Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Novel skewed usage of B-cell receptors in COVID-19 patients with various clinical presentations

Junpeng Ma\textsuperscript{a,c,1}, Han Bai\textsuperscript{a,b,1}, Tian Gong\textsuperscript{c,1}, Weikang Mao\textsuperscript{d,1}, Yijun Nie\textsuperscript{c}, Xuan Zhang\textsuperscript{c}, Yanyan Da\textsuperscript{c}, Xiaorui Wang\textsuperscript{a,b}, Hongyu Qin\textsuperscript{a,b}, Qiqi Zeng\textsuperscript{a,b}, Fang Hu\textsuperscript{a,b}, Xin Qi\textsuperscript{a,b}, Bingyin Shi\textsuperscript{c}, Chengsheng Zhang\textsuperscript{a,b,c,1,g,*}

\textsuperscript{a} Precision medicine center, The First Affiliated Hospital of Xi’an Jiaotong University, 277 Yanta West Road, Xi’an 710061, China
\textsuperscript{b} Department of Endocrinology, The First Affiliated Hospital of Xi’an Jiaotong University, 277 Yanta West Road, Xi’an 710061, China
\textsuperscript{c} Center for Molecular Diagnosis and Precision Medicine, and The Department of Clinical Laboratory, The First Affiliated Hospital of Nanchang University, 17 Yongwei Zhengjie, Nanchang 330000, China
\textsuperscript{d} The MED-X Institute, The First Affiliated Hospital of Xi’an Jiaotong University, Building 21, Western China Science and Technology Innovation Harbor, Xi’an 710000, China
\textsuperscript{e} Center for Molecular Diagnosis and Precision Medicine, and The Department of Clinical Laboratory, The First Affiliated Hospital of Nanchang University, 17 Yongwei Zhengjie, Nanchang 330006, China
\textsuperscript{f} LC-BIO TECHNOLOGIES (HANGZHOU) CO., LTD., Hangzhou 310000, China
\textsuperscript{g} The Jackson Laboratory for Genomic Medicine, Farmington, CT 06032, USA

A R T I C L E   I N F O

Keywords:
SARS-CoV-2
COVID-19
Single-cell B cell receptor (BCR) sequencing
B cell receptor repertoire
Asymptomatic infection
Antibody

A B S T R A C T

B cell-mediated immune responses play important roles in controlling SARS-CoV infection. Here, we performed the single-cell B cell receptor sequencing (scBCR-seq) of the PBMC samples from eleven healthy controls, five asymptomatic subjects and 33 symptomatic COVID-19 patients with various clinical presentations, and subsequently analyzed the abundance and diversity of the BCR repertoires in different groups, respectively. We revealed the skewed usage of the IGHV, IGLV and IGKV genes and identified a number of heavy or light chain VDJ gene pairs and combinational preference in each group, such as IGHV3-7 and IGKV2-24 enriched in the asymptomatic subjects, whereas IGHV3-13, IGHV3-23-IGHJ4, IGHV1-18-IGLV3-19, IGHV1-18-IGLV3-21, and IGHV1-18-IGLV3-25 enriched in the recovery patients with severe diseases. We also observed the differential expression of IGHV3-23 in various B cell clusters by analysis of the scRNA-seq data. Additional dock analysis indicated that IGHV3-13 could bind to the spike protein of SARS-CoV-2. These findings may advance our understanding of the humoral immune responses in COVID-19 patients and help develop novel vaccine candidates as well as therapeutical antibodies against SARS-CoV-2 infections.

1. Introduction

The pandemic of the coronavirus disease 2019 (COVID-19) has posed unprecedented challenges to the international communities. As of August 5, 2022, the cumulative number of coronavirus disease-19 (COVID-19) has exceeded 579 million confirmed cases, including over 6.4 million deaths worldwide [1]. In addition, the emerging variants of SARS-CoV-2, such as the Alpha, Beta, Gamma, Delta, and Omicron have generated great concerns across the international communities [2–6]. In particular, the existing vaccines and convalescent serums have shown reduced neutralization activities against these SARS-CoV-2 variants [7,8].

