMCELL). Purified B cells were eluted and washed with PBS containing 2% (v/v) fetal bovine serum (FBS) and 1 mM/L EDTA.

### 2.5 RNA Extraction and cDNA Synthesized by 5'-RACE

According to the manufacturer's instructions for sorted B cells, total RNA was extracted using the Rapture Total RNA Micro Kit (Magen, R4012-02, Guangzhou, Guangdong, China), then the cDNA was synthesized by 5'-RACE using SMARTer® RACE 5'/3' Kit (Takara, 634859, Mountain View, CA, USA) and generated a complete cDNA copy with the additional specific sequence at the 5' end (universal linker sequence).

### 2.6 Amplification of BCR Transcripts with Barcoded Primers

The upstream primer was targeted on the universal linker sequence from 5'-RACE, and the downstream primer targeted constant-region for *IGHG*, *IGHA*, *IGHM*, *IGHD*, *IGK*, and *IGL* genes in both of first-round and second-round PCR. Especially, barcodes were added to the second-round PCR primers that is convenient to distinguish BCRs from different individuals. PCR program for both rounds were: 5 cycles at 94 °C for 30 s, 72 °C for 3 min, 5 cycles at 94 °C for 30 s, 70 °C for 30 s, 72 °C for 3 min, and 25 cycles at 94 °C for 30 s, 68 °C for 30 s, 72 °C for 3 min (first-round PCR) and 5 cycles at 94 °C for 30 s, 72 °C for 90 s, 5 cycles at 94 °C for 30 s, 70 °C for 30 s, 72 °C for 90 s, and 25 cycles at 94 °C for 30 s, 68 °C for 30 s, 72 °C for 90 s (second-round PCR). The amplified DNA products were recovered from the agarose gel using a DNA Recovery Kit and sent to Novogene company for sequencing.

In addition, PCR was also performed to detect the expression of GAPDH. PCR condition for both cycles were: 95 °C for 5 min, 30 cycles at 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s, followed by a final extension at 72

°C for 7 min. The PCR products were separated on 1.5% agarose gel by electrophoresis. The primers used for PCR are listed in **Supplementary Tables 3,4,5**.

### 2.7 PacBio Sequencing and Data Analysis

We measured the concentration of PCR products with barcode and mixed the same amount of PCR products from COVID-19 patients (products of COVID-19 patients of 2 weeks convalescence were mixed as one sample while products of COVID-19 patients of 6 months convalescence were mixed as another one sample). The mixed PCR products were separated with 2% agarose gel. Gel extraction was performed and products were used for Pacbio sequencing. The obtained data was split according to barcodes by using the FASTX-toolkit (version 0.0.13, LongIsland, NY, USA). IMG/High V-QUEST (version 1.7.1, Montpellier, France) was used for sequence annotation to determine the V(D)J genes and sequence of the variable region.

### 2.8 Statistical Analysis

All data were analyzed by normality and lognormality tests to identify whether the data belong to a normal distribution, which was decided by the *p*-value of the Shapiro-Wilk test. Unpaired *t*-test was used in the condition of normal distribution, or Mann-Whitney test was used in the non-normal distribution (\* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.005, \*\*\*\* *p* < 0.0001). These are all executed in Graph-Pad Prism.

## 3. Results

### 3.1 Spike Protein-Specific Antibodies with Neutralizing Activity were Detectable in Recovered COVID-19 Patients 2 Weeks and 6 Months Post-Infection

COVID-19 patients at 2 weeks of convalescence were discharged from Fifth Medical Center of Chinese PLA General Hospital from February 13 to March 3, 2020 were used in this study. Eligible participants consisted of 7 female and 7 male patients who ranged from age 39 to 74, with a median of 64 years of age. All COVID-19 patients were confirmed to be SARS-CoV-2 infected by respiratory RT-PCR tests. According to the guidelines for diagnosis and management of COVID-19 released by National Health Commission of China, discharge criteria were listed as follows: (1) Normal body temperature for more than three days; (2) Significant improvement in respiratory symptoms; (3) Pulmonary imaging shows a substantial reduction in pneumonia; (4) Two consecutively negative SARS-CoV-2 RT-PCR assays of respiratory tract specimens with a one-day sampling interval. A peripheral blood sample was collected in a BD Vacutainer and the plasma was separated by centrifugation at 2000 rpm for 10 min (**Supplementary Table 1**).

COVID-19 patients with 6 months convalescence were discharged from General Hospital of Central Theater Command of PLA, Wuhan, China.

The serum samples of COVID-19 patients 2 weeks, and 6 months, after recovery were obtained, and the level of specific IgM, IgG and IgA against the spike protein of SARS-CoV-2 were measured by ELISA. Serum from 27 healthy subjects were used as negative controls. The results showed a higher level of specific IgM and IgA against SARS-CoV-2 spike protein in recovered COVID-19 patients 2 weeks when compared to patients at the 6 month mark. However, the IgG level was elevated in COVID-19 patients 6 months after recovery, while no differences were observed in viral-specific IgM and IgA levels between COVID-19 patients 6 months after recovery and healthy controls (Fig. 1A). Subsequently, serum samples from 26 COVID-19 patients after 2 weeks of convalescence and 139 COVID-19 patients after 6 months of convalescence were used to analyze neutralization activity by blocking SARS-CoV-2 pseudovirus infection of HEK293 cells overexpressing human ACE2. The results showed that 27% of 2 weeks convalescent COVID-19 patients and 37.4% of the 6 months convalescent COVID-19 patients showed more than 25% blocking activity. Moreover, 20% of 6 months convalescent COVID-19 patients displayed greater than 50% blocking activity (Fig. 1B). Of note, approximately 42% of 6 months convalescent COVID-19 patients showed elevated binding of the pseudovirus with HEK293 cells overexpressing human ACE2; however, the reason for this observation is not clear. We also purified IgA and IgG from the 2-week convalescent serum and analyzed the neutralizing activity. Results showed that, 84.4% of IgG and 71.9% of IgA taken from these COVID-19 patients showed more than 25% blocking activity, and 65.6% of IgG and 50% of IgA in COVID-19 patients showed more than 50% blocking activity (Fig. 1C).

### 3.2 Study Design for Analysis of the BCR Repertoire of COVID-19 Patients

Viral infection can trigger virus-specific memory B-cell production, and this response can linger longer in the body post-infection. While different from naive B-cells, memory B-cells can express not only virus-specific IgM and IgD immunoglobulins, but also highly effective, high affinity virus-specific IgA, IgG or IgE antibodies [4,20,21].

We first collected the peripheral Blood Mononuclear Cell (PBMC) from 35 COVID-19 donors at different stages of recovery, including 15 donors 2 weeks after recovery, 10 donors with neutralizing antibodies 6 months after recovery, 10 donors without neutralizing antibodies 6 months after recovery, and 11 healthy donors were used as controls. CD19<sup>+</sup> B-cells were sorted via immunomagnetic beads, and total RNA were extracted from these cells. Considering the diversity of BCR variable regions and a biased amplification problem caused by common use of multiple RT-PCR steps, we used a 5'-RACE reverse transcription kit to generate a complete cDNA copy with additional specific sequence at the 5' end of the amplicon. We designed the up-

