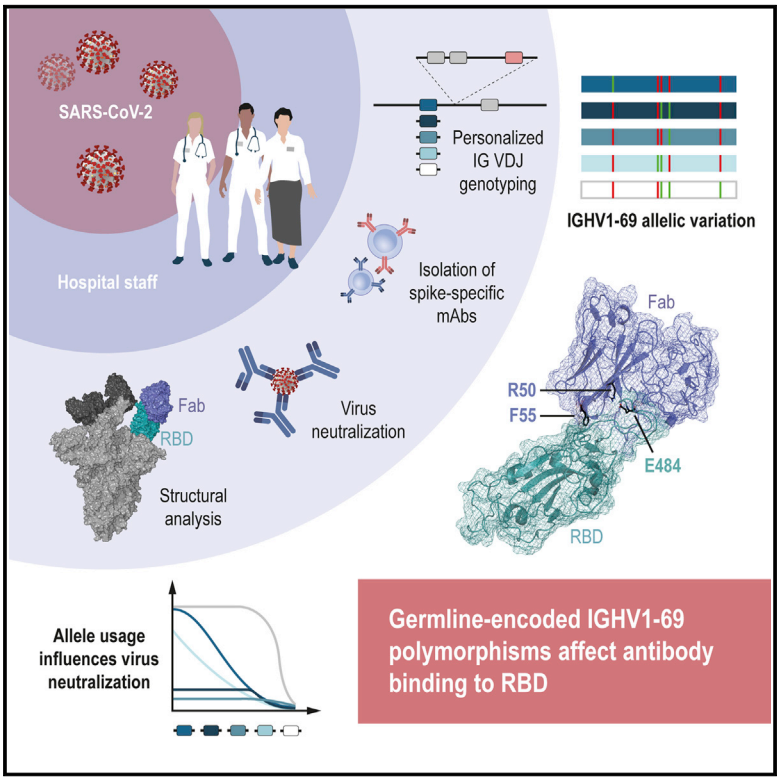
Article

Immunity

Immunoglobulin germline gene polymorphisms influence the function of SARS-CoV-2 neutralizing antibodies

Graphical abstract



Highlights

- IGH genes with high allelic diversity are used in anti-SARS-CoV-2 antibodies
- IGH allele usage can influence the activity of neutralizing antibodies
- Cryo-EM analysis confirms the role of germline-encoded residues in antigen binding
- IGH genotyping can uncover differences in Ab responses due to allelic variation

Authors

Pradeepa Pushparaj, Andrea Nicoletto, Daniel J. Sheward, ..., Ben Murrell, Martin Corcoran, Gunilla B. Karlsson Hedestam

Correspondence

gunilla.karlsson.hedestam@ki.se

In brief

The genes and alleles of the antigen receptor loci are highly variable between individuals, which may affect the quality of the immune response to different pathogens. Here, Pushparaj et al. use immunoglobulin genotyping and monoclonal antibody engineering to illustrate how heritable differences in such genes can modulate anti-SARS-CoV-2 antibody function.

Pushparaj et al., 2023, Immunity *56*, 193–206
January 10, 2023
https://doi.org/10.1016/j.immuni.2022.12.005

CellPress

Immunity

CellPress
OPEN ACCESS

Article

Immunoglobulin germline gene polymorphisms influence the function of SARS-CoV-2 neutralizing antibodies

Pradeepa Pushparaj,[1] Andrea Nicoletto,[1,4] Daniel J. Sheward,[1,4] Hrishikesh Das,[2] Xaquin Castro Dopico,[1] Laura Perez Vidakovics,[1] Leo Hanke,[1] Mark Chernyshev,[1] Sanjana Narang,[1] Sungyong Kim,[1] Julian Fischbach,[1] Simon Ekström,[3] Gerald McInerney,[1] B. Martin Hällberg,[2] Ben Murrell,[1] Martin Corcoran,[1] and Gunilla B. Karlsson Hedestam[1,5,*]

[1]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, 171 77 Stockholm, Sweden
[2]Department of Cell and Molecular Biology, Karolinska Institutet, 171 77 Stockholm, Sweden
[3]Department of Biomedical Engineering, Lund University, 221 84 Lund, Sweden
[4]These authors contributed equally
[5]Lead contact
*Correspondence: gunilla.karlsson.hedestam@ki.se
https://doi.org/10.1016/j.immuni.2022.12.005

SUMMARY

The human immunoglobulin heavy-chain (IGH) locus is exceptionally polymorphic, with high levels of allelic and structural variation. Thus, germline IGH genotypes are personal, which may influence responses to infection and vaccination. For an improved understanding of inter-individual differences in antibody responses, we isolated SARS-CoV-2 spike-specific monoclonal antibodies from convalescent health care workers, focusing on the IGHV1-69 gene, which has the highest level of allelic variation of all IGHV genes. The IGHV1-69*20-using CAB-I47 antibody and two similar antibodies isolated from an independent donor were critically dependent on allele usage. Neutralization was retained when reverting the V region to the germline IGHV1-69*20 allele but lost when reverting to other IGHV1-69 alleles. Structural data confirmed that two germline-encoded polymorphisms, R50 and F55, in the IGHV1-69 gene were required for high-affinity receptor-binding domain interaction. These results demonstrate that polymorphisms in IGH genes can influence the function of SARS-CoV-2 neutralizing antibodies.

INTRODUCTION

There is an increasing interest in understanding how common variation in immune-related genes influences our ability to control infections, not the least in the context of SARS-CoV-2. As for many other viral infections, inborn deficiencies in type I interferon responses are major risk factors for developing severe coronavirus disease 2019 (COVID-19)[2,3], while an archaic-derived isoform of oligoadenylate synthetase (OAS), an interferon-induced effector molecule, was shown to confer a protective effect against disease development.[4] These and many other findings add to a large body of work describing how variations in the innate immune system influence our response to viruses.[5] However, less is known about how variations in adaptive immune receptor genes influence our response to infections.[6] The genes that encode the antigen-binding portions of our B cell receptors (BCRs) and T cell receptors (TCRs), the variable (V), diversity (D), and junctional (J) genes, are among the most polymorphic in the human genome, although how this variation imparts functional consequences is largely unknown.

The most variable of the three loci that encode the IG genes is the immunoglobulin heavy-chain (IGH) locus, which is present at the telomeric end of chromosome 14. The heavy-chain (HC) locus is characterized by extensive structural variation involving frequent segmental deletions and duplications with particular regions, and the genes contained within, being more affected than others.[7] So far, the number of genomic assemblies, which span a complete human IGH locus, is limited.[8–10] However, recent studies using BCR repertoire sequencing (Rep-seq) and germline gene inference tools have revealed high inter-individual allelic and structural diversity within the region.[9,11,12] Importantly, IGHV allele usage determines the initial states of the HC complementarity determining regions 1 and 2 (HCDR1 and HCDR2) and contributes to the HCDR3, all of which shape epitope binding. The degree to which the presence or absence of specific IGH alleles influences the development of antibody responses to different pathogens remains largely unknown. To date, only a few examples of allele-specific responses have been described. These include a requirement for IGHV1-2*02 or IGHV1-2*04 for the generation of VRC01 class neutralizing antibodies directed against the CD4 binding site of HIV-1[13] and the propensity of specific IGHV1-69 alleles to be utilized in influenza hemagglutinin (HA) stem-directed neutralizing antibodies.[14,15]

Previous SARS-CoV-2 antibody studies demonstrated that certain IGHV genes are frequently used among S-specific neutralizing antibodies, including IGHV1-69, IGHV3-30, IGHV3-30-3, IGHV3-53, IGHV3-66, and IGHV5-51.[16–19] Of these, IGHV1-69 is of particular interest as it has the highest number of known allelic variants and is associated with structural variations resulting in frequent gene duplications.[9,10,15,20] Critically, IGHV1-69 is among the most highly utilized IGHV genes in the naive B cell repertoire.[12] Therefore, there is a high probability that B cells using this gene will encounter antigens after infection. However, because of the high allelic variation, antibody responses involving IGHV1-69 may differ considerably between individuals.

Importantly, studies describing S-specific monoclonal antibodies (mAbs) isolated from convalescent individuals early after exposure showed that few somatic hypermutations (SHMs) are sufficient to confer neutralizing activity against SARS-CoV-2.[21–23] Many of these antibodies target the receptor-binding domain (RBD) of the spike trimer, and research has also shown that RBD-binding B cells can be isolated from SARS-CoV-2 seronegative individuals.[24] If unmutated or modestly mutated antibodies that use IGHV genes with high allelic variation display neutralizing activity, it follows that this variation could impact the quality of both the initial and subsequent humoral response to the virus.