SARS-CoV-2 infection may exhibit diverse clinical presentations, ranging from asymptomatic infection (AS) to critical conditions. While many studies have been conducted to understand the pathogenesis of the moderate disease (MD) and the severe disease (SD), the AS remains poorly characterized [9–15]. Further characterization of the immune responses in AS may help us understand the protective immunity against SARS-CoV-2 infection [16]. Previous studies have shown that asymptomatic individuals developed specific IgM, memory B cells and T cells against the S1 and N proteins of SARS-CoV2, whereas the recovered patients developed SARS-CoV-2-specific IgG antibodies as well as

\textsuperscript{1} Corresponding author.
\textsuperscript{2} E-mail address: cszhang99@xjtu.edu.cn (C. Zhang).
\textsuperscript{3} These authors contributed equally to this work and shared the first authorships.
Here, we performed the single-cell BCR sequencing (scBCR-seq) and analyzed the BCR repertoires in the asymptomatic (AS) and symptomatic (SM) COVID-19 patients as well as the healthy controls (HC). We investigated an ongoing disease outbreak.

The definition and classification of all COVID-19 patients in this study follow the Guidelines of the World Health Organization and the “Guidelines on the Diagnosis and Treatment of the Novel Coronavirus Infected Pneumonia” developed by the National Health Commission of People’s Republic of China [25-27]. The study cohort included eleven healthy controls (HC, n = 11) and 38 SARS-CoV-2-infected individuals (COV, n = 38) consisting of five asymptomatic subjects (AS, n = 5) and 33 symptomatic patients (SM, n = 33) (Table 1, Supplementary Table S1). Moreover, severe disease (SD, n = 10) and SD recovery (SDR, n = 10) group could be further classified from SM. The peripheral blood was collected into the standard EDTA-K2 Vacutte Blood Collection Tubes (Jiangsu Yuli Medical Equipment Co., Ltd, China; Cat. Y09012282) and stored at room temperature or 4 °C until processed. The peripheral blood mononuclear cells (PBMCs) were prepared from these study subjects, including two patients (ID: C19 and C26) with blood samples collected at two different time points during this study. PBMCs were stored in the liquid nitrogen storage tank (-196 °C) until being used for the studies. All experimental procedures were completed inside a biosafety level 2 (BSL-2) laboratory at the Department of Clinical Diagnostic Laboratories, Renmin Hospital of Wuhan University.

2. Materials and method

### 2.1. Study cohort and blood sample collection

This study was approved by the Ethics Committee of The First affiliated Hospital of Xi’an Jiaotong University (XJTU1AF20201SK-015) and The Renmin Hospital of Wuhan University (WDRY2020-K130). All participants enrolled in this study were informed consent, or the oral consent from the patients or their family members, whereas written informed consent was waived by the Provincial and National Health Commissions in China under the exceptional circumstances for investigation of an ongoing disease outbreak.

### 2.2. Preparation of single-cell suspensions

Retrieving the PBMC-containing cryo-tubes from the liquid nitrogen storage tank and placing them in a 37 °C water bath for rapid thawing. PBMCs were mixed with 10 mL washing medium (90% DMEM+10% FBS) in a 15-mL polypropylene tube and centrifuged at 500 g for 20 min. The supernatant was then aspirated (repeat twice). The cell pellets were resuspended with 500 μL FBS (0.04% BSA) in the sterile RNase-free vacutainer tubes and added with 5 mL 1× red blood cell lysis buffer (MACS 130–094–183, 10×) and incubated at room temperature for 10 min to lyse remaining red blood cells. After incubation, the suspension...
was centrifuged at 500 g for 20 min at room temperature. The suspension was resuspended in 100 μl Dead Cell Removal MicroBeads (MACS 130–090–101) and remove dead cells using Miltenyi Dead Cell Removal Kit (MACS 130–090–101). Then the suspension was resuspended in 1× PBS (0.04% BSA) and centrifuged at 300 g for 3 min at 4 °C (repeat twice). The cell pellet was resuspended in 50 μl of 1× PBS (0.04% BSA). The overall cell viability was confirmed by trypan blue exclusion, which needed to be above 85%, and the single-cell suspensions were counted using a Countess II Automated Cell Counter and the cell concentration was adjusted to 700–1200 cells/μl.