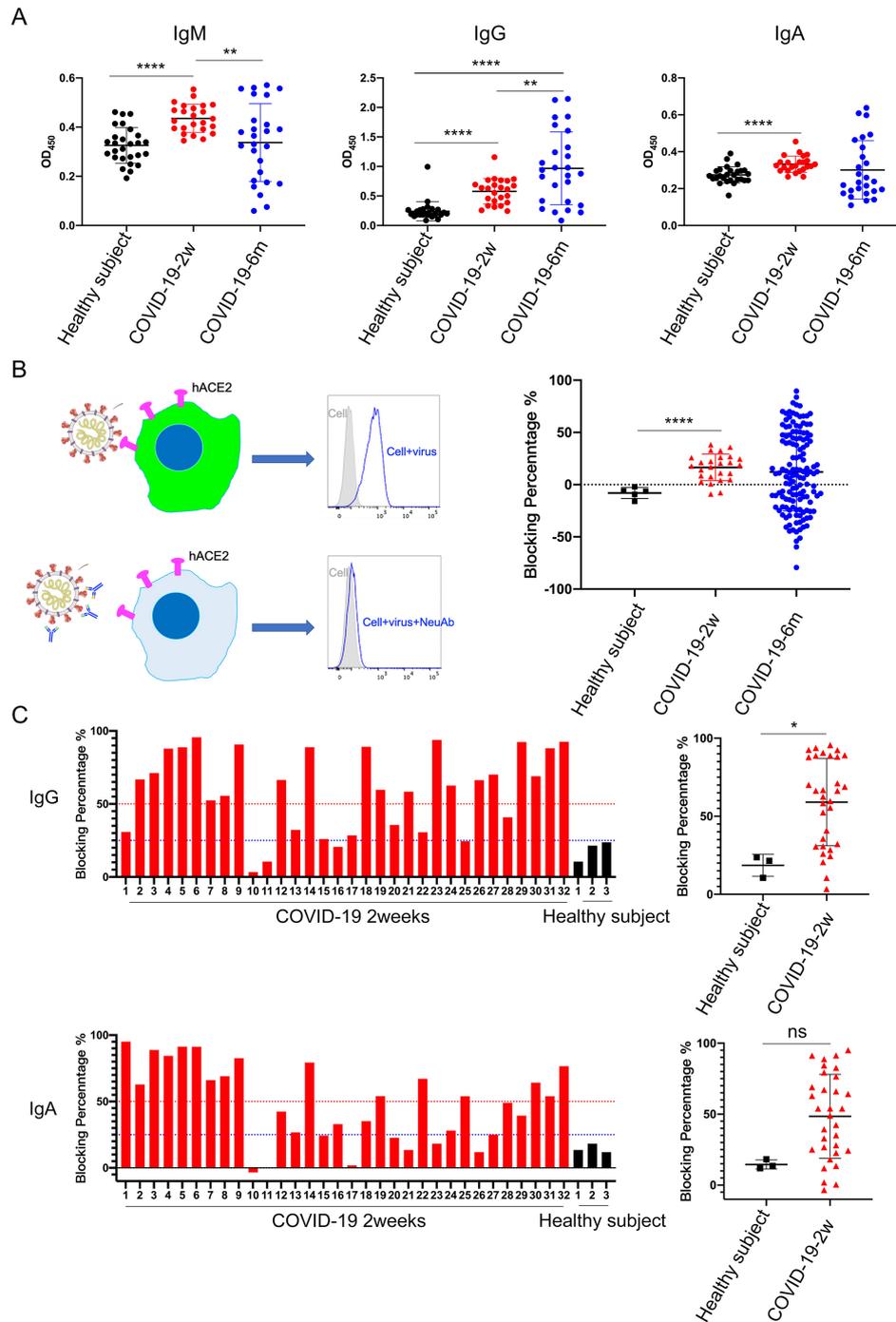
stream primer to target 5' terminal specific nucleotides sequences, and downstream primers to target BCR constant regions in 5 classes of BCR heavy chain, or the 2 types of BCR light chains. In addition, because the length of 5'-RACE products exceeded limits of some next-generation sequencing platforms, we used the PacBio sequencing technology which can sequence the full length of the BCR variable region (Fig. 2A).

We first amplified and obtained the V(D)J transcripts of four classes of IgH (except IgE), and VJ transcripts of two types of IgL (Ig $\kappa$  and Ig $\lambda$ ) (**Supplementary Fig. 1**). Using the PacBio sequencing platform and IMGT/HighV-QUEST (version 1.7.1), a website that analyzes thousands of immunoglobulin gene nucleotide sequences, we obtained a total of 157,245 BCR transcript reads from the 35 patients we studied. As expected, sufficient diversity was observed in the BCR transcripts that were obtained from each individual's B-cells, and detailed data are shown in **Supplementary Table 2**.

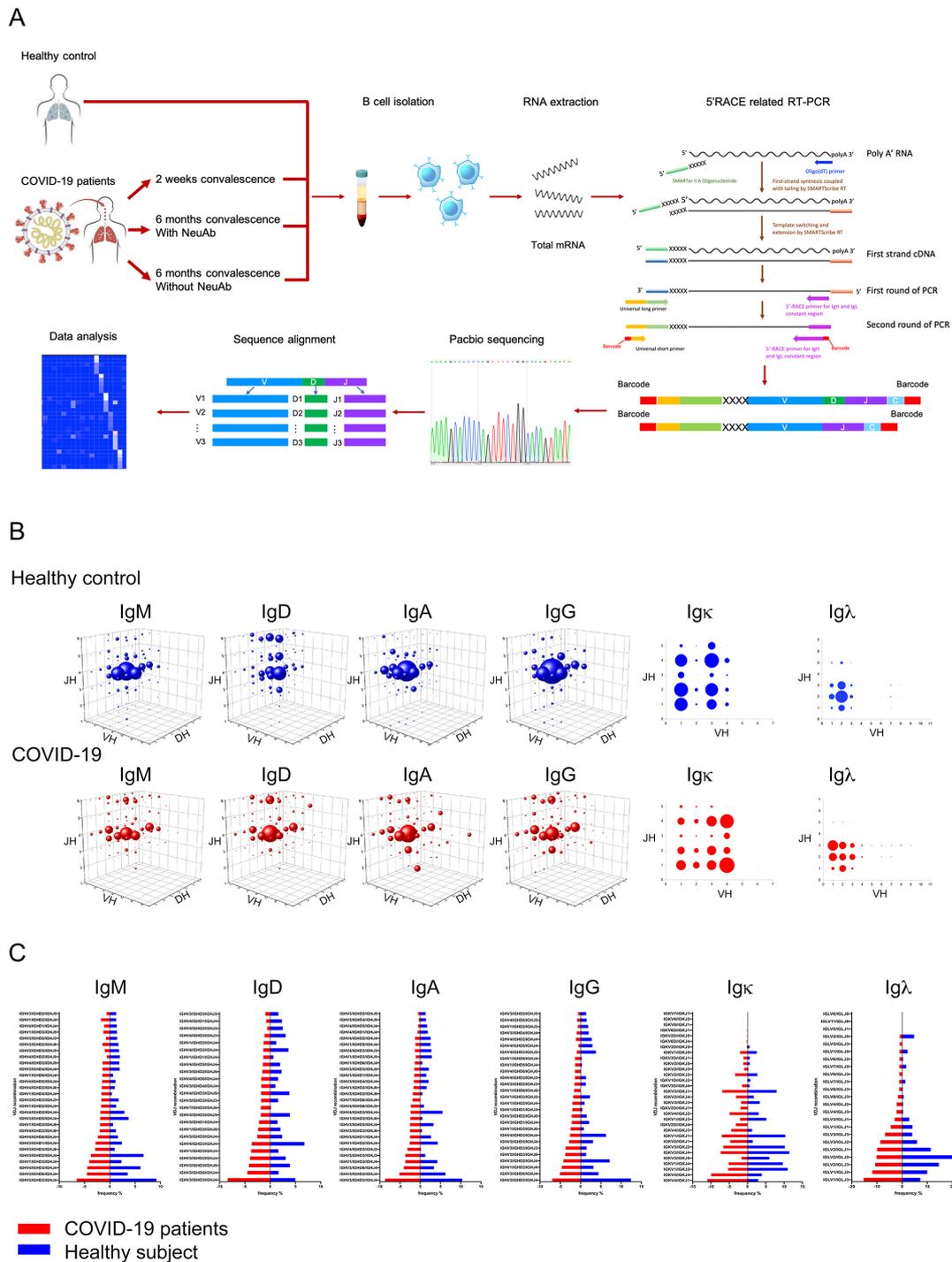
Next, we analyzed whether the relative abundance and inter-status variability of the V, D, J gene family was different between patients and healthy people (Fig. 2B,C). Our results showed that although there is no significant difference in the overall landscape and inter-status variability of each V(D)J gene family, some dominant recombination of V, D and J families are significantly different between patients and healthy donors. For example, the VH3DH3JH4 recombination pattern was dominant in both COVID-19 patients and healthy subjects. In contrast, some increased VDJ recombination patterns were more frequently observed in COVID-19 patients than in healthy controls, such as the increased VH4DH6JH6 and VH1DH6JH6 patterns usually found in IgG from COVID-19 patients, but not in healthy controls. Increased VH7DH6JH4 and VH4DH3JH2 patterns were usually observed in IgA from COVID-19 patients, and increased VH7DH3JH6 was observed in IgD from COVID-19 patients.