To address this question, we carried out personalized IG genotyping of previously infected healthcare workers to identify allelic and structural variations, particularly those affecting IGHV1-69. We first selected a case with high IGHV allelic content and a high frequency of circulating S-specific memory B cells from whom we isolated and characterized a set of SARS-CoV-2 S-specific mAbs (n = 29). Individualized genotyping enabled us to assign all cloned antibodies to the specific IGHV, IGHD, and IGHJ alleles present in this person. Five of the mAbs that we found to be potently neutralizing used the IGHV1-69 gene, of which three, CAB-I12, CAB-I47, and CAB-J39, were assigned to a recently described IGHV1-69*20 allele.[12] We investigated the impact of IGHV allele usage by reverting one of the most potent IGHV1-69*20 using antibodies, CAB-I47 to the IGHV1-69*20 germline sequence and to an additional set of five IGHV1-69 alleles that are common in the global human population. Assessment of the neutralizing activities of these antibody versions revealed that allele usage was critical for functional activity. Two additional IGHV1-69*20-using neutralizing antibodies, CAB-M77 and CAB-N86, which were isolated from an independent donor, showed very similar allele dependence. A high-resolution cryo-EM structure of the CAB-I47 Fab bound to the spike, revealed the molecular details of its interaction with the RBD. Overall, our results demonstrate that IGHV allele usage can profoundly influence the neutralizing antibody response after SARS-CoV-2 infection. Further studies aimed at defining how germline gene variation influences anti-viral B cell responses at individual level and population levels will improve our understanding of adaptive immunity.

RESULTS

Personalized IGH genotyping and haplotype analyses reveal a high degree of allelic and structural variation

A cohort of healthcare workers (n = 14), who tested positive for SARS-CoV-2 by RT-PCR in May 2020 was established to study the evolution of S-specific antibodies and memory B cell responses. Peripheral blood mononuclear cells (PBMCs) and serum samples were collected 7 months after primary infection and used for serological analyses, IG genotyping, and mAb isolation (Figure 1A). All study participants (SP01–SP14) were seropositive for S-IgG (Figure 1B), and donors displayed detectable neutralizing activity at this time point (Figure 1C).

To determine the IGHV genotypes of the participants, we produced libraries from the expressed IgM repertoire from the same time point, bulk sequenced these using the MiSeq platform, and analyzed them using our germline allele inference tool, IgDiscover.[11] The results revealed high IGHV diversity between individuals, both in terms of the number of alleles, ranging from 44 to 61 per person, and the content. The complete IGHV, IGHD, and IGHJ genotypes are shown in Table S1, demonstrating that the number of IGHV1-69 alleles varied greatly between the study participants. One individual was homozygous for IGHV1-69*01, while most study participants had two or three different alleles and two individuals had as many as four different IGHV1-69 alleles. Amino acid sequence alignment of all functional IGHV1-69 alleles shows that there are six variant positions between the alleles (Figure 1D). Altogether, six different alleles were found in the 14 study participants, IGHV1-69*01, IGHV1-69*02, IGHV1-69*04, IGHV1-69*06, IGHV1-69*09, and IGHV1-69*20 (Figure 1E).

To determine which alleles were present on each chromosome in the study participants, we performed inferred haplotype analysis.[25,26] This method takes advantage of the fact that VDJ recombination occurs locally on a given chromosome and heterozygous IGHJ or IGHD alleles as anchors to link IGHV alleles to either chromosome. Examining IGHD and IGHJ allele content, we found that SP01, SP02, SP06, SP08, and SP11 were heterozygous for IGHJ (IGHJ6*02/IGHJ6*03), 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) (Figures 1F and S1A). Use of these anchors to haplotype IGHV alleles demonstrated that the presence of gene duplications resulting in structural variation between the chromosomes (Figures 1F and S1A). Overall, the haplotype analysis of eight individuals in this cohort confirmed high allelic and structural variation in the IGHV1-69 region of the IGH locus.

To expand upon this in a larger cohort, we analyzed available IgM repertoire data from 32 unrelated cases not previously infected with SARS-CoV-2.[12] In this dataset, IGHV1-69 was the third most frequently used gene in the IgM repertoire (Figure S1B). Of the 32 individuals examined, nine could be haplotyped using heterozygous IGHJ6 as an anchor. This revealed that three of these cases had more than two alleles and one study participant had four IGHV1-69 alleles (Figure S1B). Together, these data demonstrated high allelic and structural variation in the IGHV1-69 region within the populations studied.

S-specific antibodies display preferential IGHV gene and allele usage and comprise highly potent neutralizers

We selected one of the participants from the SARS-CoV-2 convalescent cohort, SP14, who had three IGHV1-69 alleles (IGHV1-69*01, IGHV1-69*02, and IGHV1-69*20) (Figure 2A). The participant's kappa and lambda light-chain (LC) V germline alleles, IGKV and IGLV, were also determined (Figure 2A). Single

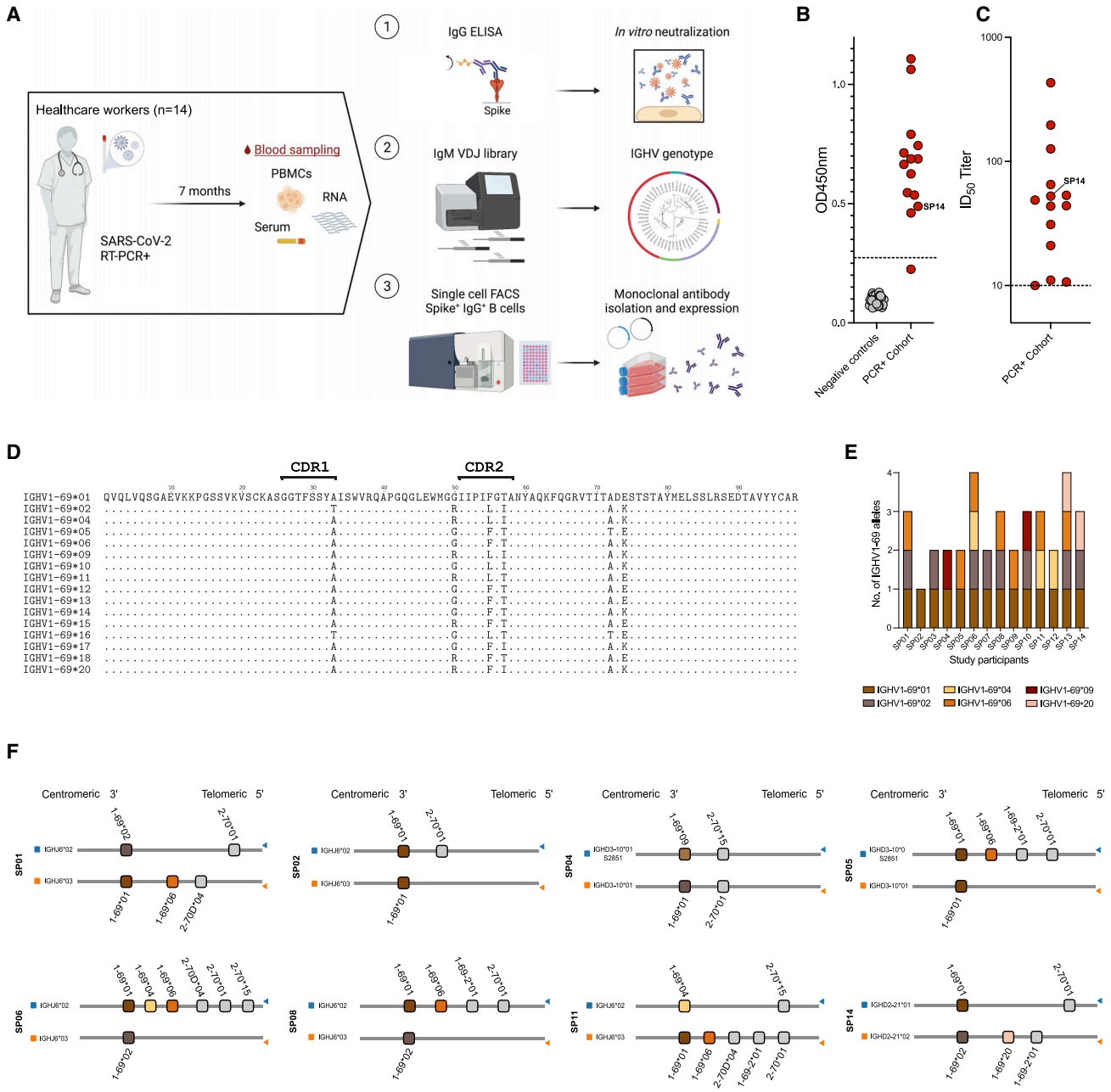


Figure 1. Study design and demonstration of high IGHV1-69 diversity among the donors

(A) Schematic representation of the sampling of 14 SARS-CoV-2 RT-PCR+ study participants followed by IGH genotyping and memory B cell sorting 7 months after the infection.

