



## OPEN ACCESS

## EDITED BY

Anamika Patel,  
Emory University, United States

## REVIEWED BY

Aditi Agrawal,  
Chan Zuckerberg Biohub, United States  
David Forgacs,  
Sanofi U.S., United States

## \*CORRESPONDENCE

Gunilla B. Karlsson Hedestam  
[✉ Gunilla.Karlsson.Hedestam@ki.se](mailto:Gunilla.Karlsson.Hedestam@ki.se)

## SPECIALTY SECTION

This article was submitted to  
Virus and Host Immunity,  
a section of the journal  
Frontiers in Virology

RECEIVED 20 December 2022

ACCEPTED 31 January 2023

PUBLISHED 01 March 2023

## CITATION

Pushparaj P, Nicoletto A, Castro Dopico X, Sheward DJ, Kim S, Ekström S, Murrell B, Corcoran M and Karlsson Hedestam GB (2023) Frequent use of IGHV3-30-3 in SARS-CoV-2 neutralizing antibody responses.  
*Front. Virol.* 3:1128253.  
doi: 10.3389/fviro.2023.1128253

## COPYRIGHT

© 2023 Pushparaj, Nicoletto, Castro Dopico, Sheward, Kim, Ekström, Murrell, Corcoran and Karlsson Hedestam. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Frequent use of IGHV3-30-3 in SARS-CoV-2 neutralizing antibody responses

Pradeepa Pushparaj<sup>1</sup>, Andrea Nicoletto<sup>1</sup>, Xaquín Castro Dopico<sup>1</sup>, Daniel J. Sheward<sup>1</sup>, Sungyong Kim<sup>1</sup>, Simon Ekström<sup>2</sup>, Ben Murrell<sup>1</sup>, Martin Corcoran<sup>1</sup> and Gunilla B. Karlsson Hedestam<sup>1\*</sup>

<sup>1</sup>Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden,

<sup>2</sup>Department of Biomedical Engineering, Lund University, Lund, Sweden

The antibody response to SARS-CoV-2 shows biased immunoglobulin heavy chain variable (IGHV) gene usage, allowing definition of genetic signatures for some classes of neutralizing antibodies. We investigated IGHV gene usage frequencies by sorting spike-specific single memory B cells from individuals infected with SARS-CoV-2 early in the pandemic. From two study participants and 703 spike-specific B cells, the most used genes were IGHV1-69, IGHV3-30-3, and IGHV3-30. Here, we focused on the IGHV3-30 group of genes and an IGHV3-30-3-using ultrapotent neutralizing monoclonal antibody, CAB-F52, which displayed broad neutralizing activity also in its germline-reverted form. IGHV3-30-3 is encoded by a region of the IGH locus that is highly variable at both the allelic and structural levels. Using personalized IG genotyping, we found that 4 of 14 study participants lacked the IGHV3-30-3 gene on both chromosomes, raising the question if other, highly similar IGHV genes could substitute for IGHV3-30-3 in persons lacking this gene. In the context of CAB-F52, we found that none of the tested IGHV3-33 alleles, but several IGHV3-30 alleles could substitute for IGHV3-30-3, suggesting functional redundancy between the highly homologous IGHV3-30 and IGHV3-30-3 genes for this antibody.

## KEYWORDS

SARS-CoV-2, immunoglobulin germline genes, allelic diversity, copy number variation, antibodies

## Introduction

Since the emergence of SARS-CoV-2 at the end of 2019, research groups around the world have characterized the antibody responses to the virus surface glycoprotein, spike (S), using serological assays and by isolating monoclonal antibodies (mAbs). It is now well-established that the receptor binding domain (RBD) of S is the main target for neutralizing antibodies (1–3), though rare neutralizing antibodies that target the N-terminal domain (NTD) or subunit 2 (S2) of S have also been reported (4–6).

In November 2020, the first variant of concern (VoC), Alpha, emerged. Alpha was followed by a series of VoCs, many of which displayed partial resistance to neutralizing antibodies, so-called immune escape variants. The sub-lineages of Omicron that now contribute to the majority of infections, are considerably more neutralization resistant than the ancestral virus (7–11). The fact that antibody escape variants emerge in populations with a high level of seropositivity, obtained from either infection, vaccination or a combination of both, demonstrates that neutralizing antibodies exert selection pressure on the virus.

The antibody response to S is highly polyclonal and engages many different IGHV genes (12–16). Several studies have reported the over-representation of certain IGHV genes, such as IGHV1-69, IGHV3-30 and IGHV3-53, in SARS-CoV-2 S-binding antibodies compared to in the total antibody repertoire (12, 14–18). Of these, IGHV1-69 and IGHV3-30 are also often used in antibody responses to other antigens as they are frequent in the IgM repertoire, and thus there is a high likelihood that naïve B cells encoding one of these genes will encounter antigen. In contrast, IGHV3-53 is frequently used in the response to SARS-CoV-2 RBD despite being infrequent in the total IgM repertoire, suggesting preferential usage (19). Studies have shown that germline-encoded amino acid motifs in IGHV3-53 make direct contacts with the RBD, explaining the potency of this class of antibodies despite low levels of somatic hypermutation (SHM) (3, 19, 20).

There are 52 known IGHV genes in humans (21, 22), many of which display extensive allelic and copy number variation between individuals (22–25). Thus, the IGHV genotypes and haplotypes across individuals are highly personal and, together with the IGHD, IGHJ and the light chain genes, shape the composition of the expressed IgM repertoire in each individual.

The IGHV region encompassing IGHV3-30, IGHV3-30-3, IGHV3-30-5 and IGHV3-33 are particularly variable between individuals. These genes are highly similar at the nucleotide level, consistent with a recent evolutionary origin due to gene duplication events. IGHV3-33 is more distantly related at the nucleotide level but differs from IGHV3-30 and IGHV3-30-3 by only one or two amino acids. These genes, often referred to as the IGHV3-30 group of genes, are some of the most highly used IGHV genes in SARS-CoV-2 neutralizing antibodies (15, 16, 18). Several antibodies that use these genes bind and neutralize the virus even in their germline state, suggesting a role for the germline-encoded residues in epitope binding (17). The frequent usage of these closely related yet polymorphic genes in the antibody response to SARS-CoV-2 raises the question of how this germline-encoded diversity influences antigen binding properties of antibodies.

Here, we studied of a cohort of healthcare workers who were infected with SARS-CoV-2 during spring 2020, from which we recently reported the isolation of a set of mAbs including the ultrapotent mAb, CAB-F52, which used IGHV3-30-3 (16). Here, we characterized this mAb further by reverting the mAb heavy chain (HC) to the germline sequence to elucidate the role of SHM for the neutralizing activity. We also investigated the presence of IGHV3-30 family genes and alleles in this cohort and in an additional cohort of healthy individuals to better understand the germline-encoded variation in the region around IGHV3-30-3. We observed a high allelic and copy number variation of these genes with many

individuals lacking the IGHV3-30-3 gene on both chromosomes. Finally, we performed IGHV-region swaps to ask if IGHV3-30 or IGHV3-33 can substitute for IGHV3-30-3 in CAB-F52, and thus potentially compensate for the frequently observed homozygous absence of IGHV3-30-3 in the neutralizing Ab response to SARS-CoV-2.

## Materials and methods

### Sample collection

All data were collected as previously described in Pushparaj et al. (16). In brief, healthcare staff at the Karolinska University Hospital in Stockholm, Sweden, were invited to take part in an antibody characterization study following a SARS-CoV-2 infection. PCR-positive participants that were infected in May 2020 provided blood samples in December 2020. Informed consent was obtained from all the study participants as part of ethics approvals (Decisions# 2020-01620, 2020-02881 and 2020-05630) from the National Ethical Review Agency of Sweden.