### 2.3. Chromium 10x genomics library construction and sequencing

Approximately 5000 single cells each sample were captured using the Chromium Single-Cell 5’ kit (V1) according to the manufacturer’s instructions (10x Genomics), followed by cDNA amplification and library construction performed according to the standard protocols. The libraries were sequenced on an Illumina NovaSeq 6000 sequencing platform (paired-end multiplexing run, 150 bp) by LC-Bio Technology Co. Ltd. (Hangzhou, China) at a minimum depth of 20,000 reads per cell. To avoid batch effects, the scRNA-seq data sets were generated by the same operators at the same laboratories using the standard operation protocols (SOPs) for cell dissociation, library preparation and sequencing.

2.4. Single-cell RNA-seq data processing

The sequencing results were demultiplexed and converted to FASTQ format using Illumina bcl2fastq software. Sample demultiplexing, barcode processing and single-cell 5’ gene counting were completed by using the Cell Ranger pipeline (https://support.10xgenomics.com/single-cell-geneexpression/software/pipelines/latest/what-is-cell-ranger, version 3.1.0) and the scRNA-seq data were aligned to Ensembl genome GRCh38 reference genome. A total of 222,457 single cells captured from eleven healthy controls and 42 COVID-19 patient samples were processed using 10x Genomics Chromium Single Cell 5’ Solution.

The Cell Ranger Seurat (version 3.1.1) was used for dimensional reduction, clustering, and analysis of scRNA-seq data. Overall, 207,718 cells passed the quality control threshold: all genes expressed in less than one cell were removed, number of genes expressed per cell >500 as low cut-off, UMI counts less than 500, and the percentage of mitochondrial-DNA derived gene-expression <25%. To visualize the data, we further reduced the dimensionality of all 207,718 cells by Seurat and used t-Distributed Stochastic Neighbor Embedding (t-SNE) to project the cells into 2D space. The steps included: (1) Using the LogNormalize method of the “Normalization” function of the Seurat software to calculate the expression level of genes; (2) The principal component analysis (PCA) was performed using the normalized expression level, within all the PCs, the top 10 PCs were used to do clustering and t-SNE analysis; (3) Using weighted Shared Nearest Neighbor (SNN) graph-based clustering method to find clusters. The marker genes for each cluster were identified with the “bimod” (Likelihood-ratio test) with default parameters via the FindAllMarkers function in Seurat. This selects marker genes that were expressed in more than 10% of the cells in a cluster and the average log (Fold Change) of greater than 0.26. To further avoid interference of putative multiplets (where more than one cell was loaded into a given well on an array), cells in a defined cluster that had high expression of more than one cell type canonical marker gene were filtered to ensure the data quality. As a result, a total of 119,799 cells were used for the final analysis in this study. The nine cell types were integrated for further sub-clustering. After integration, genes were scaled to unit variance. Scaling, principal component analysis and clustering were performed as described above.