(B) Spike-specific IgG in serum samples from the 14 study participants.

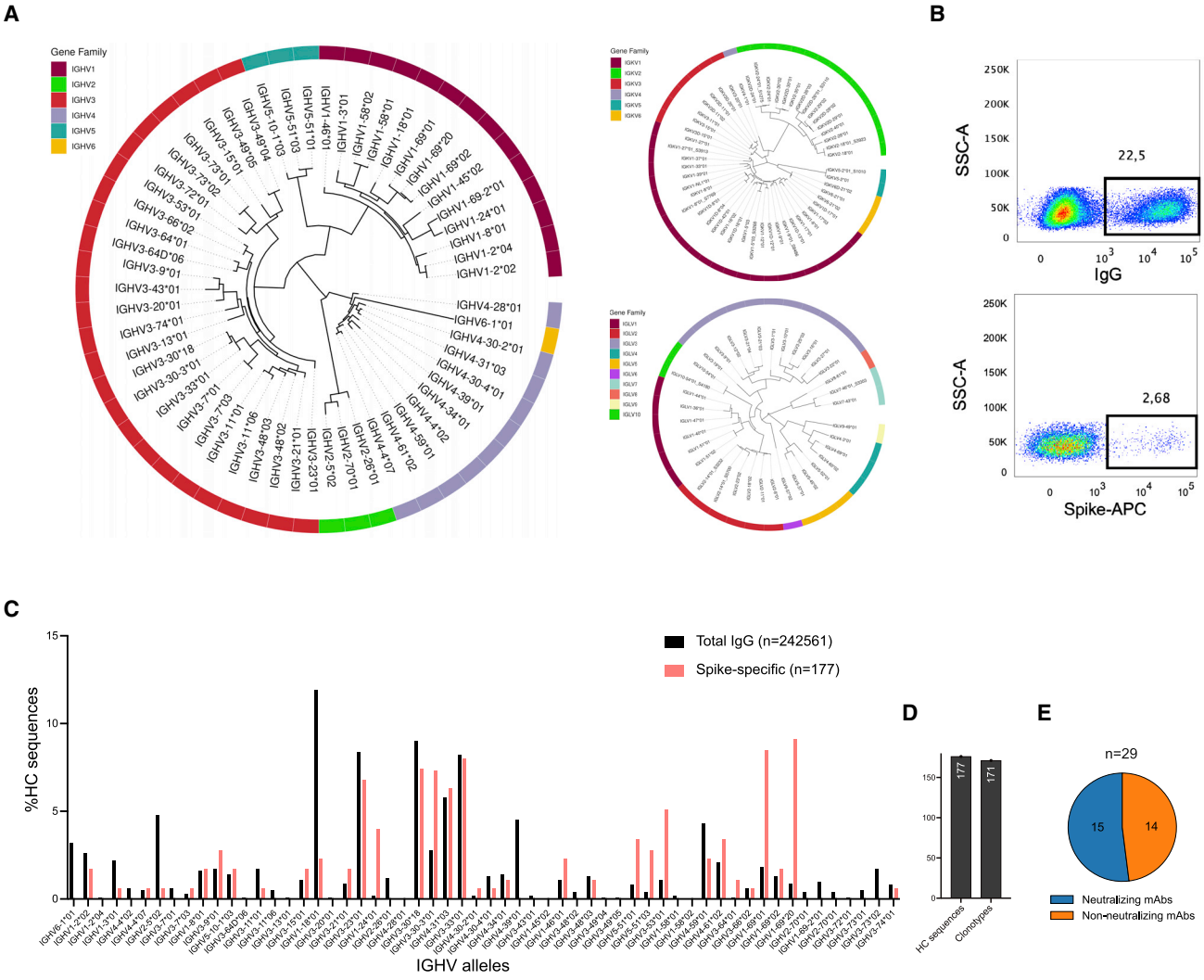
(C) Serum ID_{50} neutralization values against the ancestral virus in the serum samples.

(D) Amino acid sequence alignment of functional human IGHV1-69 alleles.

(E) A summary plot showing which IGHV1-69 alleles were present in each study participant.

(F) Schematic illustration of IGHV1-69 alleles present on each chromosome of eight study participants. Neutralization measurements were repeated twice. See also Figure S1.

SARS-CoV-2 S-specific IgG^+ memory B cells were sorted into 96-well plates for antibody sequence analysis and mAb isolation. Of the $CD20^+CD27^+IgG^+$ memory B cells, 2.7% were S specific (Figures 2B and S2A). Amplification of antibody V(D)J regions from the 247 sorted cells was performed by single-cell RT-PCR using a previously published comprehensive primer set for human HC and LC isolation.[27] Sanger sequencing identified 177 full-length HC sequences and 146 paired HC and LC sequences, a subset of which were cloned into expression vectors[28] to produce mAbs.



F

ID	VH gene	DH gene	JH gene	%VH SHM		HCDR3	VL gene	JL gene	%VL SHM		LCDR3	Spike	RBD	IC50 (μg/ml)
				nt	aa				nt	aa				
CAB-A17	IGHV3-53*01	IGHD3-10*01	IGHJ4*02	8,9	12,4	ARDYGDFYFDY	IGKV3-20*01	IGKJ1*01	3,8	7,4	QQYGSSPRT	+	+	0.01
CAB-A18	IGHV3-53*01	IGHD3-10*01	IGHJ4*02	4,4	10,3	ARDFGEFYFDY	IGKV3-20*01	IGKJ1*01	5,5	8,3	QQYGSSPRT	+	+	0.01
CAB-A49	IGHV3-53*01	IGHD3-10*01	IGHJ4*02	4,4	6,2	ARDFGEFYFDF	IGKV3-20*01	IGKJ1*01	2,8	5,2	QQYGSSPRT	+	+	0.01
CAB-B37	IGHV3-53*01	IGHD3-9*01	IGHJ6*02	6,9	13,5	ARLPYGTDV	IGKV1-33*01	IGKJ5*01	3,9	7,4	QQYDNLPIT	+	+	0.01
CAB-C19	IGHV3-53*01	IGHD6-19*01	IGHJ3*02	4,1	13,4	ARVLPVAGDSFDF	IGKV1-33*01	IGKJ1*01	6,4	13,8	QHYNNFLVTWT	+	+	0.01
CAB-D16	IGHV3-53*01	IGHD3-9*01	IGHJ6*02	9,2	11,3	AREVYALDV	IGKV1-9*01	IGKJ4*01	5,6	13,4	QRVNSHPPA	+	+	0.01
CAB-E24	IGHV3-53*01	IGHD2-8*01	IGHJ3*02	6,1	12,4	ARDRGNGAFDI	IGKV1-9*01	IGKJ2*01	1,8	2,1	QQLNSYPGDT	+	+	0.01
CAB-E30	IGHV3-53*01	IGHD3-10*01	IGHJ3*02	6,8	13,4	ARDRGNAAFDF	IGKV1-9*01	IGKJ2*01	3,5	7,4	QQLDTYPGDT	+	+	0.08
CAB-F52	IGHV3-30-3*01	IGHD2-15*01	IGHJ6*02	3,4	6,1	ARDLVVVVAASSHYYNGMDV	IGKV1D-12*01	IGKJ5*01	3,5	5,3	QQVKSFPLT	+	+	0.005
CAB-G01	IGHV1-69*02	IGHD5-12*01	IGHJ3*02	6,1	9,2	ARERGRSGYGSAWAFDL	IGLV1-40*01	IGLJ2*01	2,7	6,1	QSYDSSLTGVV	+	+	0.01
CAB-H50	IGHV1-69*02	IGHD6-6*01	IGHJ6*02	5,7	7,1	ARDPRPLIPPASQISHGMDV	IGKV2-28*01	IGKJ5*01	2	3,7	MQALQIPPAT	+	+	0.1
CAB-I12	IGHV1-69*20	IGHD3-22*01	IGHJ4*02	4,1	8,2	ARANYYYESGGYHFDN	IGKV3-20*01	IGKJ1*01	2,8	5,2	QQYHNSPRT	+	+	0.01
CAB-I47	IGHV1-69*20	IGHD3-22*01	IGHJ4*02	4,4	7,1	ARSAYYSESSGYYLDY	IGKV3-20*01	IGKJ2*01	2,4	4,2	QQYVNSPRT	+	+	0.01
CAB-J39	IGHV1-69*20	IGHD3-22*01	IGHJ4*02	1,7	5,1	ARPPSPYYYDRSGHPQADHYFDY	IGLV1-40*01	IGLJ2*01	1,3	3	QSYDSSLSGSV	+	+	0.5
CAB-K04	IGHV5-51*01	IGHD6-13*01	IGHJ4*02	6,8	16,3	ARSNTGYTSTWFRD	IGLV3-21*04	IGLJ2*01	5,9	7,3	QVWDSSIDHFVV	+	+	0.03

Figure 2. Spike-specific memory B cell sorting yields several IGHV1-69*20-using neutralizing mAbs

(A) Dendrogram showing IGHV, IGKV, and IGLV alleles present in SP14.