### Cell culture

HEK293F cells, HEK293T cells (ATCC CRL-3216), and human ACE-2 expressing HEK293T-ACE2 cells were cultured in Dulbecco's Modified Eagle Medium (glucose, with sodium pyruvate) containing with 10% FBS, 100 U/ml penicillin, and 100 µg/ml streptomycin. Cells were maintained in a humidified incubator under culture conditions of 5% CO<sub>2</sub> at 37°C. Of these, HEK293F cells were used for mAb expression while the others were used for pseudo virus neutralization assay.

### Cloning of germline-reverted mAbs

Using Gibson assembly mixture, germline reverted HC VDJ sequences were cloned into expression vectors with the human constant regions IgG1 (26). Briefly, VDJ germline sequences were designed and synthesized with overhangs matching the ends of the linearized vectors. The Gibson reaction master mix (New England Biolabs) consisted of 50 ng of vector, 30 ng of insert in a 20 µl reaction. This reaction mixture was then incubated at 50°C for 1 hour and transformed into XL10-Gold ultracompetent cells according to the manufacturer's protocol (Agilent Technologies). Colonies were screened and the positive ones were identified by Sanger sequencing (Genewiz) and expanded using plasmid Midi-prep kit (Qiagen) according to the manufacturer's instructions.

### Expression and purification of germline-reverted mAbs

Germline-reverted CAB-F52 was generated by co-transfecting the germline HCs with the mature CAB-F52 LC. Antibody expression was performed by co-transfecting 15 µg of each HC and LC plasmid using 30 ml of FreeStyle 293-F cells at a concentration of 1 million

cells/ml ( $\geq 90\%$  viability) maintained in FreeStyle 293-F medium (Life Technologies) with 30  $\mu$ l of Max reagent (Life Technologies). Purification of antibodies was performed 7 days post-transfection using Protein G Sepharose columns (GE Healthcare) and purity was confirmed by SDS-PAGE.

## Preparation of Fab fragments

Fab fragments were obtained by IgG digestion using immobilized papain (Pierce Fab Preparation Kit, ThermoFisher Scientific) according to the manufacturer's instructions.

## Pseudovirus neutralization assay

Expression vectors encoding the spike proteins of various SARS-CoV-2 strains, including B.1 (D614G), B.1.1.7 (Alpha), B.1.351 (Beta), P.1 (Gamma), B.1.617.2 (Delta), and B.1.621 (Mu) were obtained from the G2P-UK National Virology consortium and Barclay Lab (Imperial College London, UK). B.1.1.529 spike (Omicron) strain was also included and was obtained by molecular cloning from an infected individual as previously (10). To generate pseudovirus particles, the spike-encoding plasmid, a HIV gag-pol packaging plasmid (Addgene #8455), and a lentiviral transfer plasmid encoding firefly luciferase (Addgene #170674) were co-transfected into HEK293T cells using polyethylenimine. The assays were performed by incubating pseudoviruses with serially diluted mAbs for 60 minutes at 37°C, then adding 10,000 HEK293T-ACE2 cells to each well. The cells were incubated for 44–48 hours, and the neutralizing activity of the mAbs was determined by measuring luminescence using Bright-Glo (Promega) on a GloMax Navigator Luminometer (Promega). The neutralizing activity was calculated relative to the average of 8 control wells infected in the absence of serum.

## Hydrogen–deuterium exchange mass spectrometry

The experimental setup consisted of a LEAP H/D-X PAL™ platform, used for automated sample preparation, interfaced to an LC-MS system, comprising an Ultimate 3000 micro-LC coupled to an Orbitrap Q Exactive Plus MS. HDX was performed on RBD with and without Fab obtained from the digestion of CAB-F52 in 10 mM PBS. Apo state (unbound RBD) and epitope mapping (RBD bound to either one of the Fabs) samples were incubated for  $t = 0, 30, 300, 3000$  at 20°C in PBS or HDX labelling buffer of the same composition prepared in D<sub>2</sub>O. The analysis was performed in a single, continuous run and 3 replicates were run for each state and timepoint. Quenching of the labelling reaction was achieved by dilution with 1% TFA, 0.4 M TCEP, 4 M urea, at 1°C. The quenched sample was directly injected and subjected to online pepsin digestion at 4°C. A flow of 50  $\mu$ L/min 0.1% formic acid was applied for 4 minutes for online digestion and trapping of the samples. Digestion products were subjected to online solid phase extraction and washing with 0.1% FA for 60s on a PepMap300 C18 trap column, which was switched in-line with a reversed-phase analytical column, Hypersil GOLD. Separation was performed at 1°C, with

mobile phases of 0.1% formic acid (A) and 95% acetonitrile/0.1% formic acid (B), using a gradient of 5–50% B over 8 minutes and then from 50 to 90% B for 5 minutes. Separated peptides were analyzed on a Q Exactive Plus MS, equipped with a heated electrospray source (HESI) operated at a capillary temperature of 250°C with sheath gas 12 au, auxiliary gas 2 au, and sweep gas 1 au. For HDX analysis MS full scan spectra were acquired at 70000 resolution, automatic gain control (AGC) 3e6, Max ion injection time 200ms and scan range 300–2000 m/z. For identification of generated peptides separate undeuterated samples were analyzed using data dependent MS/MS.

## HDX analysis

A library of peptides containing peptide sequence, charge state and retention time was generated for the HDX analysis by running pepsin-digested, undeuterated samples against the RBD sequence on PEAKS Studio X (Bioinformatics Solutions Inc.). HDX data analysis and visualization were performed using HDEaminer v3.1.1 (Sierra Analytics Inc.). Bound states were analysed in comparison to apo states, using a single charge state per peptide. This analysis allowed only for EX2 kinetics, and the assumption was made that the two first residues of a peptide were unable to hold deuteration. Due to the comparative nature of the measurements, the deuterium incorporation levels for the peptic peptides were derived from the observed relative mass difference between the deuterated and non-deuterated peptides without back-exchange correction using a fully deuterated sample (27). The spectra for all time points were manually inspected; low scoring peptides, obvious outliers and peptides for which retention time correction could not be made consistent were removed. The HDX-MS experiments presented here are reported in accordance with established recommendations (28). A summary of the HDX experimental detail can be found in [Supplementary Table 1](#). In addition, the mass spectrometry raw data and HDEaminer analysis files have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository (29) with the dataset identifier PXD031945.

## Analysis of IgM libraries from Gidoni et al.

Individual VDJ databases were produced using IgDiscover v0.12.4.dev22+g0bc3365 with the IMGT release 202209-1 (28 February 2022) as the starting database from bulk IgM sequences reported in Gidoni et al. (24).

## Quantification and statistical analysis

Neutralizing IC<sub>50</sub> values were calculated in Prism v9 (GraphPad Software) by fitting a four-parameter logistic curve, to neutralization by serial 3-fold dilutions. Neutralization was bounded between 0 and 100%. KD and IC<sub>50</sub> correlations were determined using Pearson's R (GraphPad Software). To determine the statistical significance between the groups in the bar plot showing combined IGHV3-30/IGHV3-30-3 frequencies in IgM libraries, Wilcoxon rank sum Q1 exact test was used and the p value was determined to be 0.002.