### 3.3 Clonal Expansion and Common CDR3 Shared among the COVID-19 Patients

Antibodies commonly display large levels of diversity within the same individual or among different individuals. Theoretically, there are few convergent clonotype antibodies in an individual or between different individuals under physiological conditions. However, after an antigenic challenge, the convergent clonotypes antibodies or memory B-cells can be increased. In this study, we first analyzed the clonal proliferation of B-cells from each individual, as a whole, by analysis of the percentage of Top 10 CDR3 sequences (Fig. 3A). We found that B-cells expressing different types of Ig showed a certain degree of clonal bias within each individual patient, an even in healthy individuals, because they were previously infected by pathogens other than the SARS-CoV-19 virus. As expected, higher clonal bias of IgH chain and Ig $\lambda$  were found in COVID-19



**Fig. 1. S-specific antibodies in the serum of COVID-19 patients 2-weeks and 6-months convalescence and the neutralizing activity.** (A) The concentration of S-specific IgM, IgG and IgA in serum of COVID-19 2-weeks and 6-months convalescence by ELISA assay. Unpaired *t*-test, data are presented as mean  $\pm$  SEM, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001. (B) The neutralizing activity of serum of COVID-19 patients 2-weeks and 6-months convalescence. Serum of healthy controls and COVID-19 patients 2-weeks and 6-months convalescence were pre-incubated with SARS-CoV-2 pseudovirus labeled with GFP and then incubated with HEK293T overexpressed human ACE2. FACS was used to detect the fluorescence signal of GFP. Unpaired *t*-test, data are presented as mean  $\pm$  SEM, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001. (C) The neutralizing activity of IgG and IgA purified from serum of COVID-19 patients 2-weeks and 6-months convalescence. Unpaired *t*-test, data are presented as mean  $\pm$  SEM, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001.



**Fig. 2. Study design and the schematic diagram of 5'RACE.** (A) Schematics of the experimental design for Ig V(D)J sequencing. PBMC were obtained from healthy people and COVID-19 patients and CD19<sup>+</sup> B cells were sorted by immunomagnetic positive selection and total RNA was extracted for Ig variable region amplification. SMARTer IIA Oligonucleotide was random add to the 5' end of RNA. Then universal upstream primer to target 5' end specific nucleotides sequences and the downstream primers to target constant region of Ig. After two round of PCR, complete variable region was amplified and sequenced by pacbio sequencing. (B) VDJ family gene used in healthy control and COVID-19 patients by bubble chart. X, Y, Z axis represented for VH, JH and DH family gene separately, and the size of bubble represented for the proportion of the VHDHJH gene. (C) Frequency of the VHDHJH gene in healthy people (blue) and COVID-19 patients (red).

convalescent patients, especially in patients 2 weeks after recovery, while Ig $\kappa$  showed no difference between healthy subjects and COVID-19 patients. We noted that IgG displayed an even higher clonal bias in COVID-19 patients at 6 months, rather than 2 weeks, after recovery, and that this finding is quite different from other IgH chains.

In addition, many convergent clonotypes were frequently observed among the COVID-19 patients. Interestingly, we found a time-shift appearance of convergent clonotype pattern between different Ig classes in COVID-19 patients after 2 weeks versus 6 months of recovery (Fig. 3B,C). Shown is the shared convergent clonotypes that were significantly increased in IgD, IgA, IgG and Ig $\kappa$  in COVID-19 patients 6 months rather than 2 weeks after recovery. Of note, IgM showed an increased proportion of shared convergent clonotypes in COVID-19 patients 2 weeks after recovery when compared to the 6 month convalescence group. These results suggest that at least 2 weeks after COVID-19 convalescence, expansion of B-cell clones is dominated by the expression of IgM, which is converted to IgG and IgA at the 6 month time point. In contrast with COVID-19 patients, healthy subjects do not share CDR3 sequence among individuals.

#### 3.4 Different Types of BCR Repertoire in an Individual Show Their Unique Specificity; However, BCR in Patients with COVID-19 Still Share V(D)J Usage Characteristics

In this study we analyzed the ratio of top 5 V(D)J recombination patterns in COVID-19 patients and healthy subjects (Fig. 4A), and found that the most frequently used V(D)J recombination pattern in *IGH* genes were different among individuals, while Ig $\lambda$  showed conserved recombination pattern in both healthy subjects and COVID-19 patients. The percentage of the IgG-biased V(D)J recombination pattern displayed increases in COVID-19 patients 6 months after recovery when compared to healthy subjects, while Ig $\lambda$  showed a decreased ratio of biased VJ recombination patterns in COVID-19 patients versus healthy subjects.

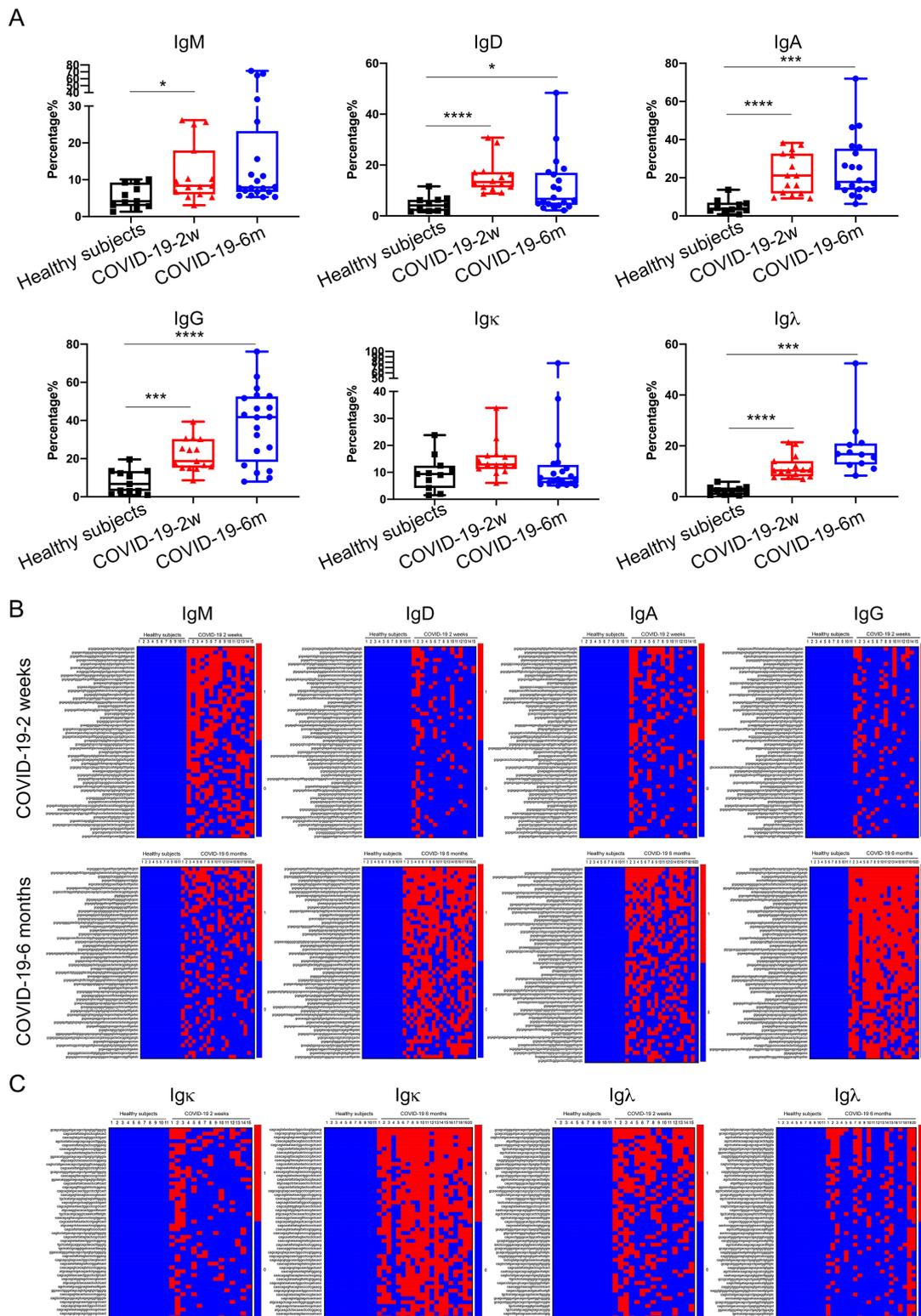
We found that, among the individuals, significant diversity in recombination patterns within IgM, IgA, IgG, IgD as well as Ig $\kappa$ , but not in Ig $\lambda$  (Fig. 5A). This finding indicates that, under physiological conditions, the variable region of Ig $\lambda$  is conserved in different individuals. In contrast, convergent recombination patterns were frequently observed among the COVID-19 patients. Interestingly, we found a time shift change of appearing convergent recombination patterns between different Ig classes in COVID-19 patients after recovery of 2 weeks to 6 months. These shared convergent recombination pattern were significantly increased in IgM and Ig $\kappa$ , and moderately increased in IgA in COVID-19 patients after a 2-week recovery. However, when 6-month recovery COVID-19 patients were analyzed, increased proportions of shared convergent recombination pattern were converted into IgD, IgG and IgA class im-

munoglobulins. These results suggest that at least 2 weeks into COVID-19 convalescence, the expansion of the B-cell clone is dominated by the expression of IgM and IgA, but not IgG.