(B) Panels showing the gating strategy used to sort single S-specific $CD20^{+}CD27^{+}IgG^{+}$ B cells.

(C) Comparison of IGHV allele usage in the total and spike-specific IgG repertoires of SP14.

(D) Number of HC sequences from sorted B cells and clonal lineages derived from the total HC sequences.

(E) Pie chart showing the proportion of neutralizing (blue) and non-neutralizing (orange) antibodies among the 29 isolated mAbs.

(F) Genetic and functional properties of the 15 neutralizing mAbs isolated from SP14. Neutralization measurements were repeated twice.

See also Figure S2.

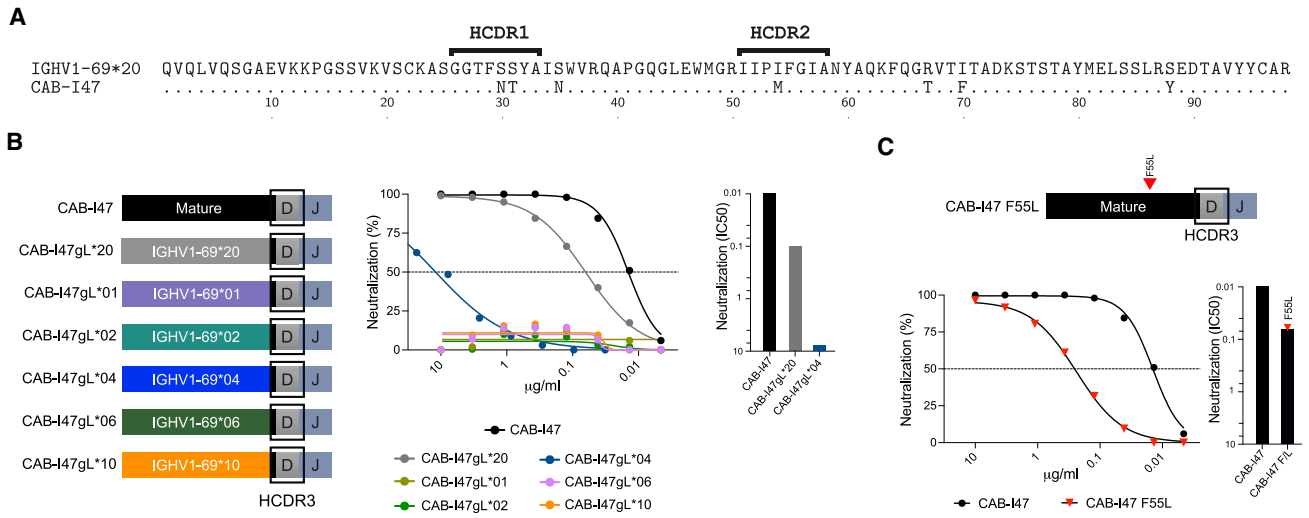


Figure 3. IGHV1-69 allele usage influences CAB-I47 neutralizing activity

(A) Amino acid sequence alignment of the IGHV1-69*20 germline allele with the V gene region of CAB-I47.

(B) The design of CAB-I47 germline-reverted HC sequences and swaps to the IGHV1-69 germline alleles *20, *01, *02, *04, *06, or *10, which were paired with the mature CAB-I47 LC for functional testing (left). Neutralizing activity of the variant mAbs against the ancestral SARS-CoV-2 strain is shown as curves (middle) and IC50 values (μg/mL) (right).

(C) Design of the mature CAB-I47 with the HCDR2 F55L substitution, paired with the CAB-I47 mature LC (top panel). Neutralization activity of CAB-I47 and CAB-I47 F55L against the ancestral SARS-CoV-2 strain (bottom left) and IC50 values (μg/mL) (bottom right). Neutralization measurements were repeated twice. See also Figure S3.

Examination of the IGHV allele usage of the 177 spike-specific HCs and comparisons with the study participant's total IgG repertoire, obtained by bulk PBMC RNA sequencing, revealed that several IGHV genes were over-represented in the S-specific B cells, including IGHV1-69, IGHV3-30-3, IGHV3-53, and IGHV5-51, consistent with previous reports[19,29–34] (Figure 2C). The S-specific response was highly polyclonal, with 171 unique clones of the 177 total HC sequences analyzed (Figure 2D) using the same IGHV allele, IGHJ allele, length of HCDR3, and 80% aa HCDR3 identity as a definition for clonal relatedness.[35] From the 146 paired sequences, 29 spike-specific mAbs were expressed, of which 52% (n = 15) were neutralizing (Figure 2E). All neutralizing mAbs bound the RBD (Figure S2B). The neutralization curves and IC_{50} values against the ancestral SARS-CoV-2 strain are shown in Figure S2C. A summary of the genetic and functional properties of all neutralizing mAbs is shown in Figure 2F and of the non-neutralizing mAbs in Table S2.

Knowledge of the specific V, D, and J alleles present in the study participant allowed for precise genetic characterization of the isolated neutralizing antibodies. Several mAbs used IGHV3-53 and had short HCDR3s, suggesting that they were of the previously described class 1 group of SARS-CoV-2-specific antibodies.[32,36] The IGHV3-53*01-using mAbs isolated here, including three clonally related mAbs, CAB-A17, CAB-A18, and CAB-A49, are separately described.[37] One ultrapotent neutralizing mAb, CAB-F52, used IGHV3-30-3*01, a gene that is present in only a subset of individuals (Table S1). Several of the most potently neutralizing mAbs used IGHV1-69, either IGHV1-69*02 or the recently identified IGHV1-69*20 allele.[12] Two of the IGHV1-69*20-using mAbs, CAB-I12 and CAB-I47, were clonally related (Figure 2F) and narrow in their neutralization profile against different SARS-CoV-2 variants (Figure S2D). When inspecting the sequence alignment of the IGHV1-69 alleles, we found that IGHV1-69*20 was closest to IGHV1-69*04 with a single-nucleotide difference resulting in a leucine (L) in *04 to phenylalanine (F) substitution at position 55 in the HCDR2 of *20. Several other IGHV1-69 alleles, *01, *05, *06, *12, *13, *15, and *17, also have an F in this position but differ from IGHV1-69*20 at other variant positions (Figure 1D). The average SHM for the neutralizing mAbs was 5.5% in the nucleotide sequence and 10.1% in the aa sequence, reflecting the time point at which they were isolated, 7 months after primary infection. None of the mAbs used previously unreported LC alleles. These results demonstrate that potent neutralizing mAbs were elicited in SP14, several of which used IGHV1-69.

IGHV1-69 allele usage influences neutralizing antibody activity

To examine the role of IGHV1-69 allelic variation, we selected an IGHV1-69*20-using mAb, CAB-I47, for in-depth analyses. The IGHV gene of CAB-I47 was modified by SHM at 7 aa positions, 3 of which were in the HCDRs (Figure 3A). To determine the contribution of SHM toward the neutralization potency of CAB-I47, we reverted the V gene region to the IGHV1-69*20 germline configuration. The HCDR3 region was kept intact and there was no SHM in the J gene (Figure 3B, left). The V gene germline-reverted (gL) version of CAB-I47, designated CAB-I47gL*20, paired with the mature CAB-I47 LC, neutralized the ancestral SARS-CoV-2 spike, albeit approximately 6-fold less potently than the corresponding mature antibodies (Figure 3B, right). We also reverted the V gene region of CAB-I47 to IGHV1-69*01, IGHV1-69*02, IGHV1-69*04, IGHV1-69*06, and IGHV1-69*10, which are common in the population and have coding differences. The germline-reverted HCs were co-expressed

with the mature CAB-I47 LC and designated CAB-I47gL*01, CAB-I47gL*02, CAB-I47gL*04, CAB-I47gL*06, and CAB-I47gL*10 (Figure 3B, left) and had their neutralizing capacity assessed. All germline-reverted allele-swapped mAbs lost neutralizing activity, except CAB-I47gL*04, which displayed substantially reduced but detectable neutralization (Figure 3B, right).