## Data availability

All data were collected as previously described (16). The HC and LC sequences of CAB-F52 and the IgM repertoire data from the 14 donors are available at GenBank under the accession numbers ON086926 and ON086941. Repertoire data is available from SciLifeLab at: <http://doi.org/10.17044/scilifelab.19317512>. The HDX data have been deposited to the ProteomeXchange Consortium *via* the PRIDE (29) partner repository with the dataset identifier PXD031945.

## Software availability

The IgDiscover software can be found at <http://docs.igdiscover.se/en/stable/>. The plot allele module and the rep-seq analysis tools can be found under IgDiscover v0.12.4.dev22+g0bc3365.

## Results

### SARS-CoV-2 spike-specific memory B cell sequencing revealed preferential IGHV gene usage

We used a previously described cohort of healthcare workers ( $n = 14$ ) (16), who tested positive for SARS-CoV-2 by RT-PCR in Spring 2020. As described, peripheral blood mononuclear cells (PBMCs) were collected seven months after the primary infection and all study participants (SPs) were individually genotyped for their IG V, D and J allele content using the germline gene inference tool IgDiscover (23). Single S-specific IgG memory B cells were sorted from two SPs and mAbs were isolated as described (Figure 1A) (16). A total of 703 S-specific HCs were obtained from the sorted cells. Examination of IGHV gene usage showed that 93 of the 703 S-specific sequences (13%) used the IGHV3-30-3 gene, making it the second most utilized IGHV gene in these individuals after IGHV1-69. IGHV3-30 and IGHV3-33 were the third and fourth most used genes in S responses, respectively (Figure 1B). Preferential IGHV gene usage was particularly evident for IGHV3-30-3, where the percentage of HCs using this gene was markedly higher in the S-specific response than in total IgG libraries from the same individual as we showed previously (16).

Amplification of antibody V(D)J regions from the S-specific B cells was performed by single-cell RT-PCR using previously described primers (30) for production of a panel of mAbs, as recently described (16). From this panel, we selected the IGHV3-30-3-using mAb, CAB-F52, for further analysis since this mAb displayed exceptional neutralization potency with an IC<sub>50</sub> value of 0.005 against D614G. We tested the neutralization potency of CAB-F52 against several VoCs including Alpha, Beta, Gamma, Delta, Mu, and Omicron (BA.1) alongside a commercial IGHV3-30-3-using mAb, REGN10987 (31) and found that the potency of CAB-F52 was consistently higher when compared to that of REGN10987 against all VoCs except Omicron (BA.1) as demonstrated in the full dilution curves (Figure 1C) and by half maximal inhibitory concentration (IC<sub>50</sub>) values (Figure 1D). Several IGHV3-30-3-using neutralizing mAbs were also identified in other SARS-CoV-2 studies (13, 18, 32,

33), underscoring that such antibodies are frequently elicited following exposure to S (Figure 1E).

### CAB-F52 retains neutralization and breadth upon germline reversion

To map the epitope specificity of CAB-F52, we generated a Fab version of the antibody and used hydrogen-deuterium exchange mass spectrometry (HDX-MS) to map the binding site on the ancestral RBD. The results demonstrated that the main interaction sites for CAB-F52 involved amino acid (aa) residues 450-480 and 485-500 of the RBD (Figure 2A). When modelling the epitope on the RBD structure, the binding mode overlapped that of SARS-CoV-2 neutralizing class 2 or 3 antibodies described by Barnes et al. (20) (Figure 2B). This likely explains why CAB-F52 is unable to neutralize Omicron BA.1, which harbors the T478K and G446S mutations that are known to affect binding by class 2 and class 3 antibodies, respectively (34). However, CAB-F52 retains neutralization capacity against Beta, demonstrating that it is insensitive to E484K.

B cells undergo affinity maturation of their recombined antigen receptors through the germinal center (GC)-mediated SHM process. To evaluate the contribution of the HC SHM for the neutralizing activity displayed by CAB-F52, the antibody was reverted to its germline configuration, keeping the mature HC complementarity determining region 3 (HCDR3) intact. An alignment of the CAB-F52 IGHV region and the germline IGHV3-30-30\*01 sequence shows the positions in which the mature antibody had acquired SHM (Figure 2C). At the nucleotide level, the IGHV SHM level of CAB-F52 was 3.4%: a total of 6 aa substitutions, one of which was in the HCDR1 and one in the HCDR2. The aa substitutions in the HCs of CAB-F52 were reverted to the germline residues by replacing the VH-region with its assigned germline sequence, IGHV3-30-30\*01 keeping the HCDR3 intact (Figure 2D). One SHM substitution in the HC J gene outside of the HCDR3 was also reverted to the germline state. The germline-reverted HC was paired with the mature CAB-F52 LC. Assessment of neutralization of the germline-reverted CAB-F52 (gLCAB-F52) compared to the mature mAb against the ancestral SARS-CoV-2 strain demonstrated that gLCAB-F52 retained its ability to neutralize the virus, though the potency was reduced by approximately ten-fold compared to the mature CAB-F52 (Figure 2E). gLCAB-F52 also retained neutralization breadth against several VoCs, except Delta as shown in the titration curves (Figure 2E) and IC<sub>50</sub> values (Figure 2F). These results highlight the role of SHM for improving the potency of CAB-F52 such that it also neutralized the highly infectious Delta variant.

### Homozygous absence of IGHV3-30-3 affects the composition of the IgM repertoire

We next examined the germline-encoded variation in the IGHV3-30 family of genes in our study participants. We previously determined the IG V, D and J allele content in the 14 study participants (16) and here we employed the process of inferred