### 2.5. BCR sequencing and analysis

The Barcoded, full-length BCR V(D)J segments were enriched from the amplified 5’ cDNA libraries using the Chromium Single-Cell V(D)J Enrichment kit according to the manufacturer’s protocol (10X Genomics). If both T cells were expected to be present in the partitioned cell population, BCR transcripts were enriched in separate reactions from the same amplified cDNA material. Cell Ranger pipeline (https://support.10xgenomics.com/single-cell-vdj) that processes Chromium single cell 5’ RNA-seq output for V(D)J was employed to assemble, quantify, and annotate paired V(D)J transcript sequences. The workflow of Cell Ranger started by demultiplexing the Illumina sequencer’s base call files (BCLs) for each flow cell directory into FASTQ files. Cell calling was performed independently of V(D)J read filtering and assembly. A clonotype was a set of cells that shared the same paired receptor sequences (presumably derived from the same progenitor cell). Cell barcodes were grouped into clonotypes if they shared the same set of productive CDR3 nucleotides sequences by exact match. The repertoire information of all samples was exhibited using a set of basic features: CDR3 abundance, CDR3 bases length, Variable (V) and Joining (J) segment usage (the frequency of associated reads for each V/J gene presented in the samples), the length of V/J gene in CDR3 region, and V-J gene paired frequency in CDR3 junctions. We used the hierarchical cluster (https://stat.ethz.ch/R-manual/Rdevel/library/stats/html/hclust.html) and Kruskal’s Non-metric Multidimensional Scaling (isoMDS, https://cran.r-project.org/web/packages/MASS/MASS.pdf) to estimate Immune Repertoires similarity. To avoid batch effects, as recommended by the standard 10X protocol, the data sets were generated with the same operators and were also produced in the same laboratories using the same cell dissociation protocols, library preparation approaches and/or sequencing platforms.

2.6. Dock analysis

The homology modeling of IGHV3–13 was performed based on the experimental structure (P01766) using SWISS-MODEL (https://www.swissmodel.expasy.org). The S protein of SARS-CoV-2 and IGHV3–13 were docked using HADDOCK server (https://bianca.science.uu.nl/haddock2.4/) with the default parameters (28,29). The detailed statistical data are displayed for each cluster, representing the average values calculated over the top four best-scoring structures within each cluster. The most reliable cluster was selected by the Z-score (the lower the better) according to HADDOCK. The Z-score indicates how many standard deviations from the average this cluster is located.

2.7. Laboratory tests

All laboratory tests were conducted in the Renmin Hospital of Wuhan University, including white blood cells, neutrophils and lymphocytes (XN2800, SYMIX, Japan); CRP, IL-2, IL-4, IL-6, IL-10 (AUS5400, Beckman Coulter, Inc. USA); CD16/56 NK cells, CD19 B cells (CytoFLEX LX, Beckman Coulter, Inc. USA); and the specific IgM antibodies (Cat#20,203,400,769, YHLO Biotech) and IgG antibodies (Cat#20,203,400,770, YHLO Biotech) against the S and/or N proteins of SARS-COV-2.

2.8. Statistical analysis

All statistical analyses were performed by SPSS (Statistical Package for the Social Sciences) version 23.0 software (SPSS Inc) unless otherwise stated. Diagrams plotting were performed using GraphPad Prism 8.0.2 and R [30]. R packages ggplot2, pheatmap, UpSetR and ggrepel were used. Categorical variables were described as frequency rates and percentages, and continuous variables were described using mean with standard deviation. For two groups, using independent group t-tests when the data were normally distributed; otherwise, the Mann-Whitney
test was used. For three groups, One-way Anova test conducted which was normally distributed and homoscedasticity, otherwise, Kruskal-Wallis was used. Besides, for the multiple comparisons, Bonferrion correction was used. Proportions for categorical variables were compared using the $\chi^2$ test, although the Fisher exact test was used when the data were limited. The statistical details of the experiments were provided in the respective table/figure legends.

3. Data availability

The raw sequence data reported in this paper have been deposited in GEO, under accession code GSE165080 and GSE180118 are publicly accessible at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE165080 [31] and https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE180118. Other supporting raw data are
4. Results

4.1. The study cohort and the single-cell BCR sequencing of PBMCs

To identify the immune cell alternations in COVID-19 patients, we first performed deep single-cell RNA sequencing on peripheral blood

available from the corresponding author upon reasonable request. Source data are provided with this paper.