To further understand this outcome, we investigated the role of F55 in the context of the mature CAB-I47 mAb. F55 in CAB-I47 was substituted with an L and the F55L HC was co-expressed with the mature LC of CAB-I47 to produce CAB-I47 F55L (Figure 3C). The neutralization potency of CAB-I47 F55L was reduced more than 10-fold compared with CAB-I47, demonstrating the importance of the phenylalanine at this position for virus neutralizing activity. We also determined the binding kinetics and affinity constants for CAB-I47, CAB-I47gL*20, CAB-I47gL*04, and CAB-I47 F55L Fabs by surface plasmon resonance. The results revealed that the affinity constant was markedly lower for the mature CAB-I47 Fab compared with CAB-I47gL*20 (0.23 nM compared with 17 nM), illustrating the importance of SHM for affinity maturation. Further, the CAB-I47 F55L Fab displayed a similar on rate to CAB-I47, but a very fast off rate (KD of 18.7 nM), highlighting the importance of F55 for rapid association. The CAB-I47gL*04 Fab displayed a very-low affinity to RBD with a KD of 505 nM, reiterating the impact of allelic variation on antibody binding and function. The KD measurements correlated well with the IC_{50} values from the neutralization assay (Figures S3A–S3C). Together, these results demonstrate that germline-encoded polymorphisms in IGHV1-69 can influence the function of SARS-CoV-2 neutralizing antibodies.

IGHV1-69*20-using mAbs from an independent donor display similar characteristics

To investigate whether similar antibodies could be isolated from an unrelated donor, we used samples from study participant SP13, the only other donor in our cohort who had the IGHV1-69*20 allele (Figure 4A; Table S1). PBMC and serum from SP13, who was infected in the spring of 2020 and subsequently vaccinated, were available in January 2022. As expected, we observed elevated serum neutralizing activity at this time compared with June 2020 (Figure 4B). Spike-specific memory B cells were sorted into 96-well plates using dual fluorescently labeled spike probes yielding 526 HC sequences (Figures 4C and S4A). Analysis of the IGHV allele usage of the spike-specific HCs compared with the study participant's total IgG repertoire, obtained by bulk PBMC Rep-seq sequencing, revealed a similar over-representation of IGHV genes in the S-specific B cells as observed in SP14, with frequent usage of IGHV1-69, IGHV3-30-3, and IGHV5-51 (Figure 4D). Of the IGHV1-69-using HC sequences, the most frequently used allele of the donor's four alleles was IGHV1-69*20, with 28 of 66 HCs (Figure 4E). Of the IGHV1-69*20-using HCs that also had a productive LC sequence, we selected pairs that matched CAB-I47 regarding LC V gene usage (IGKV3-20*01) and HCDR3 length (16 aa). Of four such pairs evaluated, we selected two that were sensitive to the F55L change in the HCDR2, CAB-M77, and CAB-N86. Both mAbs bound the RBD and exhibited neutralizing activity against the ancestral SARS-CoV-2 (D614G) variant. The genetic and functional characteristics of CAB-M77 and CAB-N86 are summarized in Figure 4F, and the F55L sensitivity is shown in Figure 4G.

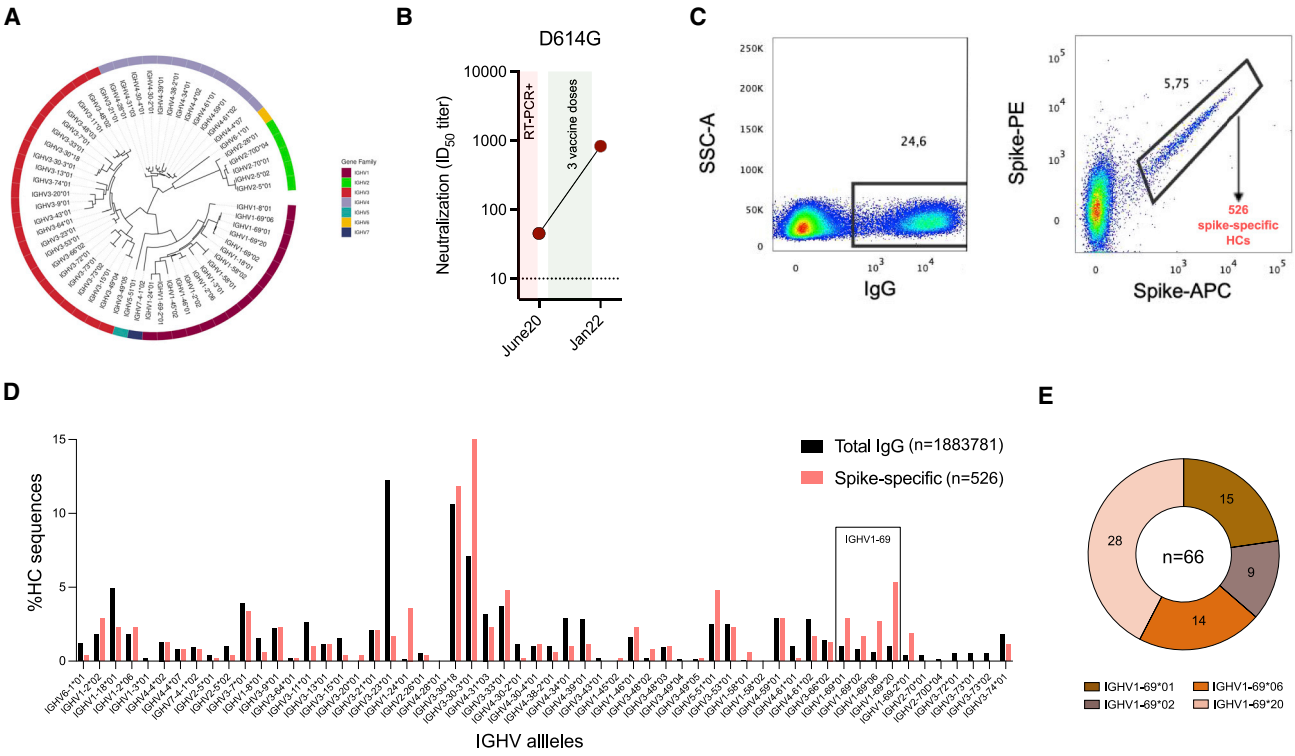
We next generated versions of CAB-M77 and CAB-N86 where the HC V gene region was reverted to the naturally used allele, IGHV1-69*20, or to IGHV1-69*01, IGHV1-69*02, IGHV1-69*04, IGHV1-69*06, and IGHV1-69*10. The germline-reverted HCs were co-expressed with the mature LC of the respective antibody and evaluated for their neutralizing activities. We found that CAB-M77 was only functional in the context of its natural germline allele with no detectable neutralizing activity observed for the allele swaps. CAB-N86 behaved very similarly to CAB-I47 in that the IGHV1-69*04 allele retained some neutralizing activity, but none of the other alleles yielded antibodies with measurable neutralization (Figure 4H). Both CAB-I47 and CAB-M77 also showed a greater reliance on their HC SHM than LC SHM as shown by testing the neutralizing activities of either or both HC and LC V gene germline-reverted versions of these mAbs compared with the mature Abs (Figures S4D–S4F).

To evaluate the target specificities of CAB-I47 and CAB-M77, we used hydrogen-deuterium exchange mass spectrometry (HDX-MS) to map the binding to the ancestral RBD. The main interaction site for both mAbs was mapped to aa 470–490 (Figure S4B), overlapping the site targeted by class 1 antibodies.[36] To further evaluate CAB-I47, CAB-M77, and CAB-N86, we tested their capacity to neutralize a D614G virus encoding the RBD E484K mutation (Zeta), and we found that all three mAbs were sensitive to this mutation (Figure S4C). Thus, isolation of spike-specific antibodies from an independent donor yielded IGHV1-69*20-using neutralizing mAbs that had similar characteristics to CAB-I47.

Structural analysis reveals the basis for IGHV allele requirement

To understand the structural basis of CAB-I47 binding, we determined a cryo-EM structure of the Fab bound to the ancestral SARS-CoV-2 spike protein. Extensive classification of picked particles resulted in two different states, a "1-up" state with three Fabs bound (Figure 5A) and a "2-up" state with three Fabs bound to each RBDs (Figure S5A). The structures demonstrated that the Fab can bind the RBD both in the "up" and "down" conformation. The overall resolution of the cryo-EM reconstructions of 1-up and 2-up was 2.5 Å (FSC 0.143) and 2.6 Å (FSC 0.143), respectively. However, due to the flexibility of the RBD-Fab part, the resolution of this region was not sufficient to analyze molecular details. Therefore, we performed localized reconstruction and refinement of the RBD-Fab (variable heavy [VH] and variable light [VL]) part, resulting in a higher resolution map of overall 2.5Å (FSC 0.143) (Figure S5B), which helped us dissect the atomic details of the Fab-RBD interaction.