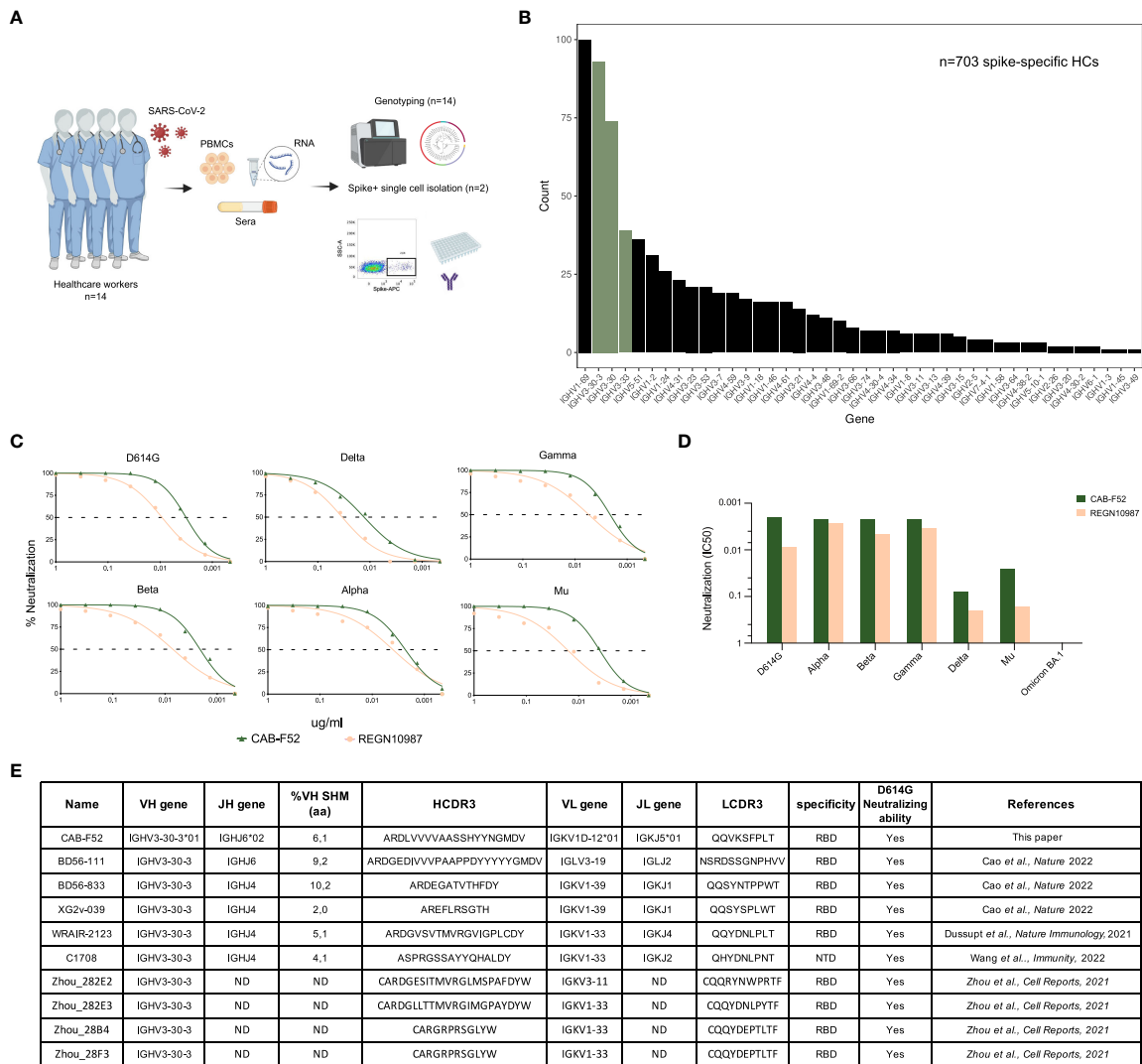
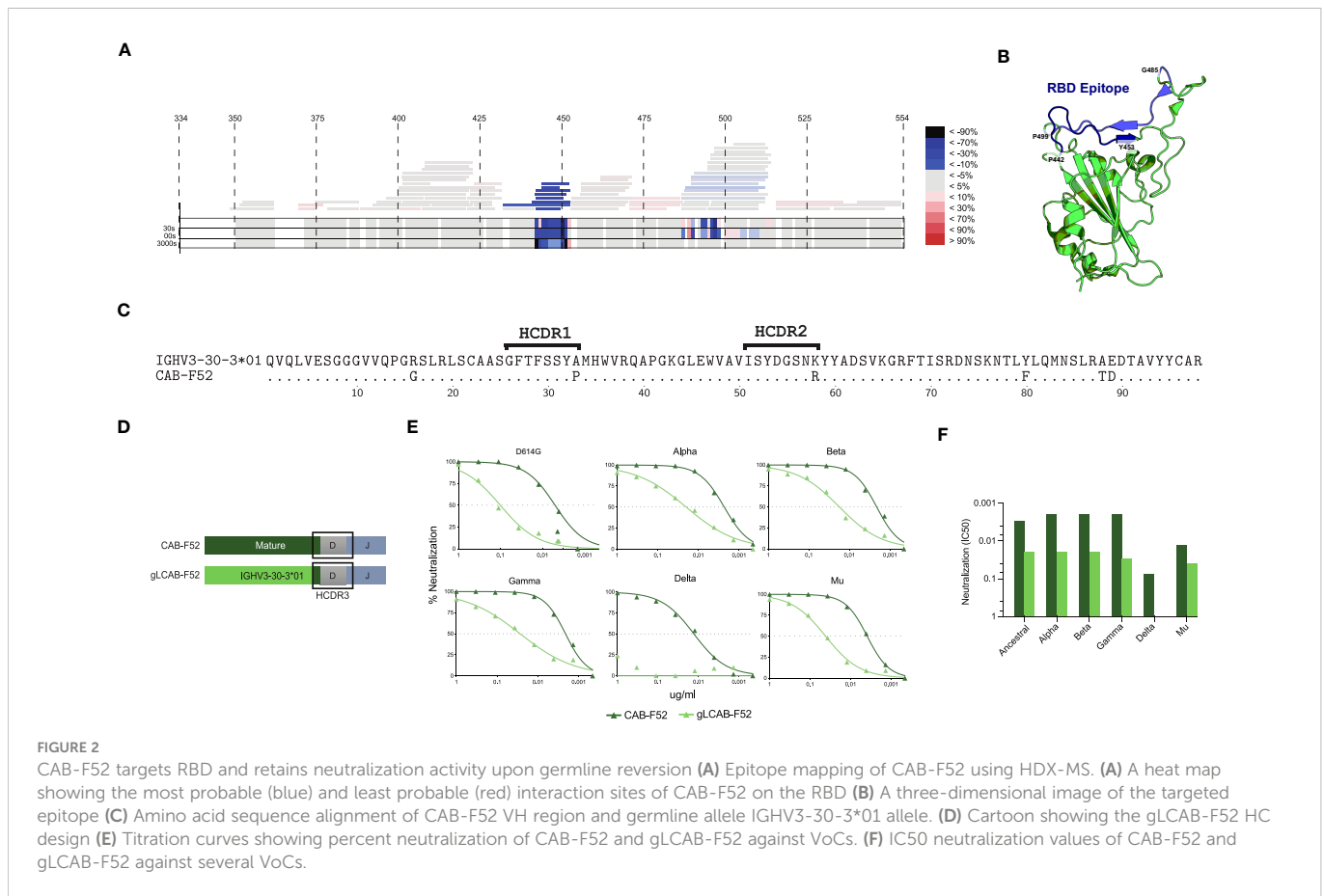


FIGURE 1

Spike HCs frequently use IGHV3-30 group of genes. (A) A schematic representation of the SARS-CoV-2 cohort consisting of 14 individuals. PBMCs, sera and RNA samples were isolated from each of these individuals. The genotyping of 14 study participants was done using the RNA sample. Two of the 14 individuals' PBMCs were used to isolate S-specific HCs and subsequent antibody expression. (B) A bar plot showing the IGHV gene usage of S-specific HCs with IGHV3-30 group highlighted in olive green. (C) Titration curves showing percent neutralization of CAB-F52 and REGN10987 against VoCs. (D) IC50 neutralization values of CAB-F52 and REGN10987 against the VoCs. (E) A table showing the genetic properties, specificity, and neutralization activity of CAB-F52 and several IGHV3-30-3-using mAbs from the CoVAbDab.