Fig. 2. Detection of BCR clonal expansions in various groups. (A) The t-SNE plot shows the B cell expansions in the HC, AS, SM, SD and SDR, clonal was marked with red color. (B) The BCR clonal expansion status and frequency of the clonotype in each sample. The number of color blocks represents the complexity of the clonal states. Some individuals’ data were not available due to the limitation of the sequencing technology. (C) The number of highly expanded clones (above the dotted line) in HC, AS, and SM, respectively. The highly expanded clone was defined as the clone comprising 10% or more of all BCR sequences. (D) The number of highly expanded clones (above the dotted line) in SD and SDR, respectively. InFig. 2C and 2D, the bars showing mean ± sd, HC (n = 4), AS (n = 3), SM (n = 15), SD (n = 3) and SDR (n = 6). Samples with the percentage of expanded clones <0 were not shown.
Fig. 3. BCR repertoire in asymptomatic and symptomatic COVID-19 patients. (A) The skewed usage of IGKV genes in HC, AS and SM group, respectively. (B) The motif logo plot showing the CDR3 length distribution and amino acids frequency of IGKV1–9 in AS. The relative size of the letters indicated their frequency in the sequences and the total height of the letters depicts the information content of the position in bits. (C) The Heatmaps showing the VJ gene pairs of IGH, IGK and IGL in AS and SM. The color range indicating the usage counts of specific V-J gene pairs. (D) The Upset view showing the common and unique VJ pairs of IGH, IGK and IGL in HC, AS and SM group, respectively. (E). The motif logo plot showing the CDR3 length distribution and amino acids frequency of IGHV3–23.IGHJ4 in AS and SM. (F) The Heatmaps showing the V gene heavy and light chain combinations in AS. The color range indicating the usage counts. The error bars in Fig. 3A represent mean ± sd. *p < 0.05 was considered significant. The samples include HC (n = 11), AS (n = 5) and SM (n = 35, total 33 patients and two patients were sampled twice at different disease stages).
mononuclear cells (PBMCs) and illustrated the immune landscape of B cells (Fig. 1A). We subsequently completed the scBCR-seq and analyzed the abundance and diversity of the BCR repertoires in eleven health controls (HC), five asymptomatic subjects (AS), and 33 symptomatic COVID-19 patients (SM) that included ten patients with severe disease (SD) and ten recovery patients from the severe disease (SDR) (Fig. 1A, Supplementary Table S1). The T-distributed stochastic neighbor embedding (t-SNE) plots showed that four clusters of B cells: naive B, memory B, activated B and marginal zone B cells. Moreover, representative marker genes included ITGAM (naive B), IGHJ3 (memory B), CRIP2 and ITGAM (activated B) as well as FCRL5 and FGR (marginal zone B) (Fig. 1B) [31]. The ridge plot shows the distribution identity of these marker genes in four clusters of B cells (Fig. 1C). Furthermore, more than 90% BCRs were matched to the B cells, and there was no biased proportion of BCRs from a single donor here (Supplementary Fig. S1A, 1B, Supplementary Table S1). We then calculated the relative percentage of B cells in each group. While some of these groups (e.g., SM vs AS, SDR vs SD) had relatively higher percentage (mean) of B cells, there were no significant difference among these groups (Fig. 1D, 1E). However, when we compared the routine laboratory testing results, we found that CD19⁺ B cells were significantly higher in AS than SM (383.8 ± 116.5, 199.1 ± 147.9, p = 0.036), and the count of CD19⁺ B cells decreased in SD compared with SDR patients (244.6 ± 181.0, 124.4 ± 88.6, p = 0.09) (Table 1). These results suggested that B cell alternations were associated with SARS-CoV-2 infections.