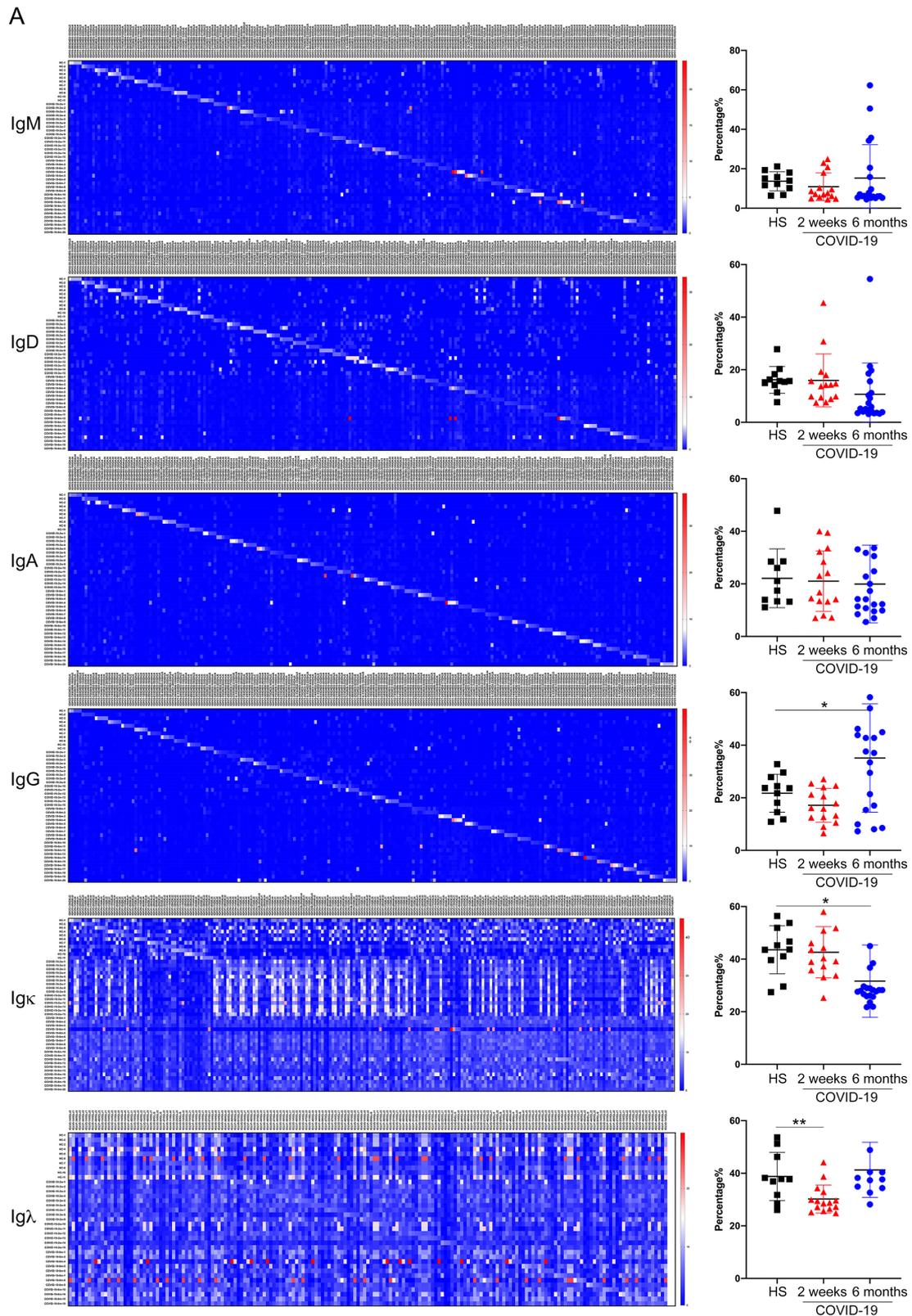
We next analyzed and compared the shared V(D)J patterns between the different classes of Ig, which exist only among COVID-19 patients after 2 weeks of recovery (Fig. 5B,C). We found that the IgD class frequently utilizes IGHV3-74\IGHD6-13\IGHJ4 (5/15), whereas IgA tend to use IGHV4-39\IGHD3-22\IGHJ2 (9/15). IgG were likely to use IGHV1-8\IGHD3-22\IGHJ4 of (6/15), and Ig $\lambda$  usually use IGLV3-21\IGLJ6 (8/15). Note that the IGHV1-18\IGHD3-10\IGHJ6 pattern was shared by IgM class immunoglobulins in all COVID-19 patients (15/15), and the IGKV2D-29\IGKJ1 of Ig $\kappa$  was shared by 14 of 15 COVID-19 patients. In addition, we also analyzed for differences in the convergent clonotypes between moderate and severe individuals, and found no significant difference between the recovered moderate and severe individuals (data not shown).

#### 3.5 Different Classes of Ig Heavy Chain Showed Quite Different CDR3 Sequences within the Same Individual in COVID-19 Patients and Healthy Subjects