haplotype analysis to determine which alleles are present on which chromosome. This process relies on the fact that VDJ recombination occurs locally within the set of V, D and J genes present on the maternally or paternally derived IGH locus and therefore, heterozygous IGHJ or IGHD alleles can be used as anchors to position IGHV alleles to either of the two chromosomes (35–37). We found that SP01, SP02, SP06, SP08, and SP11 were heterozygous for IGHJ (IGHJ6\*02/IGHJ6\*03), while SP14 was heterozygous for IGHD2-21 (IGHD2-21\*01/IGHD2-21\*02), and SP04 and SP05 were heterozygous for IGHD3-10 (IGHD3-10\*01/IGHD3-10\*01\_S2851) (Figure 3A). Thus, we used these heterozygous IGHJ or IGHD alleles as anchors to haplotype eight of the 14 SPs. The inferred haplotype analysis demonstrated great inter-individual variability in the IGHV3-30 and IGHV3-30-3 allele content (Figures 3A, S1A). Several of the neighboring genes, IGHV4-30-2, IGHV4-30-4, and IGHV3-33, were also absent from these haplotypes, consistent with

high structural variation in this genomic region (Figure 3A). Two cases, SP02 and SP05, were homozygous for IGHV3-30\*04/IGHV3-30-3\*03, which cannot be distinguished as they are identical the nucleotide level. To investigate which gene this sequence was likely to represent, we examined the total IGHV3-30\*04/IGHV3-30-3\*03 expression counts present on each chromosome in SP02, a donor who also lacked IGHV3-30-3\*01 (Figure S1B). This revealed that there were more than twice as many IGHV3-30\*04/IGHV3-30-3\*03 read counts associated with the IGHJ6\*02 chromosome compared to the IGHJ6\*03 chromosome, suggesting that both IGHV3-30\*04 and IGHV3-30-3\*03 are present on the IGHJ6\*02 chromosome in SP02 while only one of the genes is present on the IGHJ6\*03 chromosome (Figure S1B). While we could not definitely confirm that this is the case, we concluded that the haplotype result of SP02 is consistent with a genotype containing both IGHV3-30\*04 and IGHV3-30-3\*03 on the IGHJ6\*02 anchored chromosome. SP02 therefore likely contained



one copy of the IGHV3-30-3 gene, while 4 of 14 donors lacked IGHV3-30-3 altogether (Figure 3B).

To understand the consequence of homozygous IGHV3-30-3 absence, we computed the combined frequencies of transcripts assigned to either IGHV3-30 or IGHV3-30-3 in the IgM libraries of each of the 14 SPs, having ensured that the sizes of the IgM libraries were comparable (Figure S1C). The results of the analysis showed that the individuals who lacked IGHV3-30-3 had distinctly fewer IGHV3-30 and IGHV3-30-3 IgM sequences (Figures 3C, S1D), which may impact their capacity to make CAB-F52-like antibodies and other responses involving these genes.

## IGHV3-30 group gene variation in an expanded study group and evaluation of functional redundancy

IGHV3-30, IGHV3-30-3 and IGHV3-33 are closely related at both nucleotide and amino acid levels, ranging from one to five polymorphic coding positions (Figure 4A). As mentioned above, IGHV3-30\*04 and IGHV3-30-3\*03 have identical nucleotide sequences and cannot easily be distinguished from one another. Here, to indicate this ambiguity, we labeled these sequences as IGHV3-30\*04/IGHV3-30-3\*03.

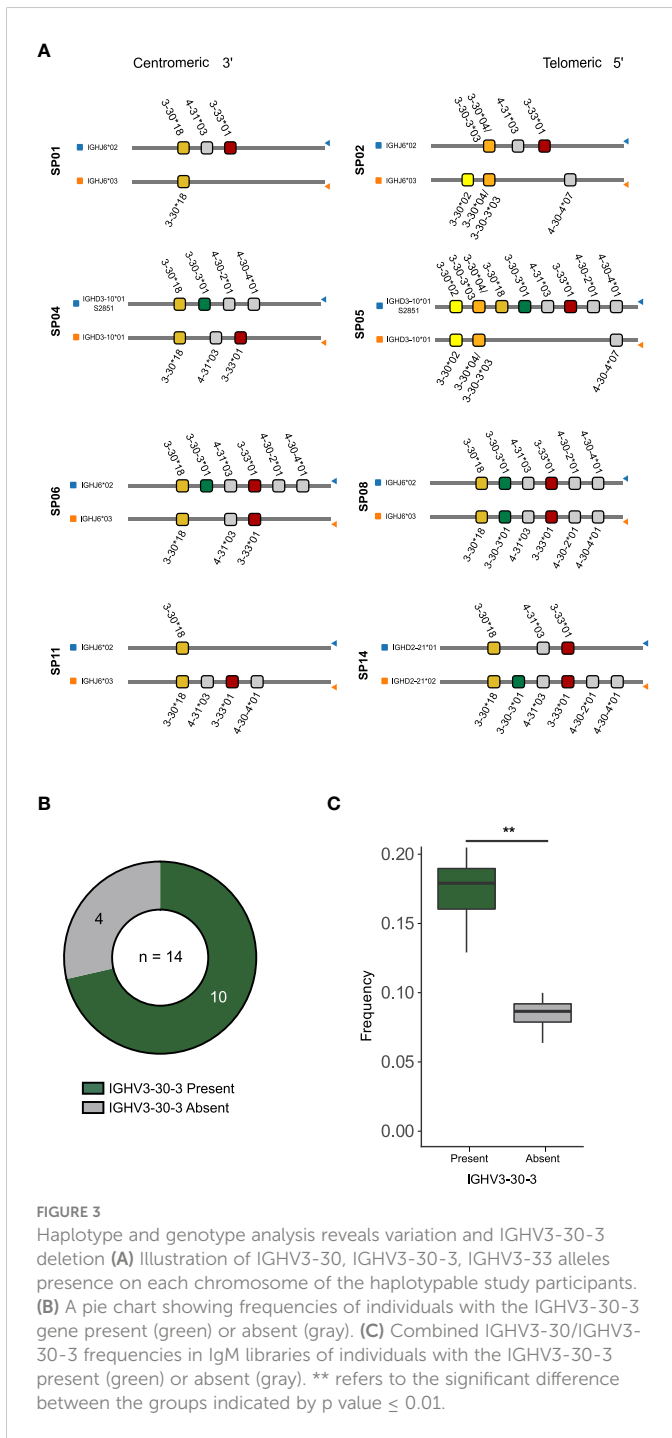
To expand the analysis of the IGHV3-30 group germline variation, we analyzed publicly available IgM repertoire data from an additional cohort of 32 individuals from an independent study (24). Examination of the combined 46 cases revealed that 9 individuals (19.5%) lacked IGHV3-30-3\*01 on both chromosomes

(Figure 4B). In addition to these 9 cases, 3 more individuals had the ambiguous IGHV3-30\*04/IGHV3-30-3\*03 sequence. We also examined the IGHV3-30 and IGHV3-33 allele content in the 46 cases. The IGHV3-30 allele content varied greatly between the cases (Figure 4B). IGHV3-30\*18 was the most common IGHV3-30 allele and was found in all but one case (Figure 4B). The IGHV3-33 gene was less variable, with the IGHV3-33\*01 allele present in 44 cases and homozygous absence of the gene only observed in two cases.

To test whether IGHV3-30-3 can be functionally replaced by other genes of the IGHV3-30 group we designed mutated versions of CAB-F52 HCs, replacing its VH-region with alleles IGHV3-30\*02, IGHV3-30\*03, IGHV3-30\*04, IGHV3-30\*18, IGHV3-33\*01 and IGHV3-33\*06/08 allele germline sequences and then expressed them as mAbs using mature CAB-F52 LCs (Figure 4C). All the IGHV3-30 allele-swapped mAbs retained neutralization except for the IGHV3-30\*02 demonstrating the functional similarities between the IGHV3-30 and many IGHV3-30-3 alleles. On the other hand, replacing the CAB-F52 VH-region with the IGHV3-33 alleles resulted in a complete loss of neutralization capability illustrating the functional impact of even a single germline-encoded amino acid difference (Figures 4D, E).