4.2. Identification of BCR clonal expansions in SM and SDR patients

Here, we also performed scBCR-seq to assess the status of clonal expansions in PBMCs derived from the COVID-19 patients and the healthy controls. We found that more clonal expansions occurred in SM group compared with HC and AS. Of note, SDR had much more clonal cells than SD patients. In addition, memory B and activated B cells exhibited high clonal expansions in SM and SDR (Fig. 2A, Supplementary Fig. S1C). We revealed that most SM patients exhibited highly expanded clones at the individual level (more than 10%) compared with the healthy or asymptomatic subjects (less than 5%) (Fig. 2B, Supplementary Table S2). We also observed that the recovery patients (SDR) had higher expanded clones than the SD patients (Fig. 2B, Supplementary Table S2). Moreover, quantification of the most highly expanded clones for each patient and healthy control found that SM had the highest percentages of the maximum clones (2/15) compared with HC and AS group. Furthermore, we revealed that the maximum clone numbers were higher in SDR (1/6) than SD group (Fig. 2C, D).

4.3. B-cell immune repertoire in asymptomatic and symptomatic COVID-19 patients

To study the specific changes and preference gene usages of BCR in COVID-19 patients with various clinical presentations, we compared the usage of VDJ genes in asymptomatic subjects (AS), symptomatic patients (SM) and health controls (HC). We found that IGVK2-24 and IGVK3-7 were highly enriched in the AS subjects compared with the SM and HC groups, whereas IGHV1-9 was preferred in the HC group (Fig. 3A, Supplementary Fig. S2A-C, Supplementary Table S3). CDR3 (complementary determining region 3) is the most important region of an antibody structure, which interacts with the antigen. Usually, CDR3 region has a conserved length of amino acids with antigen specificity. To explore the relationship of gene usage and V-J pairs with the clonal expansions resulting from SARS-CoV-2 infection, we identified the CDR3 structures of IgG-BCR expressed on IGVK1-9. We found that eleven amino acids in IGVK1-9 had highest frequencies at specific lengths from HC and AS (Fig. 3B). Furthermore, IGHV3-23 IGJ4, IGLV1-44 IGLJ3 and IGVK3-20 IGJ1 were the top V-J pairs in all groups, and the highest frequency was observed in SM (Fig. 3C, Supplementary Fig. S2D-F, Supplementary Table S4-6). The Upset View showed the common and unique VJ pairs in HC, AS and SM group. There were three IGH, five IGK and two IGL VJ pairs in AS patients, respectively. However, the frequency of these VJ pairs was low (Fig. 3D, Supplementary Table S4-6). We also analyzed the expression of IGHV3-23 and IGJ4 pair in AS and SM, and found ten (12, 13, 15–22 amino acids) and four (14, 16, 17 and 20 amino acids) types of amino acids lengths in AS and SM group, respectively (Fig. 3E, Supplementary Fig. S2G), suggesting that these CDR3 region structures had a high variability and SARS-CoV-2 specificity. Additional analysis coupled with the scRNA-seq data demonstrated that IGHV3-23 was differentially expressed in naive and memory B cells among four clusters (Supplementary Fig. S2K, Supplementary Table S8). Moreover, the skewed usage and the frequency of VH/VL combinations were detected among different groups. For instance, the IGHV4-44 IGLV1-44 combination was highly enriched in AS and SM group, the IGHV4-34 IGVK19-39 combination had the highest frequencies in SM (Fig. 3F, Supplementary Fig. S2H), whereas IGLV3-9 IGVH6-1 and IGLV3-27 IGVH7-4–1 combinations were specifically enriched in AS (Supplementary Fig. S2J, Supplementary Table S7), suggesting that the expanded BCR repertoire were associated with the viral infection states. Some of these findings are consistent with the published data (CoV-AbDab) [32], demonstrating the VJ gene usage and bias in different groups (Table 2).