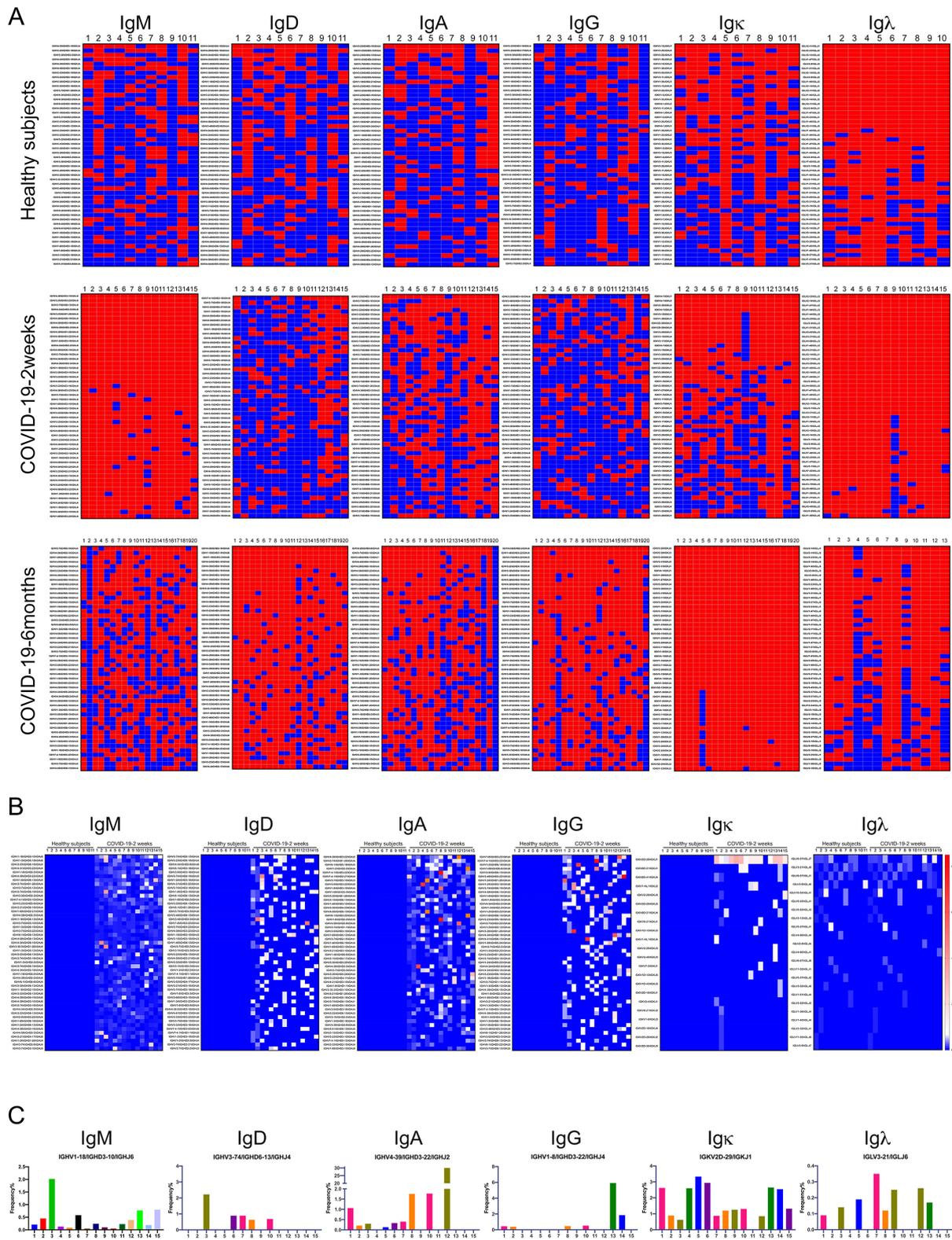
According to the classical clonal selection theory, different types of Ig produced by one B-cell all present the same V(D)J combination and sequence. In other words, different types of Ig detected from the same B-cell cluster should present the same V(D)J combination pattern. However, our previous finding, revealed by single cell sequencing, was that one B-cell can express more than one classes of Ig and, moreover, different classes of Ig express their own unique variable region sequence, which is somewhat at odds with the classical concept of Ig class switching [19]. In this study, we further addressed if the same population of B-cell-derived Ig classes display different V(D)J recombination patterns (Fig. 6A). Results showed that approximately 80% of V(D)J recombination patterns were uniquely expressed in one Ig heavy chain within the same healthy individual (IgM:  $77.39 \pm 11.79$ ; IgD:  $78.25 \pm 11.21$ ; IgA:  $78.16 \pm 11.87$ ; IgG:  $78.37 \pm 11.45$ ). In contrast, the percentage of uniquely expressed V(D)J recombination patterns were decreased in COVID-19 patients, especially in COVID-19 patients 2 weeks after recovery (COVID-19 2 weeks: IgD:  $34.54 \pm 10.56$ ; IgA:  $43.57 \pm 7.521$ ; IgG:  $35.57 \pm 10.84$ ; COVID-19 6 months: IgD:  $60.55 \pm 8.2$ ; IgA:  $54.01 \pm 8.991$ ; IgG:  $51.66 \pm 8.883$ ). It should be noted that the percentage of unique V(D)J recombination patterns in IgM were similar in COVID-19 patients 2 weeks after recovery and healthy subjects (2weeks:  $74.14 \pm 9.214$ ) (Fig. 6A). The percentage of uniquely expressed CDR3 of each Ig in the same individual was significantly decreased in COVID-19 patients especially in the 2 week convalescent patient (2 weeks: IgM:  $92.93 \pm 3.037$ ; IgD:  $72.93 \pm 13.64$ ; IgA:  $82.75 \pm 8.926$ ; IgG:  $74.4 \pm 8.53$ ; 6 months: IgM:  $98.25$



**Fig. 3. Clonal expansion and common CDR3 shared among the COVID-19 patients.** (A) The top 10 CDR3 ratio in COVID-19 patients compared with healthy people. COVID-19 6m-A represents for patients with neutralizing antibodies and COVID-19 6m-B represents for patients without neutralizing antibodies. Unpaired *t*-test, data are presented as mean  $\pm$  SEM, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001. (B) COVID-19 patients specific CDR3 sequences of Ig heavy chain that are shared between every individual in COVID-19 patients 2 weeks and 6 months after recovery. Red cell represents for the CDR3 exists in this individual while blue cells represent for the CDR3 does not exists in this individual. (C) COVID-19 patients specific CDR3 sequences of Ig light chain that are shared between every individual in COVID-19 patients 2 weeks and 6 months after recovery.



**Fig. 4. V(D)J recombination pattern in healthy subjects and COVID-19 patients.** (A) The top 5 V(D)J rearrangement patterns that are used by B cells of healthy people and COVID-19 patients (left) and the percentage of top 5 V(D)J rearrangement in healthy people and COVID-19 patients. Unpaired *t*-test, data are presented as mean  $\pm$  SEM, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001.



**Fig. 5. Shared V(D)J recombination pattern in healthy subjects and COVID-19 patients.** (A) Common V(D)J recombination pattern between individuals of healthy subjects and COVID-19 patients 2 weeks and 6 months after recovery. (B) COVID-19 patients specific V(D)J recombination pattern that are shared between every individual in COVID-19 patients 2 weeks after recovery. Heats maps show the expression frequency of Ig heavy chain and light chain. (C) The percentage of the most shared V(D)J recombination pattern in every patient by column charts.

$\pm 0.9742$ ; IgD:  $97.98 \pm 0.7597$ ; IgA:  $96.49 \pm 1.441$ ; IgG:  $96.3 \pm 1.48$ ; healthy controls: IgM: 1.0; IgD: 1.0; IgA: 1.0; IgG: 1.0) (Fig. 6B).

We further analyzed if some convergent clonotypes were shared by different classes of Ig in the same individual (Fig. 6C). As expected, unlike the healthy donors that do not share convergent clonotypes among the different classes of Ig in an individual, we found that the same clonotype patterns frequently shared different classes of Ig in COVID-19 patients. For example, IgM showed increased frequency of the same CDR3 sequence with IgD, IgG and IgA in 2 weeks of convalescence COVID-19 patients, while IgD showed increased frequency of the same CDR3 sequence with IgG and IgA in COVID-19 patients with 6 months convalescence.

### 3.6 Shared V(D)J Recombination Patterns and CDR3 Sequences between COVID-19 Patients at 2 Weeks and 6 Months of Convalescence

In theory, memory B-cells at 2 weeks and 6 months of convalescence may exhibit the same or similar V(D)J or VJ rearrangement patterns. We first excluded the V(D)J and VJ rearrangements shared with healthy individuals, then compared the VDJ and VJ rearrangements of different IgH and Ig $\lambda$  between COVID-19 patients with 2 weeks versus 6 months of convalescence (Fig. 7A,B). We found that identical V(D)J patterns shared by different COVID-19 patients at 2 weeks and 6 months of convalescence were easy to identify in all types of Ig heavy and light chains, while shared CDR3 exist only in IgG and Ig light chains. In detail, IGHV1-18\IGHD3-10\IGHJ6 of IgM (26/35), IGHV3-23\IGHD6-19\IGHJ4 of IgD (20/35), IGHV3-74\IGHD1-26\IGHJ4 of IgA (19/35), IGHV4-39\IGHD2-2\IGHJ6 of IgG (21/35) was frequently shared by COVID-19 patients at 2 weeks and 6 months of convalescence. Of note, IGKV2D-29\IGKJ1 of Ig $\kappa$  was shared by 28 of 35 COVID-19 patients.

This work also identified numerous V(D)J and CDR3 sequences that are shared by COVID-19 patients convalescent for 2 weeks and 6 months who demonstrate neutralizing antibodies but not in COVID-19 patients 6 months convalescent without neutralizing antibodies nor in healthy subjects (Fig. 7C,D). These sequences may contribute to screening neutralizing antibodies.

### 3.7 Some V(D)J Recombination Sequence of Recovered COVID-19 Patients can Map to the Sequence of Spike Protein-Specific Neutralizing Antibodies Analyzed by Protein Mass Spectrometry