## Discussion

Inter-individual variation in immune genes is a characteristic of mammals. This leads to heterogeneity in responses at individual levels that benefits the population, for example during pandemics. The most



variable of the three loci that encode the IG genes is the IGH locus, present at the telomeric end of chromosome 14. This locus is characterized by extensive structural variation involving frequent segmental deletions, insertions and duplications of a subset of IGH genes, suggesting higher genomic instability in these locations (22–25). Regions within the IGH locus displaying especially high variability include the region spanning the IGHV3-30, IGHV3-30-3, and IGHV3-33 genes, resulting in several haplotypes that vary in length by over 50 kb (38–42). Furthermore, the IGHV3-30 gene displays an unusually high level of allelic variation, with additional novel alleles discovered as the IG germline allele content in additional individuals are defined.

IGHV3-30, IGHV3-30-3 and IGHV3-33 are some of the most highly utilized genes in the naive B cell repertoire. Since the Covid-19 pandemic, several independent studies that have characterized the S-specific antibody response indicate preferential usage of IGHV3-30 and IGHV3-30-3 in addition to several other genes including IGHV1-69, IGHV3-9, IGHV3-53, and IGHV5-51 (15–18). Many mAbs isolated from early time points following infection display low levels of SHM yet moderate-strong neutralizing activity, suggesting that their naïve, germline-encoded states have sufficient affinity to the targeted neutralizing epitopes (12, 17). Since some of these antibodies use IGHV genes that are highly variable between individuals, such as IGHV1-60, IGHV3-30, or IGHV3-30-3, it follows that differences in IG genotypes may impact the early immune control of primary SARS-CoV-2 infection.

The fact that an individual’s antigen-specific antibody repertoire can be influenced by IGHV germline gene composition and usage underscores the importance of personalized IG genotyping for a comprehensive understanding of B cell responses, as demonstrated by recent studies (16).

In the current study, we isolated an IGHV3-30-3 using mAb that was ultrapotent and showed breadth against VOCs, adding to the list of several additional IGHV3-30-3-using neutralizing antibodies reported in the literature. The functional characterization of CAB-F52 revealed its epitope on the spike RBD to be overlapping with the ACE-2-binding region and with that of class 2 and 3 antibodies (20). HC germline-reversion of CAB-F52 reduced the neutralizing activity compared to the mature antibody, but the germline-reverted antibody retained considerable potency and breadth. Due to the high sequence homology between IGHV3-30 and IGHV3-30-3, we asked if CAB-F52 could have used IGHV3-30, which is almost always present in individuals, sometimes in more than one copy per chromosome.

Studies have shown that gene copy numbers can influence IGHV gene expression in the repertoire (43, 44). Although homozygous absence of IGHV3-30-3 may be functionally replaced by IGHV3-30, our analysis of the combined frequencies of IGHV3-30 and IGHV3-30-3 sequences in IgM repertoires revealed that genomic absence of IGHV3-30-3 can greatly decrease the total number of IgM B cells encoding these highly similar V genes and may, therefore negatively influence responses that rely on IGHV3-30 group gene usage. In contrast, if several different IGHV3-30 genes and alleles show functional redundancy, a higher allelic copy number will result in some individuals showing a higher early response, particularly in situations where germline-encoded antibodies have high affinity to their target epitope in the absence of SHM.

Overall, the biological consequence of variation in IGHV genes remain incompletely understood. While we observed redundant function between several IGHV3-30 and IGHV3-30-3 alleles in the context of CAB-F52, despite their overall similarity, the IGHV3-33 alleles were sufficiently different that they could not substitute for the native IGHV3-30-3 germline gene. We observe that functional redundancy existing between similar V genes or alleles can, however, be abrogated by even single amino acid differences. This has important implications for the high level of IGHV allelic variation found in the human population. The requirements for specific germline gene and allele usage are expected to depend greatly on the specific binding modes, which vary greatly between different antibodies. Further cohort studies that investigate the impact of IG

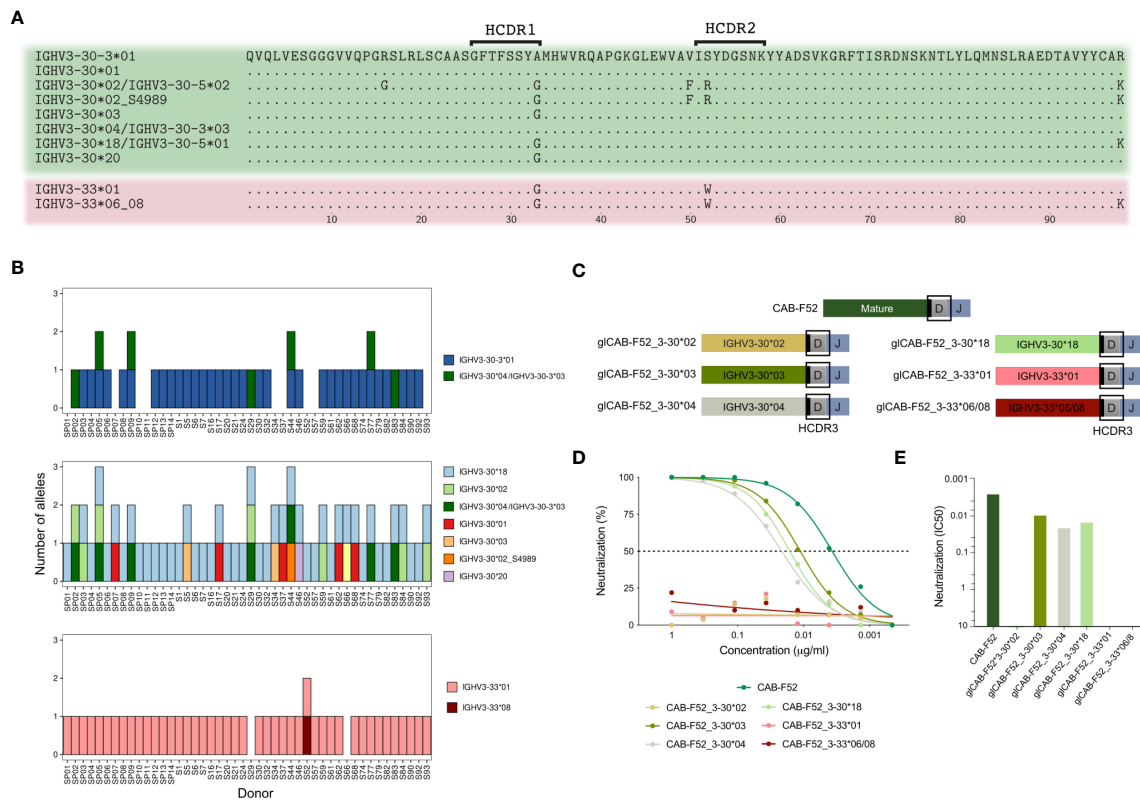


FIGURE 4

IGHV3-30 group of genes are highly variable and can impact CAB-F52 function. (A) Alignment of all the amino acid sequences of IGHV3-30, IGHV3-30-3 and IGHV3-33 alleles present in the extended group of 46 cases. (B) Allelic content in the 46 cases: IGHV3-30 (top graph), IGHV3-30-3 (middle graph) and IGHV3-33 (lower graph). The IGHV3-30\*04/IGHV3-30-3\*03 (dark green) is shown in both the IGHV3-30-3 and IGHV3-30 plots since they cannot be distinguished (C) Cartoons showing the design of the germline-reverted CAB-F52 HC variants. (D) Titration curves showing percentage neutralization of the allele-swapped CAB-F52 mAb variants. (E) IC<sub>50</sub> neutralization values of the allele-swapped CAB-F52 mAb variants.

germline gene variation to biological outcomes in the context of infections and vaccinations will offer further insight. Such studies will be particularly important for vaccine design aimed at eliciting specific classes of antibodies, so-called germline targeting vaccine strategies (45), where knowledge of IG allele frequencies in the targeted population is of particular importance.