Table 2

| BCR clonotypes identified in this study compared with the antibodies reported in the CoV-AbDab. |
|---|---|---|---|---|---|---|
| ID | IG V gene | IG J gene | Binds | Protein & Epitope | Group | Source |
| No.1 | IGVK2-24 | IGKJ2 | SARS-CoV2 | Unknown | AS | human | This study |
| No.2 | IGVK3-7 | IGKJ1 | SARS-CoV2 | Unknown | AS | human | This study |
| No.3 | IGVK3-7 | IGKJ1 | SARS-CoV1 | Unknown | mouse | Michael Gubbins et al. [40] |
| No.4 | IGVK3-20 | IGKJ1, 2 | SARS-CoV2 | Unknown | SM, SDR | human | This study |
| No.5 | IGVK1D-39 | IGKJ1 | SARS-CoV2 | Unknown | SM, SDR | human | This study |
| No.6 | IGLV1-44 | IGLJ3 | SARS-CoV2 | Unknown | SM, SDR | human | This study |
| No.7 | IGLV2-14 | IGLJ2 | SARS-CoV2 | Unknown | SM, SDR | human | This study |
| 12E; 15D et al | IGLV1-44 | IGLJ3 | SARS-CoV2 | Unknown | SM, SDR | human | [43,44] |
| BLN2 | IGLV2-14 | IGLJ2 | SARS-CoV2 | Unknown | SM, SDR | human | This study |
| RA | IGLV2-14 | IGLJ2 | SARS-CoV2 | Unknown | SM, SDR | human | This study |

Note: This Supplementary Table Shows the V J gene usage and bias in different groups. The clonotypes were compared with public data in CoV-AbDab [32]. NA: not available. HC: Healthy control; AS: Asymptomatic subjects; SM: Symptomatic patients, including the patients with severe disease (SD) and the patients recovered from the severe disease recovery (SDR).
4.4. The skewed usage of BCR repertoire in severe (SD) and recovery (SDR) COVID-19 patients

To study the biased V(D)J rearrangements of the BCR in the severe (SD) and recovery (SDR) COVID-19 patients, we compared the usage of V(D)J genes between these groups. IGLV2–14 and IGHV1–69D were dominant IGV genes in SD, whereas IGLV4–3 and several IGHV3 family genes (IGHV3–13, IGHV3–20, IGHV3–43) were enriched in SDR (Fig. 4A, B, Supplementary Fig. S3A-C, Supplementary Table S3). We also performed dock analysis and showed that IGHV3–13 may bind to the S protein of SARS-CoV-2 (Supplementary Table S3). Moreover, we compared the common and unique VJ pairs between SD and SDR group, and found that the V-J pairs of IGKV1D39-IGHJ1, IGHV4–59-IGHJ4, IGHV4–39.IGHJ4 and IGHV3–23.IGHJ4 enriched in SDR, indicating that B cells might have undergone unique and specific V(D)J rearrangements in the recovery patients (Fig. 4C-D, Supplementary Fig. S3D, 3E, Supplementary Table S4–6). Furthermore, we identified the IGHV3–23.IGHJ4 and IGLV2–14.IGLJ2 expressed in the CDR3 of BCR in

Fig. 4. The skewed usage of BCRs in SD and SDR COVID-19 patients. (A) The differential usage of IGH V genes in SD and SDR groups. (B) The differential usage of IGL V genes in SD and SDR groups. (C) The Heatmaps showing the VJ gene pairs of IGH, IGK and IGL in SDR. The color range indicating the usage counts of specific V-J gene pairs. (D) The Venn diagram showing the common and unique VJ pairs of IGH, IGK, and IGL in SD and SDR groups. (E) The motif logo plot showing the CDR3 length distribution and amino acids frequency of IGHV3–23.IGHJ4 in SD and SDR. (F) The motif logo plot showing the CDR3 length distribution and amino acids frequency of IGLV2–14.IGLJ2 in SD and SDR. (G) The Heatmaps showing the V gene heavy and light chain combinations in SDR. The color range indicating the usage counts. The error bars in Fig. 4A and B represent mean ± sd. p < 0.05 was considered significant. The samples include SD (n = 11, total 10 patients and one patient was sampled twice) and SDR (n = 11, total 10 patients and one sample was sampled twice).
and the operational funds from The First Affiliated Hospital of Xi’an Jiaotong University (CZ and BS).