The structure revealed that the CAB-I47 Fab binds on the upper part of the RBD and belongs to the class 2 of RBD-targeting antibodies. The CAB-I47 Fab and ACE2 share binding surfaces on the RBD, therefore CAB-I47 blocks ACE2 binding due to steric conflicts. The LC is also involved, with interaction interface areas of VH and VL being ~630 and 276 $Å^2$, respectively. VH residue R50 plays a vital role in the interaction as it forms several hydrogen bonds with E484 and G482 and one salt bridge with RBD E484 (Figure 5B). The Fab VL residues N94 and R97 also interact with RBD, VL N94 forms hydrogen bonds with RBD N481 and V483, and VL R97 forms a hydrogen



F

ID	VH gene	DH gene	JH gene	%VH SHM (nt)	HCDR3	VL gene	JL gene	%VL SHM (nt)	LCDR3	Spike	RBD	IC50 (μg/ml)
CAB-M77	IGHV1-69*20	IGHD3-22*01	IGHJ4*02	4.4	ARAEFYSDSSGHHFDY	IGKV3-20*01	IGKJ1*01	4.5	HQYANSPRT	+	+	0.01
CAB-N86	IGHV1-69*20	IGHD3-22*01	IGHJ3*02	5.5	ARSNFYSDTSGYYLDI	IGKV3-20*01	IGKJ1*01	2.4	QQYDSSPRT	+	+	0.01

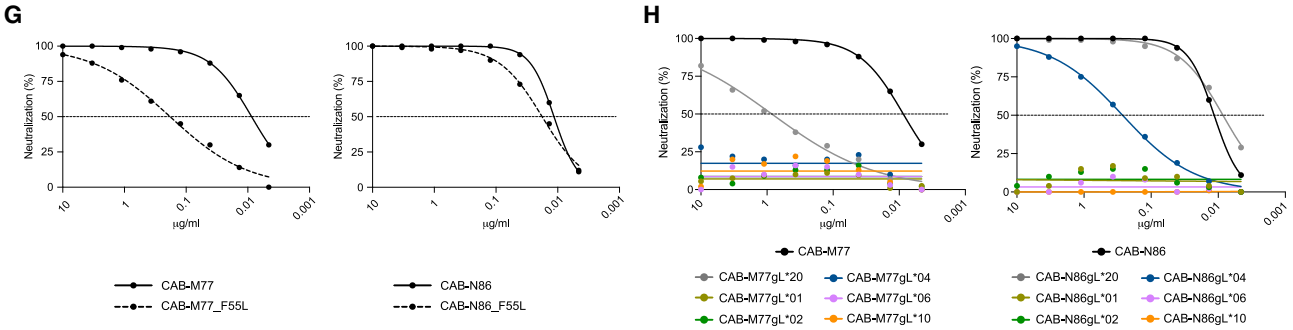


Figure 4. IGHV1-69*20-using neutralizing mAbs isolated in an independent donor

(A) Dendrogram showing IGHV alleles present in SP13.

(B) Serum ID_{50} neutralization values against ancestral SARS-CoV-2 at pre- and post-vaccination time points.

(C) Panels showing the gating strategy used to isolate single spike-specific $CD20^{+}CD27^{+}$ IgG B cells.

(D) Comparison of IGHV allele usage in the total and spike-specific IgG repertoire of SP13.

(E) Pie chart showing the number of HCs of the total IGHV1-69-using HCs using a given IGHV1-69 allele from the sorted B cells.

(F) Genetic and functional properties of the selected neutralizing mAbs.

(G) Neutralization of the ancestral SARS-CoV-2 strain by the mature CAB-M77 and CAB-N86 mAbs with and without the F55L mutation.

(H) Neutralizing activity by CAB-M77 and CAB-N86 with their HC V gene regions reverted to the IGHV1-69 germline alleles *20, *01, *02, *04, *06, or *10, tested against the ancestral SARS-CoV-2 strain. Neutralization measurements were repeated twice.

See also Figure S4.

bond and a salt bridge with RBD E484. Also, the SHM-acquired VH residue N30 forms a hydrogen bond with RBD residue N450 (Figure 5B), and Y101 as well as S103 (HCDR3-encoded) interact with F490 via backbone hydrogen bonding. Furthermore, VH F55 docks into a hydrophobic pocket formed by RBD residues L452, I468, F490, and I492 (Figure 5C).

The finding that CAB-I47 R50 interacted with E484 and G482 on the RBD led us to generate versions of this Ab carrying a

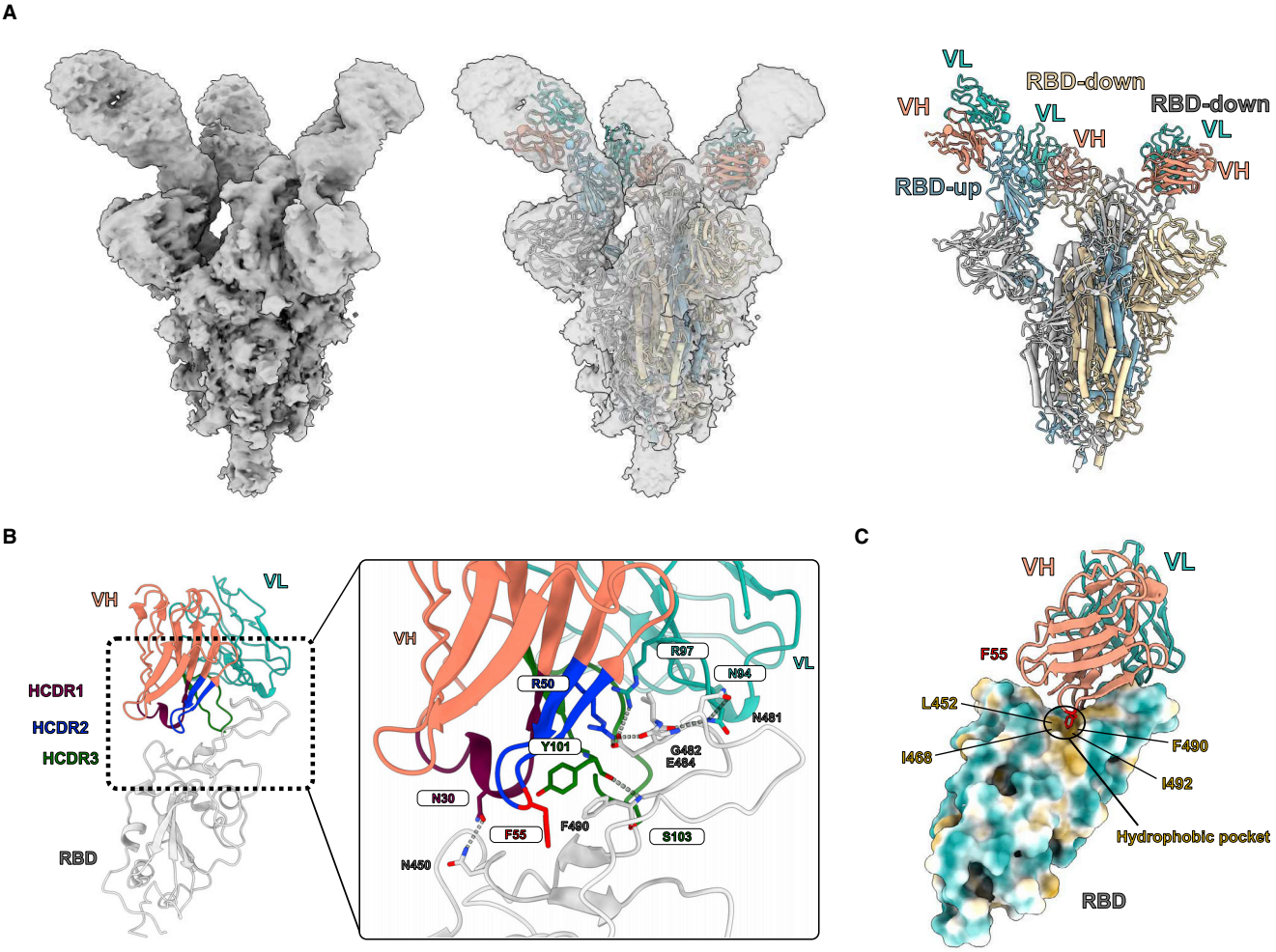


Figure 5. Cryo-EM analysis of the spike-CAB-I47 Fab complex reveals its binding mode