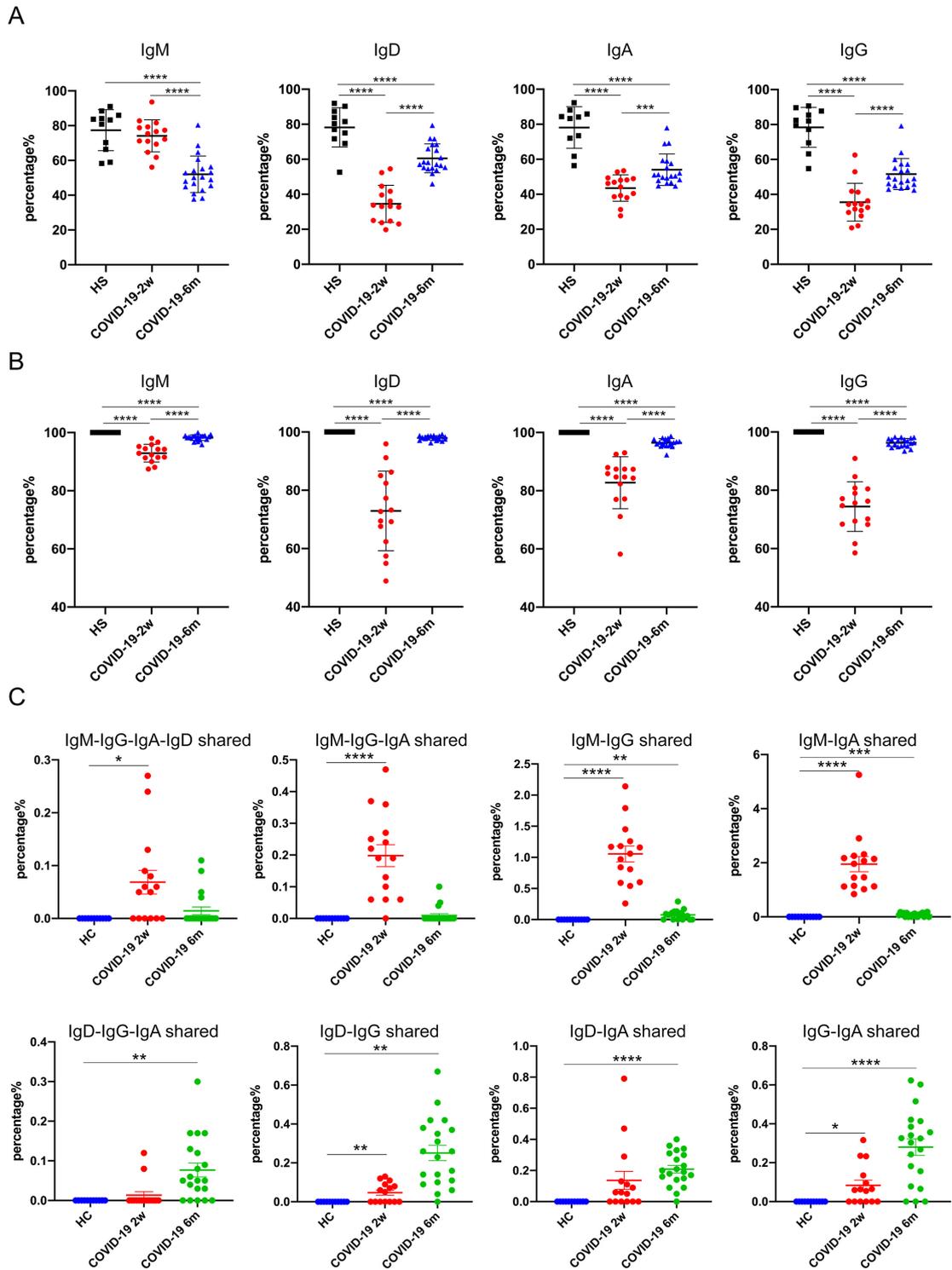
We first prepared an affinity chromatography column using the spike (S) protein of SARS-CoV-2. IgG was purified by protein G chromatography from serum taken from 15 2-weeks convalescent COVID-19 patients. Following this, the IgG from each individual was purified by S protein affinity chromatography. Next, we identified IgG which can block the binding between pseudovirus and ACE2 overexpressing HEK293 cells, and the peptide pro-

file of IgG with viral blocking activity was analyzed by protein mass spectrometry (Fig. 8). Subsequently, peptide sequence of IgG was mapped to the V(D)J sequence of recovered COVID-19 patients obtained in this study. Significantly, we found that some IgG variable domain peptides could map with the V(D)J sequence of recovered COVID-19 patients (Fig. 8), which suggested that the BCR repertoire library obtained in this study can be used to screen for effective neutralizing antibodies against SAR-CoV-2 or to prepare a phage antibody library following SARS-CoV-2 infection.

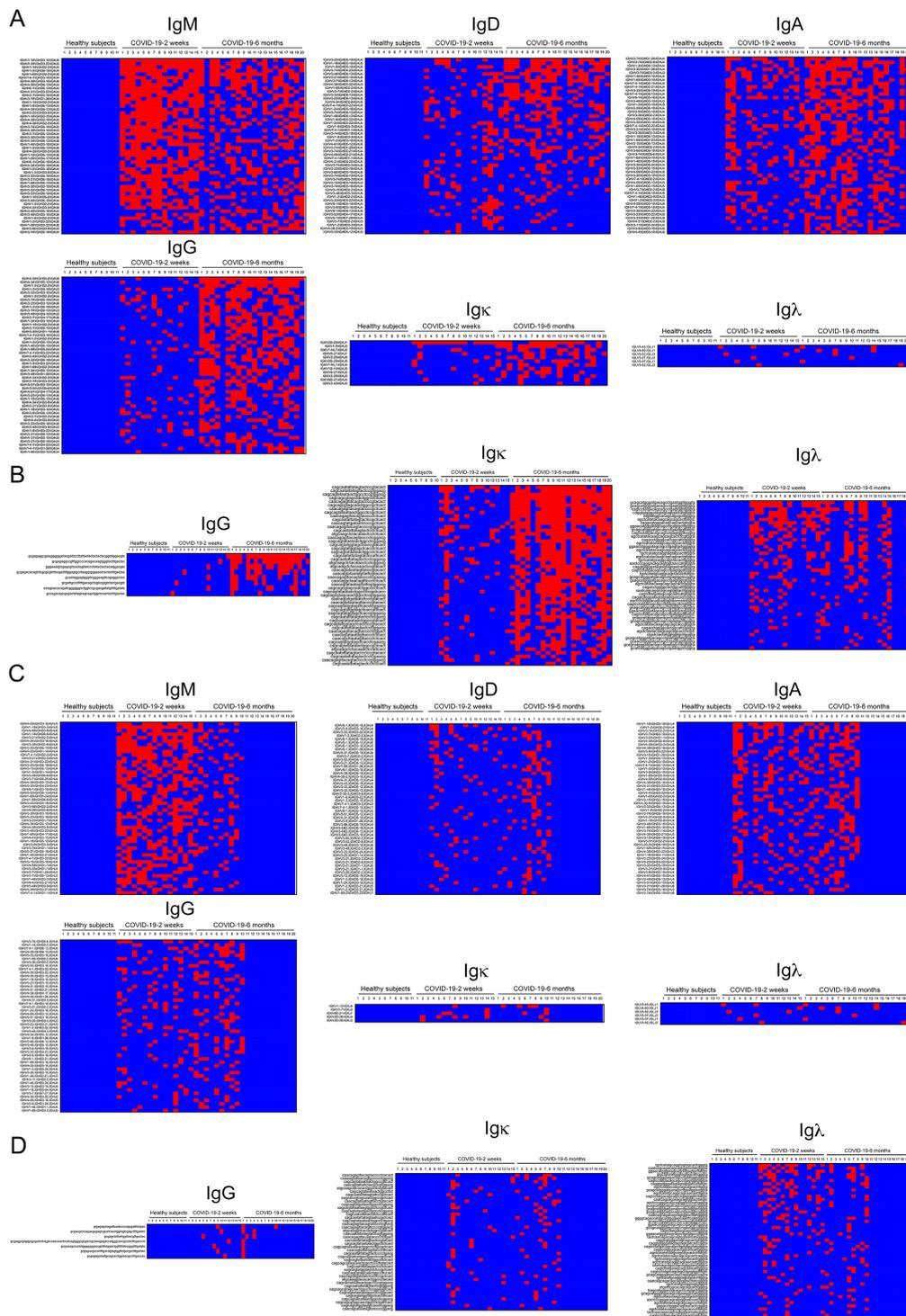
## 4. Discussion

Using the novel strategy of combining 5'-RACE and Pacbio sequencing, we obtained the entire BCR repertoire of 4 heavy chain genes and 2 light chain genes from convalescent COVID-19 patients 2 weeks (15 cases) and 6 months (20 cases) after infection. We demonstrated that clonal expansion and shared common CDR3 sequences among different individuals existed in COVID-19 patients but not healthy subjects. Furthermore, CDR3 sequences were also shared among different classes of Ig in the same individual, which was only observed in COVID-19 patients but not healthy subjects. We also identified numerous specific BCR sequences that displayed neutralization activity towards the S protein of SARS-CoV-2, especially those that exist only in the 2 weeks convalescent and 6 months convalescent samples with neutralizing antibody activity. This BCR repertoire database can also be used to identify antibodies that mediate pathological injury in COVID-19 patients.