cRediT authorship contribution statement

Junpeng Ma: Methodology, Writing – original draft, Writing – review & editing. Han Bai: Methodology, Writing – original draft, Writing – review & editing. Tian Gong: Methodology, Writing – original draft, Writing – review & editing. Weikang Mao: Methodology, Writing – original draft. Yijun Nie: Methodology. Xuan Zhang: Methodology. Yanyan Da: Methodology. Xiaorui Wang: Methodology. Hongyu Qin: Methodology. Qiqi Zeng: Methodology. Fang Hu: Methodology. Xin Qi: Methodology. Bingyin Shi: Conceptualization, Resources, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition. Chenguang Zhang: Conceptualization, Methodology, Resources, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition, Project administration.

Declarations of Competing Interest

The author Weikang Mao is employed by the LC-BIO TECHNOLOGIES (HANGZHOU) CO., LTD., China. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

We are extremely grateful to all the patients and their families for participation of this study and providing the valuable information. We also thank our colleagues at The First Affiliated Hospital of Xi’an Jiaotong University and The Renmin Hospital of Wuhan University for their kind help and strong support throughout the course of this study. We would like to acknowledge all other supporting entities, including the FP7 WeNMR (project# 261572), H2020 West-Life (project# 675858), the EOSC-hub (project# 777536), and the EGI-ALE (project# 101017567). In particular, the European e-Infrastructure projects for the use of their web portals, which make use of the EGI infrastructure with the dedicated support of CSENET-MCC, INFN-PADOVA-STACK, INFNLN-2, NCG-INGRID-PT, TW-NCHC, CESGA, IFCAC-LCG2, UA-BITP, SURFsara and NIKHEF.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.imlet.2022.08.006.

References

[1] WHO Coronavirus (COVID-19) Dashboard. https://www.who.int/. [2] P.S. Arunachalam, A.C. Walls, N. Golden, C. Atyeo, S. Fischinger, C. Li, P. Aye, M. J. Navarro, L. Lai, V.V. Edara, et al., Adjuvancing a subunit COVID-19 vaccine to induce protective immunity, Nature 594 (2021) 253–258, https://doi.org/10.1038/s41586-021-03530-2. [3] Q. Li, J. Wu, J. Nie, L. Zhang, H. Hao, S. Liu, C. Zhao, Q. Zhang, H. Liu, L. Nie, et al., The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity, Cell 182 (2020) 1284–1294, https://doi.org/10.1016/j.cell.2020.07.012, e1289. [4] F. Campbell, B. Archer, H. Laurenson-Schafer, Y. Jinmai, F. Konings, N. Batra, B. Pavlin, V. Kandemiae, M.D. Van Kerkhove, T. Jombart, et al., Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021, Euro Surveill. 26 (2021), 2100509, https://doi.org/10.2807/1560-7917.ES.2021.26.21.2100509. [5] World Health Organization. Weekly epidemiological update on COVID-19: 11 May 2021. Available online: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19—11-may-2021. [6] Y. Araf, A. Akter, Y.D. Tang, R. Fatemi, S.A. Parvez, C. Zheng, G. Hossain, Omicron variant of SARS-CoV-2: genomics, transmissibility, and responses to current COVID-19 vaccines, J. Med. Virol. 94 (5) (2022) 1825–1832, https://doi.org/10.1002/jmv.27508. Epoch ahead of print.