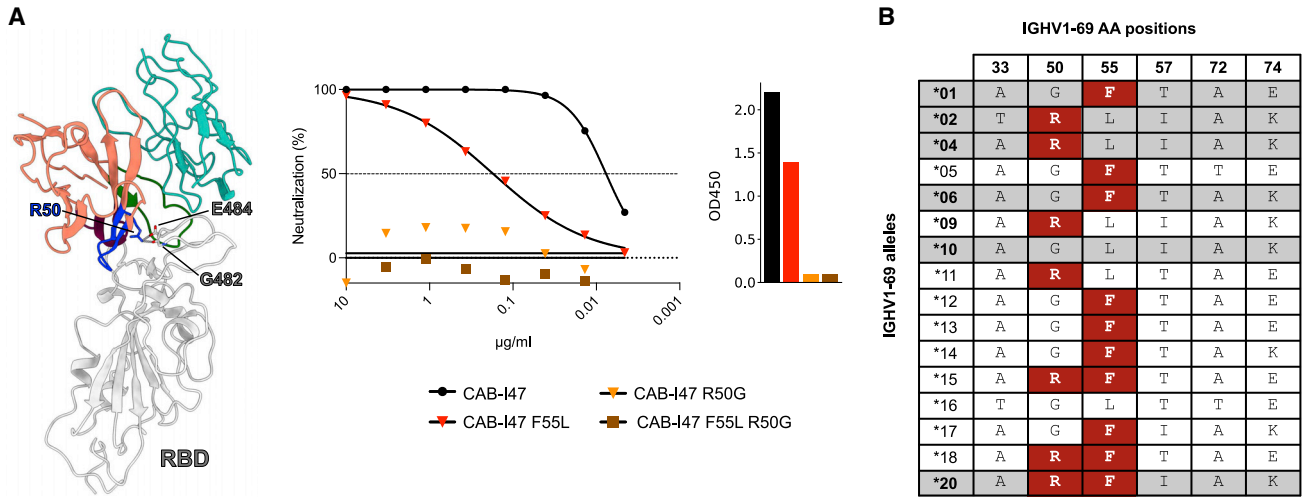
(A) Cryo-EM reconstruction of the spike-CAB-I47 Fab complex in 1-up conformation, three CAB-I47 Fabs occupy the RBDs (left). The cryo-EM map (gray transparent) overlayed with atomic model, trimeric spike with three the CAB-I47 Fab variable domains (middle). Atomic model of RBD- the CAB-I47 Fab (variable domains) with the HCDR1, HCDR2, and HCDR3 color-coded (right).

(B) Molecular details of the CAB-I47 Fab and RBD interaction with residues important for the interaction labeled.

(C) The IGHV1-69*20 F55 residue (shown in red) located in a hydrophobic pocket of the RBD. The RBD is shown in surface representation and colored according to relative hydrophobicity with various RBD residues in the binding area are labeled.

See also Figure S5.

R50G mutation alone or in combination with F55L (position 50 varies between an R and a G in different IGHV1-69 alleles). The F55L mutation reduced neutralizing potency of CAB-I47 by over 10-fold and both R50G and the combined R50G and F55L completely abolished both neutralizing and binding activity (Figure 6A). These results demonstrate that R50 is critical for the interaction with the RBD for this class of antibodies while a leucine in position 55 is partly tolerated but results in reduced activity. When inspecting the variant positions in the known functional IGHV1-69 alleles, we found that other than IGHV1-69*20, which was used in the antibodies studied here, only IGHV1-69*15 and IGHV1-69*18 encode both R50 and F55, but they differ from IGHV1-69*20 at position 57 (Figure 6B). Of note, IGHV1-69*15 and IGHV1-69*18, which have identical coding sequences but differ at the nucleotide level, were not found in the 14 participants studied here (Table S1), nor in the 32 individuals from Gidoni et al.,[12] suggesting that they are rare in the represented population groups, while IGHV1-69*20 is more common. To investigate whether IGHV1-69*20 is used in other SARS-CoV-2 RBD-targeting neutralizing antibodies, we queried CoV-Ab-Dab (http://opig.stats.ox.ac.uk/webapps/covabdab/). When examining antibodies for which nucleotide sequences were available, we identified several mAbs for which the closest assignment was the IGHV1-69*20 allele (Figure S6A). These included CV1206, CV1182,[38] and BG1-24;[39] three antibodies described to be sensitive to the E484K escape mutation, similar to the mAbs described in this study. For neutralizing IGHV1-69-using antibodies in CoV-Ab-Dab for which only the amino acid sequence was available, we identified BD57-005, BD56-031 and XGv-232,[40] REGN10977,[41] C14646P3S,[42] and C1210 and



	33	50	55	57	72	74
*01	A	G	F	T	A	E
*02	T	R	L	I	A	K
*04	A	R	L	I	A	K
*05	A	G	F	T	T	E
*06	A	G	F	T	A	K
*09	A	R	L	I	A	K
*10	A	G	L	I	A	K
*11	A	R	L	T	A	E
*12	A	G	F	T	A	E
*13	A	G	F	T	A	E
*14	A	G	F	T	A	K
*15	A	R	F	T	A	E
*16	T	G	L	T	T	E
*17	A	G	F	I	A	K
*18	A	R	F	T	A	E
*20	A	R	F	I	A	K

Figure 6. Substitution of the germline-encoded R50 residue abolishes the activity of CAB-I47

(A) Cartoon and stick representation of the IGHV1-69*20 R50 residue interactions with G482 and E484 on the RBD, color coding as in Figure 5 (left). Neutralization of the ancestral SARS-CoV-2 strain by the mature CAB-I47 and versions containing the F55L mutation, the R50G mutation or both, and ELISA binding by the same antibodies against RBD (right).

(B) Summary of variant amino acid positions in different IGHV1-69 alleles with the presence of R50 and F55 highlighted in red. Gray background indicates alleles that were evaluated in this study. Neutralization measurements were repeated twice.

See also Figure S6.

C1211,[43] which could be assigned to IGHV1-69*20 due to the presence of R50, F55, I57, A72, and K74, a combination of residues only found in IGHV1-69*20 (Figures 1D, 6B, and S6B).

Overall, the results presented here suggest that elicitation of IGHV1-69*20-using SARS-CoV-2 neutralizing antibodies is not infrequent. Furthermore, CAB-I47, CAB-M77, CAB-N86, and the engineered versions of these antibodies shown here offer clear examples of how differences in IGH allelic content between individuals can influence our response to SARS-CoV-2.

DISCUSSION

Heritable variation within components of the immune system has widely been identified as critical for the genetic arms race between pathogens and hosts.[44,45] Anti-viral antibody epitope selection was recently reported to be a heritable trait, controlled in part by the MHC class II locus.[46] However, how germline-encoded variations at the IG loci, especially coding changes in V, D, and J genes, influence our response to viruses remains poorly characterized. Such studies have been hampered by the challenges faced in genotyping these complex loci as previously discussed.[6,7,47] Studies have demonstrated a very high degree of inter-individual diversity across the IG loci, including both single-nucleotide polymorphisms and large structural variations.[9,11,12,48]

Despite these challenges, a set of previous observations support further investigations of this area. For example, a polymorphism in the recombination signal sequence (RSS) of the IGKV2D-29 LC gene was shown to be associated with increased susceptibility to *Haemophilus influenzae* type B infection in a Navajo population,[49] resulting from the pauciclonal utilization of this allele in the antibody response to infection and vaccination. Further examples were provided by studies showing that IGHV1-69 allele usage is important for the generation of influenza HA stem-directed antibodies[14] and that only certain IGHV1-2 alleles, which encode a particular germline-encoded aa motif, can generate VRC01-class broadly neutralizing antibodies against HIV-1.[13]

Here, we investigated the role of IGH germline gene variation in the antibody response against SARS-CoV-2 by focusing on IGHV1-69, which is characterized by high allelic and structural diversity and is frequently utilized in SARS-CoV-2 S-specific antibody responses. Personalized IG genotyping and haplotype analysis allowed us to define which IGHV alleles were present on each diploid IGHV locus of the study participants, demonstrating that many individuals had more than one IGHV1-69 copy on each chromosome. The presence of more than two IGHV1-69 alleles in a single individual was previously defined in the context of anti-influenza antibodies. In those studies, IGHV1-69 allele usage was shown to be important for targeting the HA stem region, a broadly neutralizing epitope, and gene copy number varied between individuals and population groups.[15]

IGHV1-69 locates in a region at the IGH locus that has been duplicated during human evolution[10] resulting in the presence of an additional gene, IGHV1-69D, in some individuals. Our results confirm that IGHV1-69 displays frequent copy-number variation. For example, in SP06, we observed three separate copies of both IGHV1-69 and the closely located IGHV2-70 on the same haplotype. The IGHV1-69 and IGHV2-70 genes are separated by less than 10 kb. The presence of two additional copies of both genes demonstrates copy-number variation of a segment that contains both genes, resulting in two additional high-frequency functional IGHV1-69 alleles on this haplotype. A total of six different IGHV1-69 alleles were expressed in the fourteen individuals, with the number of alleles per individual varying between one and four, resulting in 12 different IGHV1-69 genotypes among the 14 study participants. Thus, the

complexity of IGHV1-69 genotypes is more significant than often appreciated. A recent study using VH-restricted knockin mice expressing either the human IGHV1-69*01 (F54) allele or IGHV1-69*09 (L54) allele suggested that the L allele predisposes toward poly- and autoreactivity.[50] How this finding translates to human B cell responses given the high diversity of IGHV1-69 genotypes will be an exciting but challenging question to address. It follows that detailed genotyping of the IGH locus in populations will reveal important insights into the pathology of infectious and autoimmune diseases.