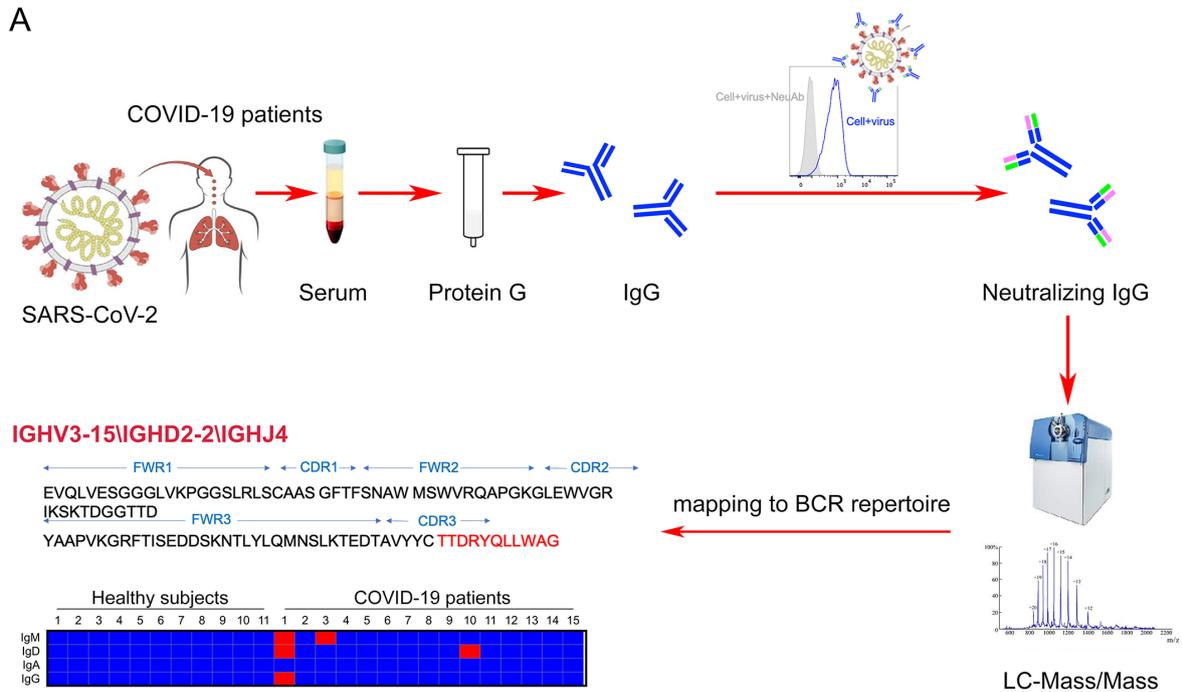
After pathogen infection, B-cells will produce high affinity neutralizing antibodies to prevent pathogen invasion as well as produce specific memory B-cells to limit subsequent infections with the same pathogen. With the SARS-CoV-2 outbreak, scientists seek to understand the genetic information that gives rise to neutralizing antibodies with the aim of improving current therapeutic regimens using genetic engineering. Although the repertoire of BCR presents a unique snapshot into the history of immune response within an infected individual [22–25], identifying signatures of a functional antibody to a given pathogen from the population of BCRs in circulation is a challenging task as several obstacles are encountered. First, Ig expressed by B-cells in the body has infinite diversity, and it is not straightforward to identify neutralizing antibodies to certain antigen epitopes from peripheral blood. Second, single cell sequencing and multiplex PCR are commonly used to analyze BCR and these two approaches have limitations making the acquisition of complete BCR variable region sequence challenging [26]. Third, current immunological theory holds that one or more Ig subtypes from the same group of B-cells have identical variable region sequence. At present, most studies only analyze IgG sequences which limits the amount of BCR information gathered.



**Fig. 6. VDJ recombination pattern and CDR3 sequences of different Ig classes in the same individual of healthy subjects and COVID-19 patients.** (A) The percentage of VDJ recombination pattern that exist only in one type of Ig in the same individual. Unpaired *t*-test, data are presented as mean  $\pm$  SEM, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001. (B) The percentage of CDR3 sequence that exist only in one type of Ig in the same individual. Unpaired *t*-test, data are presented as mean  $\pm$  SEM, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001. (C) Common CDR3 sequences shared between Ig heavy chains in COVID-19 patients compared with healthy controls. Unpaired *t*-test, data are presented as mean  $\pm$  SEM, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001.



**Fig. 7. Common V(D)J recombination pattern and CDR3 sequences that are shared by COVID-19 2 weeks and 6 months after recovery.** (A) Common V(D)J recombination pattern exist only in COVID-19 patients and shared by COVID-19 2 weeks and 6 months after recovery. (B) Common CDR3 sequences exist only in COVID-19 patients and shared by COVID-19 2 weeks and 6 months after recovery. (C) Common V(D)J recombination pattern shared by COVID-19 2 weeks and 6 months after recovery with neutralizing antibodies and do not exist in healthy subjects and COVID-19 6 months after recovery without neutralizing antibodies. (D) Common CDR3 sequences shared by COVID-19 2 weeks and 6 months after recovery with neutralizing antibodies and do not exist in healthy subjects and COVID-19 6 months after recovery without neutralizing antibodies.



**Fig. 8. Scheme to verify the presence of neutralizing antibody sequences in the BCR repertoire obtain in this study.** Purified IgG from the serum of COVID-19 patients by protein G and identified the neutralizing IgG by pseudovirus blocking experiment and then the peptide information was obtained by LC-Mass/Mass. After mapping to BCR repertoire, we obtained the amino acid sequence of neutralizing antibodies.

In this study, we first made methodological innovations to maximize the specific antibody information obtained from peripheral blood B-cells harvested from COVID-19 patients. We obtained these B-cells by magnetic cell sorting which will isolate almost all the B-cell clusters in the peripheral blood. This step overcomes the limitation of the number of cells to be analyzed using single cell sequencing technology. 5'-RACE was used to overcome amplification bias introduced by multiple rounds of thermocycling, and Pacbio sequencing provided sufficient sequencing length to obtain full-length sequencing of the BCR variable region not achievable with other next generation sequencing platforms. In this study, we found that unlike BCR variable regions of healthy individuals that showed infinite BCR diversity, aggregated clonal types were frequently observed in COVID-19 patients. This finding suggests that 5'-RACE can indeed ensure the unbiased amplification of BCR, and that the approach we have adopted in this study can be used to analyze the BCR repertoire in response to other pathogens.

Based on our recent discovery that individual B-cells can produce more than one antibody, analyzing B-cells for only one type of Ig has limitations. Our results showed that in addition to IgG, other Ig subtypes also exhibit clonal expansion and dominant rearrangements, which suggests that finding neutralizing antibodies should not be limited to IgG, but also other classes of Ig as well. In addition, we found

that the proportion of IgM and IgA immunoglobulins sharing the same V(D)J rearrangement pattern within different individuals was significantly higher in COVID-19 patients than healthy individuals, while in these patients IgG showed no difference from that of healthy controls. The proportion of IgM, IgD, IgG and IgA sharing the same V(D)J rearrangement in COVID-19 patients 6 months after recovery was significantly higher than that observed in healthy individuals. Data gathered suggests that SARS-CoV-2-specific IgM antibodies appears earlier, but persist for a relatively short time, whereas SARS-CoV-2-specific IgG and IgA appears later and persists longer which is consistent with previous findings [27]. We also found that the proportion of common CDR3 that were shared among different patients were significantly increased in IgG of COVID-19 patients 6 months after recovery which indicated that the IgG memory B cells are predominant post 6 months of infection which is consistent with previous findings [28]. These results indicate that searching for specific antibodies should focus on IgM and IgA within 2 weeks after infection, and IgG at later time points. Our study also found that IgD also showed a high proportion of shared V(D)J sequences and consistent CDR3 sequences in 6-month convalescent COVID-19 patients. Interestingly, IgD and IgG/IgA shared the same CDR3 found in 6-month convalescent COVID-19 patients but not in 2-week convalescent COVID-19 patients 2-weeks and healthy individuals. This finding suggests that

IgD may also play an important role in viral infection and long term immunologic memory and needs further investigation.