## Limitations

A limitation of our study is sample size. The germline reversion and allele swapping experiments were only done in the context of a single mAb, CAB-F52, and the functional redundancy of the IGHV3-30 and IGHV3-30-3 alleles may not be true for other antibodies. Furthermore, we did not perform germline reversions of the CAB-F52 light chain V gene since the focus of the paper was on germline heavy chain V genes. For the genotyping studies, we addressed the small sample size by extending our cohort with the individuals from Gidoni et al. Nevertheless, further studies that couple genetic and functional analyses in larger cohorts are required to better understand functional consequences of germline IG gene variation.

## Data availability statement

All data were collected as previously described (16). The HC and LC sequences of CAB-F52 and the IgM repertoire data from the 14 donors are available at GenBank under the accession numbers ON086926 and ON086941. Repertoire data is available from SciLifeLab at: <http://doi.org/10.17044/scilifelab.19317512>. The HDX data have been deposited to the ProteomeXchange Consortium via the PRIDE (29) partner repository with the dataset identifier PXD031945. The IgDiscover software can be found at <http://docs.igdiscover.se/en/stable/>. The plot allele module and the rep-seq analysis tools can be found under IgDiscover v0.12.4.dev22+g0bc3365.

## Ethics statement

The studies involving human participants were reviewed and approved by National Ethical Review Agency of Sweden (Decisions# 2020-01620, 2020-02881 and 2020-05630). The participants provided their written informed consent to participate in this study.



## Author contributions

Conceptualization: PP, MC, GKH. Methodology: PP, MC, AN. Investigation: PP, AN, DS, XCD, SK, SE, MC, GKH. Visualization: PP, AN. Funding acquisition: GKH. Supervision: GKH. Writing – original draft: PP, GKH. Editing – original draft: all authors.

## Funding

Funding for this work was provided by a Distinguished Professor grant from the Swedish Research Council (agreement 2017-00968), a European Research Council (ERC) Advanced Grant (agreement 788016), and grants from the SciLifeLab National COVID-19 Research Program (VC-2021-0026) financed by the Knut and Alice Wallenberg Foundation and the SciLifeLab Pandemic Laboratory Preparedness (PLP) program 2022-2023 (VC-2022-0028) to GKH.

## Acknowledgments

We would like to thank the study participants, as well as Robert Dyrdak, Jan Albert, Joakim Dillner, and the personnel at the Centre for Clinical Laboratory Studies (Karolinska University Hospital), for generous help with participant recruitment and blood sampling. We gratefully acknowledge support from the Swedish National Infrastructure for Biological Mass Spectrometry (BioMS) and the

SciLifeLab (Sweden) Integrated Structural Biology platform, as well as Fondation Dormeur (Liechtenstein) for their generous contribution towards equipment.

## Conflict of interest

MC and GKH are founders of ImmuneDiscover Sweden.

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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fviro.2023.1128253/full#supplementary-material>

## References

- Ju B, Zhang Q, Ge J, Wang R, Sun J, Ge X, et al. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* (2020) 584(7819):115–9. doi: 10.1038/s41586-020-2380-z
- Walls AC, Fiala B, Schäfer A, Wrenn S, Pham MN, Murphy M, et al. Elicitation of potent neutralizing antibody responses by designed protein nanoparticle vaccines for SARS-CoV-2. *Cell* (2020) 183(5):1367–1382.e17. doi: 10.1016/j.cell.2020.10.043
- Barnes CO, West AP, Huey-Tubman KE, Hoffmann MAG, Sharaf NG, Hoffman PR, et al. Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. *Cell* (2020) 182(4):828–842.e16. doi: 10.1016/j.cell.2020.06.025
- Liu L, Wang P, Nair MS, Yu J, Rapp M, Wang Q, et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* (2020) 584(7821):450–6. doi: 10.1038/s41586-020-2571-7
- Cerutti G, Guo Y, Zhou T, Gorman J, Lee M, Rapp M, et al. Potent SARS-CoV-2 neutralizing antibodies directed against spike n-terminal domain target a single supersite. *Cell Host Microbe* (2021) 29(5):819–833.e7. doi: 10.1016/j.chom.2021.03.005
- Wang Y, Yuan M, Lv H, Peng J, Wilson IA, Wu NC. A large-scale systematic survey reveals recurring molecular features of public antibody responses to SARS-CoV-2. *Immunity* (2022) 55(6):1105–1117.e4. doi: 10.1016/j.immuni.2022.03.019
- Weisblum Y, Schmidt F, Zhang F, DaSilva J, Poston D, Lorenzi JC, et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *Elife* (2020) 9:e61312. doi: 10.7554/eLife.61312
- Cao Y, Yisimayi A, Jian F, Song W, Xiao T, Wang L, et al. BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by omicron infection. *Nature* (2022) 608(7923):593–602. doi: 10.1038/s41586-022-04980-y
- Planas D, Saunders N, Maes P, Guivel-Benhassine F, Planchais C, Buchrieser J, et al. Considerable escape of SARS-CoV-2 omicron to antibody neutralization. *Nature* (2022) 602(7898):671–5. doi: 10.1038/s41586-021-04389-z
- Sheward DJ, Kim C, Fischbach J, Sato K, Muschiol S, Ehling RA, et al. Omicron sublineage BA.2.75.2 exhibits extensive escape from neutralising antibodies. *Lancet Infect Dis* (2022) 22(11):1538–40. doi: 10.1016/S1473-3099(22)00524-2
- Wang Q, Iketani S, Li Z, Liu L, Guo Y, Huang Y, et al. Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *Cell* (2023) 186(2):279–286.e8. doi: 10.1016/j.cell.2022.12.018
- Brouwer PJM, Caniels TG, van der Straten K, Snitselaar JL, Aldon Y, Bangaru S, et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* (2020) 369(6504):643–50. doi: 10.1126/science.abc5902
- Cao Y, Su B, Guo X, Sun W, Deng Y, Bao L, et al. Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients' b cells. *Cell* (2020) 182(1):73–84.e16. doi: 10.1016/j.cell.2020.05.025
- Robbiani DF, Gaebler C, Muecksch F, Lorenzi JCC, Wang Z, Cho A, et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* (2020) 584(7821):437–42. doi: 10.1038/s41586-020-2456-9
- Sakharkar M, Rappazzo CG, Wieland-Alter WF, Hsieh CL, Wrapp D, Esterman ES, et al. Prolonged evolution of the human b cell response to SARS-CoV-2 infection. *Sci Immunol* (2021) 6(56):eabg6916. doi: 10.1126/sciimmunol.abg6916
- Pushparaj P, Nicoletto A, Sheward DJ, Das H, Castro Dopico X, Perez Vidakovic L, et al. Immunoglobulin germline gene polymorphisms influence the function of SARS-CoV-2 neutralizing antibodies. *Immunity* (2023) 56(1):193–206.e7. doi: 10.1016/j.immuni.2022.12.005
- Seydoux E, Homad LJ, MacCamy AJ, Parks KR, Hurlburt NK, Jennewein MF, et al. Analysis of a SARS-CoV-2-Infected individual reveals development of potent neutralizing antibodies with limited somatic mutation. *Immunity* (2020) 53(1):98–105.e5. doi: 10.1016/j.immuni.2020.06.001
- Zhou X, Ma F, Xie J, Yuan M, Li Y, Shaabani N, et al. Diverse immunoglobulin gene usage and convergent epitope targeting in neutralizing antibody responses to SARS-CoV-2. *Cell Rep* (2021) 35(6):109109. doi: 10.1016/j.celrep.2021.109109
- Yuan M, Liu H, Wu NC, Lee CCD, Zhu X, Zhao F, et al. Structural basis of a shared antibody response to SARS-CoV-2. *Science* (2020) 369(6507):1119–23. doi: 10.1126/science.abd2321
- Barnes CO, Jette CA, Abernathy ME, Dam KMA, Esswein SR, Gristick HB, et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* (2020) 588(7839):682–7. doi: 10.1038/s41586-020-2852-1
- Matsuda F, Ishii K, Bourvagnet P, Kuma Ki, Hayashida H, Miyata T, et al. The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. *J Exp Med* (1998) 188(11):2151–62. doi: 10.1084/jem.188.11.2151
- Watson CT, Steinberg KM, Huddleston J, Warren RL, Malig M, Schein J, et al. Complete haplotype sequence of the human immunoglobulin heavy-chain variable, diversity, and joining genes and characterization of allelic and copy-number variation. *Am J Hum Genet* (2013) 92(4):530–46. doi: 10.1016/j.ajhg.2013.03.004