We purified IgG possessing neutralizing activity from the peripheral blood of COVID-19 patients and deduced the primary amino acid sequence of the variable region using LC MS/MS. After comparing the BCR repertoires obtained in this study, we determined the CDR3 sequences within the variable region and synthesized a peptide to verify its neutralizing activity using pseudovirus blocking. The results showed that the synthetic CDR3 polypeptide had neutralizing activity supporting our view that our sequencing approach can serve to identify neutralizing antibodies with high affinity. More importantly, the B cell epitopes that were used to identify neutralizing antibodies in this study were highly conserved and were not mutated in Omicron variant, which indicated that our BCR repertoire can still be useful for finding neutralizing antibody sequences against Omicron and other variants. However, there are also some limitations as the combination of heavy and light chains synthesized in each cell cannot be obtained. Moreover, it is necessary to continuously combine different heavy and light chains to obtain better blocking effects and prepare complete antibodies.

In general, the BCR repertoire analysis outlined in this report provides information on the specific nature of the B-cell response to SARS-CoV-2 infection. The information generated has the potential to aid in the treatment of COVID-19 by supporting diagnostic approaches to predict the progression of disease, informing vaccine development, and enabling the development of therapeutic antibody treatments and prophylactics.

## 5. Conclusions

In summary, we used 5'-RACE combined with Pacbio sequencing to interrogate the BCR repertoire of peripheral blood CD19<sup>+</sup> B-cells from convalescent patients after SARS-CoV-2 infection. Data gathered indicates numerous clonotypes were frequently shared between different patients and different classes of immunoglobulins. These convergent clonotypes provide a resource to identify potential therapeutic/prophylactic antibodies, or pathological antibodies, in patients infected by SARS-CoV-2.

## Author Contributions

ZZ, HD, YL, SZ and XQ conceived the study and designed the workplan. Experimental work—ZZ, MY, HY, HD. Providing human blood sample—XJ, LL, WL, YZ, GK, YL, SZ and NF. Collection of data—WX. Providing laboratory—FL, JW and NZ. Statistical analysis—ZZ, PW, FD and JH. Manuscript writing—ZZ, PW and XQ.

## Ethics Approval and Consent to Participate

The study protocol 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).

## Acknowledgment

We would like to thank the help of You Wan, Ruqun Shen, and Hongwei Cui on this study.

## Funding

This work was supported by the China Ministry of Science and Technology (2020YFA0707801), National Natural Science Foundation of China (82030044), Discipline Construction Project of Peking University Health Science Center (BMU2020HKYZX004).

## Conflict of Interest

The authors declare no conflict of interest.

## Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.31083/j.fbl2802040>.

## References

- [1] Earle KA, Ambrosino DM, Fiore-Gartland A, Goldblatt D, Gilbert PB, Siber GR, *et al.* Evidence for antibody as a protective correlate for COVID-19 vaccines. *Vaccine*. 2021; 39: 4423–4428.
- [2] Awadasseid A, Wu Y, Tanaka Y, Zhang W. Current advances in the development of SARS-CoV-2 vaccines. *International Journal of Biological Sciences*. 2021; 17: 8–19.
- [3] Shields AM, Faustini SE, Kristunas CA, Cook AM, Backhouse C, Dunbar L, *et al.* COVID-19: Seroprevalence and Vaccine Responses in UK Dental Care Professionals. *Journal of Dental Research*. 2021; 100: 1220–1227.
- [4] Eckl-Dorna J, Villazala-Merino S, Linhart B, Karaulov AV, Zhernov Y, Khaitov M, *et al.* Allergen-Specific Antibodies Regulate Secondary Allergen-Specific Immune Responses. *Frontiers in Immunology*. 2019; 9: 3131.
- [5] Janeway C. *Immunobiology. the immune system in health and disease*. 6th edn. Garland Science: New York. 2005.
- [6] Shlomchik MJ, Luo W, Weisel F. Linking signaling and selection in the germinal center. *Immunological Reviews*. 2019; 288: 49–63.
- [7] Frölich D, Giesecke C, Mei HE, Reiter K, Daridon C, Lipsky PE, *et al.* Secondary immunization generates clonally related antigen-specific plasma cells and memory B cells. *The Journal of Immunology*. 2010; 185: 3103–3110.
- [8] Burnet FM. *The clonal selection theory of acquired immunity*. Vanderbilt University Press: Nashville. 1959.
- [9] Cyster JG, Allen CDC. B Cell Responses: Cell Interaction Dynamics and Decisions. *Cell*. 2019; 177: 524–540.
- [10] Chen YW, Luo WX, Wang MQ, Wang J, Li LF, Yuan Q, *et al.* Isolation of human antibodies against hepatitis E virus from phage display library by immobilized metal affinity chromatog-

- raphy. *Biomedical and Environmental Sciences*. 2007; 20: 488–494.
- [11] Cui C, Zhu Y, Jia Z, Mao Q, Lan L. Identification of two novel anti-HCV E2 412–423 epitope antibodies by screening a Chinese-specific phage library. *Acta Virologica*. 2019; 63: 149–154.
- [12] Zhang JL, Gou JJ, Zhang ZY, Jing YX, Zhang L, Guo R, *et al*. Screening and evaluation of human single-chain fragment variable antibody against hepatitis B virus surface antigen. *Hepato-biliary & Pancreatic Diseases International*. 2006; 5: 237–241.
- [13] Jung D, Giallourakis C, Mostoslavsky R, Alt FW. Mechanism and Control of V(D)J Recombination at the Immunoglobulin Heavy Chain Locus. *Annual Review of Immunology*. 2006; 24: 541–570.
- [14] Setliff I, Shiakolas AR, Pilewski KA, Murji AA, Mapengo RE, Janowska K, *et al*. High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. *Cell*. 2019; 179: 1636–1646.e15.
- [15] Montague Z, Lv H, Otwinowski J, DeWitt WS, Isacchini G, Yip GK, *et al*. Dynamics of B-cell repertoires and emergence of cross-reactive responses in COVID-19 patients with different disease severity. *medRxiv*. (in press)
- [16] Niu X, Li S, Li P, Pan W, Wang Q, Feng Y, *et al*. Longitudinal Analysis of T and B Cell Receptor Repertoire Transcripts Reveal Dynamic Immune Response in COVID-19 Patients. *Frontiers in Immunology*. 2020; 11: 582010.
- [17] Schultheiß C, Paschold L, Simnica D, Mohme M, Willscher E, von Wenserski L, *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.
- [18] Trück J, Ramasamy MN, Galson JD, Rance R, Parkhill J, Lunter G, *et al*. Identification of Antigen-Specific B Cell Receptor Sequences Using Public Repertoire Analysis. *The Journal of Immunology*. 2015; 194: 252–261.
- [19] Shi Z, Zhang Q, Yan H, Yang Y, Wang P, Zhang Y, *et al*. More than one antibody of individual B cells revealed by single-cell immune profiling. *Cell Discovery*. 2019; 5: 64.
- [20] van den Heuvel D, Jansen MA, Bell AI, Rickinson AB, Jaddoe VW, van Dongen JJ, *et al*. Transient reduction in IgA+ and IgG+ memory B cell numbers in young EBV-seropositive children: the Generation R Study. *Journal of Leukocyte Biology*. 2017; 101: 949–956.
- [21] Seifert M, Küppers R. Human memory B cells. *Leukemia*. 2016; 30: 2283–2292.
- [22] Robins H. Immunosequencing: applications of immune repertoire deep sequencing. *Current Opinion in Immunology*. 2013; 25: 646–652.
- [23] Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H, Quake SR. The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nature Biotechnology*. 2014; 32: 158–168.
- [24] Horns F, Vollmers C, Dekker CL, Quake SR. Signatures of selection in the human antibody repertoire: Selective sweeps, competing subclones, and neutral drift. *Proceedings of the National Academy of Sciences*. 2019; 116: 1261–1266.
- [25] Nourmohammad A, Otwinowski J, Łuksza M, Mora T, Walczak AM. Fierce Selection and Interference in B-Cell Repertoire Response to Chronic HIV-1. *Molecular Biology and Evolution*. 2019; 36: 2184–2194.
- [26] Galson JD, Schaetzle S, Bashford-Rogers RJM, Raybould MIJ, Kovaltsuk A, Kilpatrick GJ, *et al*. Deep Sequencing of B Cell Receptor Repertoires From COVID-19 Patients Reveals Strong Convergent Immune Signatures. *Frontiers in Immunology*. 2020; 11: 605170.
- [27] Nielsen SCA, Yang F, Hoh RA, Jackson KJL, Roeltgen K, Lee JY, *et al*. B cell clonal expansion and convergent antibody responses to SARS-CoV-2. *Research Square*. (in press)
- [28] Gupta SL, Jaiswal RK. Neutralizing antibody: a savior in the Covid-19 disease. *Molecular Biology Reports*. 2022; 49: 2465–2474.