23. Corcoran MM, Phad GE, Bernat NV, Stahl-Hennig C, Sumida N, Persson MAA, et al. Production of individualized V gene databases reveals high levels of immunoglobulin genetic diversity. *Nat Commun* (2016) 7(1):13642. doi: 10.1038/ncomms13642
24. Gidoni M, Snir O, Peres A, Polak P, Lindeman I, Mikocziova I, et al. Mosaic deletion patterns of the human antibody heavy chain gene locus shown by Bayesian haplotyping. *Nat Commun* (2019) 10(1):628. doi: 10.1038/s41467-019-08489-3
25. Rodriguez OL, Gibson WS, Parks T, Emery M, Powell J, Strahl M, et al. A novel framework for characterizing genomic haplotype diversity in the human immunoglobulin heavy chain locus. *Front Immunol* (2020) 11. doi: 10.3389/fimmu.2020.02136
26. Tiller T, Meffre E, Yurasov S, Tsuiji M, Nussenzweig MC, Wardemann H. Efficient generation of monoclonal antibodies from single human b cells by single cell RT-PCR and expression vector cloning. *J Immunol Methods* (2008) 329(1–2):112–24. doi: 10.1016/j.jim.2007.09.017
27. Engen JR, Wales TE. Analytical aspects of hydrogen exchange mass spectrometry. *Annu Rev Anal Chem (Palo Alto Calif)* (2015) 8:127–48. doi: 10.1146/annurev-anchem-062011-143113
28. Masson GR, Burke JE, Ahn NG, Anand GS, Borchers C, Brier S, et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods* (2019) 16(7):595–602. doi: 10.1038/s41592-019-0459-y
29. Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* (2022) 50(D1):D543–52. doi: 10.1093/nar/gkab1038
30. Vázquez Bernat N, Corcoran M, Hardt U, Kaduk M, Phad GE, Martin M, et al. High-quality library preparation for NGS-based immunoglobulin germline gene inference and repertoire expression analysis. *Front Immunol* (2019) 10:660. doi: 10.3389/fimmu.2019.00660
31. Hansen J, Baum A, Pascal KE, Russo V, Giordano S, Wloga E, et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* (2020) 369(6506):1010–4. doi: 10.1126/science.abd0827
32. Dussupt V, Sankhala RS, Mendez-Rivera L, Townsley SM, Schmidt F, Wiczorek L, et al. Low-dose *in vivo* protection and neutralization across SARS-CoV-2 variants by monoclonal antibody combinations. *Nat Immunol* (2021) 22(12):1503–14. doi: 10.1038/s41590-021-01068-z
33. Wang Z, Muecksch F, Cho A, Gaebler C, Hoffmann HH, Ramos V, et al. Analysis of memory b cells identifies conserved neutralizing epitopes on the n-terminal domain of variant SARS-CoV-2 spike proteins. *Immunity* (2022) 55(6):998–1012.e8. doi: 10.1016/j.immuni.2022.04.003
34. Cao Y, Wang J, Jian F, Xiao T, Song W, Yisimayi A, et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* (2022) 602(7898):657–63. doi: 10.1038/s41586-021-04385-3
35. Kidd MJ, Chen Z, Wang Y, Jackson KJ, Zhang L, Boyd SD, et al. The inference of phased haplotypes for the immunoglobulin h chain V region gene loci by analysis of VDJ gene rearrangements. *J Immunol* (2012) 188(3):1333–40. doi: 10.4049/jimmunol.1102097
36. Kirik U, Greiff L, Levander F, Ohlin M. Data on haplotype-supported immunoglobulin germline gene inference. *Data Brief* (2017) 13:620–40. doi: 10.1016/j.dib.2017.06.031
37. Vázquez Bernat N, Corcoran M, Nowak I, Kaduk M, Castro Dopico X, Narang S, et al. Rhesus and cynomolgus macaque immunoglobulin heavy-chain genotyping yields comprehensive databases of germline VDJ alleles. *Immunity* (2021) 54(2):355–366.e4. doi: 10.1016/j.immuni.2020.12.018
38. Walter G, Tomlinson IM, Cook GP, Winter G, Rabbitts TH, Dear PH. HAPPY mapping of a YAC reveals alternative haplotypes in the human immunoglobulin VH locus. *Nucleic Acids Res* (1993) 21(19):4524–9. doi: 10.1093/nar/21.19.4524
39. Milner EC, Hufnagle WO, Glas AM, Suzuki I, Alexander C. Polymorphism and utilization of human VH genes. *Ann N Y Acad Sci* (1995) 764:50–61. doi: 10.1111/j.1749-6632.1995.tb55806.x
40. Cho CS, Wang X, Zhao Y, Carson DA, Chen PP. Genotyping by PCR-ELISA of a complex polymorphic region that contains one to four copies of six highly homologous human VH3 genes. *Proc Assoc Am Physicians* (1997) 109(6):558–64.
41. Chingé NO, Pramanik S, Hu G, Lin Y, Gao R, Shen L, et al. Determination of gene organization in the human IGHV region on single chromosomes. *Genes Immun* (2005) 6(3):186–93. doi: 10.1038/sj.gene.6364176
42. Pramanik S, Cui X, Wang HY, Chingé NO, Hu G, Shen L, et al. Segmental duplication as one of the driving forces underlying the diversity of the human immunoglobulin heavy chain variable gene region. *BMC Genomics* (2011) 12:78. doi: 10.1186/1471-2164-12-78
43. Sasso EH, Johnson T, Kipps TJ. Expression of the immunoglobulin VH gene 51p1 is proportional to its germline gene copy number. *J Clin Invest.* (1996) 97(9):2074–80. doi: 10.1172/JCI118644
44. Sasso EH, Ghillani P, Musset L, Piette JC, Cacoub P. Effect of 51p1-related gene copy number (V1-69 locus) on production of hepatitis c-associated cryoglobulins. *Clin Exp Immunol* (2001) 123(1):88–93. doi: 10.1046/j.1365-2249.2001.01406.x
45. Leggat DJ, Cohen KW, Willis JR, Fulp WJ, deCamp AC, Kalyuzhniy O, et al. Vaccination induces HIV broadly neutralizing antibody precursors in humans. *Science* (2022) 378(6623):eadd6502. doi: 10.1126/science.add6502