Here, we sorted SARS-CoV-2 spike-specific B cells and isolated a first set of mAbs from SP14 who had three IGHV1-69 alleles: IGHV1-69*01, IGHV1-69*02, and IGHV1-69*20. Several of the mAbs used IGHV1-69*20 including CAB-I47, which we characterized in detail. When evaluating the role of IGHV allele usage by performing germline allele swaps of CAB-I47, we found profound differences in affinity and neutralizing ability depending on which IGHV1-69 allele was used, demonstrating that germline-encoded differences between the alleles made critical contributions to antibody function. The second donor we isolated mAbs from, SP13, had four IGHV1-69 alleles: IGHV1-69*01, IGHV1-69*02, IGHV1-69*06, and IGHV1-69*20. We isolated two mAbs, CAB-M77 and CAB-N86, which displayed similar genetic and functional properties to CAB-I47. These mAbs also required IGHV1-69*20 for their neutralizing capacity, supporting that allele-specific germline-encoded amino acids were required. In agreement with previous studies,[17] we found that despite relatively weak SARS-CoV-2 serum neutralizing activity, potently neutralizing antibodies were readily isolated from the memory B cell compartment of these donors. The reduced neutralizing activity detected when the IGHV region was reverted to the germline configuration agrees with early studies showing the presence of Abs with no or very-low SHM with neutralizing activity against SARS-CoV-2.[21–23]

When swapping the germline V gene region of CAB-I47 from IGHV1-69*20 to IGHV1-69*04, resulting in a F55L change (the only difference between the two alleles), the neutralizing activity of CAB-I47 was greatly reduced and antigen affinity decreased 30-fold. Cryo-EM data of the CAB-I47 Fab in complex with the SARS-CoV-2 spike confirmed that F55 docked snuggly into a hydrophobic pocket of the RBD. The structure also showed that the IGHV1-69*20 germline-encoded R50 residue formed hydrogen bonds and a salt bridge with E484 on the RBD. This binding mode resembles that of Bamlanivimab, an IGHV1-69-using mAb that was shown to interact with RBD E484 partly via its germline-encoded R50.[51] However, Bamlanivimab has a leucine in the HCDR2 (position 55) and based on its amino acid sequence, is using either IGHV1-69*04 or IGHV1-69*09. Given that the E484K mutation was one of the earliest mutations to arise in SARS-CoV-2 Beta, it is likely that R50 is critical for the epitope specificity of these classes of antibodies. Altogether, our results from functional and structural studies of CAB-I47 demonstrate that both R50 and F55 contribute key contacts that mediate RBD binding by the IGHV1-69*20-using antibodies presented here.

When extending our study to the analysis of antibody sequences available in CoV-Ab-Dab, we identified 10 additional antibodies that could be assigned to IGHV1-69*20.[38,39,41–43,52] The nucleotide sequence was only available for 3 of 10 antibodies; however, it was possible to assign allele usage to all 10 due to the presence of a combination of amino acids that characterize IGHV1-69*20. Of particular note, two independent IGHV1-69*20-using antibodies, CV1206 and CV1182, were cloned from the same individual,[38] again demonstrating the propensity of IGHV1-69*20 carriers to produce multiple distinct SARS-CoV-2 neutralizing antibodies that use this allele. This analysis of antibodies from several donors, reported in independent studies, demonstrates that IGHV1-69*20-using SARS-CoV-2 neutralizing antibodies are relatively frequent in the broader population.

In addition to the well-known role of the hydrophobic IGHV1-69 HCDR2 in antibodies that bind the influenza HA stem region, antibodies using this gene were also found to be over-represented in hepatitis C virus (HCV) E2-binding neutralizing antibodies and structural studies demonstrated that HCDR2 F54 (F55 in the current study) was central for this interaction.[53,54] IGHV1-69 was also found to be frequently used in the B cell response to respiratory syncytial virus,[55] and there are likely additional examples of its function given its prominent usage in human B cell repertoire. Our knowledge about potential allele biases for different target specificities remains limited. Such studies will benefit greatly from the approach taken here, namely personalized IG genotyping coupled with mAb isolation, as this provides high confidence in antibody allele assignments.

The high allelic variation observed for IGHV1-69 is higher than that observed for other human IGHV genes. This diversity, together with the copy-number variation and frequent usage in naive IgM repertoires, suggests that this gene has played an important role during our evolutionary history. The presence of multiple alleles that vary in key positions in or near the CDRs, allows the production of antibodies with numerous binding modes, providing an extra layer of population antibody diversity, beyond that achieved by VDJ recombination and SHM within each individual. This study reveals extensive inter-individual IGHV gene and allele variation and illustrates that germline-encoded residues that are specific to given IGHV1-69 alleles shape SARS-CoV-2 neutralization. Together, these data indicate that IGH genotypes play a measurable role in our response to infection.

Limitations of the study

The B cell response against most pathogens is highly polyclonal, an intrinsic property of adaptive immunity. Therefore, limitations of this study include the number of study participants studied in-depth, as well as the number of IGHV1-69-using mAbs investigated from each donor. We chose to perform an extensive series of analyses for each spike-specific antibody to gain understanding of key germline-encoded residues differing between alleles. The goal of our study was to couple personalized IG genotyping with mAb isolation to begin to understand the role of allelic variation for functional Ab responses. Further studies of larger cohorts, different population groups, and additional targets are required for a further understanding of the role of IGH genotypes for functional Ab responses.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.immuni.2022.12.005.

ACKNOWLEDGMENTS

We would like to thank the study participants, as well as Robert Dyrdak, Jan Albert, Joakim Dillner at Karolinska Institutet, and the personnel at the Centre for Clinical Laboratory Studies (Karolinska University Hospital) for generous help with participant recruitment and blood sampling. We thank the G2P-UK National Virology consortium funded by MRC/UKRI (MR/W005611/1) and the Barclay group at Imperial College for providing spike expression plasmids. 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 toward equipment. All cryo-EM data were collected at the Karolinska Institutet's 3D-EM facility, and surface plasmon resonance analysis was performed at the Protein Science Facility at Karolinska Institutet. Funding for this work was provided by a Distinguished Professor grant from the Swedish Research Council (2017-00968), a European Research Council (ERC) Advanced Grant (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 G.B.K.H. Additional funding was provided from the Knut and Alice Wallenberg Foundation (KAW 2018.0080 and 2017.0080) and the Swedish Research Council (VR 2017-06702) to B.M.H. and from European Union's Horizon 2020 research and innovation programme (101003653, CoroNAb) to B.M. and G.B.K.H.

AUTHOR CONTRIBUTIONS

Conceptualization, P.P., M. Corcoran, and G.B.K.H.; methodology, P.P., H.D., M. Chernyshev, D.J.S., L.H., S.E., B.M.H., and M. Corcoran; investigation, P.P., A.N., D.J.S., H.D., X.C.D., L.P.V., M. Chernyshev, S.N., S.K., J.F., S.E., B.M.H., B.M., M. Corcoran, and G.B.K.H.; visualization, P.P., A.N., H.D., M. Chernyshev, and G.B.K.H.; resources, L.H., G.M., and B.M.H.; funding acquisition, G.B.K.H. and B.M.H.; supervision, M. Corcoran, B.M.H., B.M., and G.B.K.H.; writing – original draft, P.P., M. Corcoran, and G.B.K.H.; editing – original draft, all authors.

DECLARATION OF INTERESTS

M. Corcoran and G.B.K.H. are founders of ImmuneDiscover Sweden.

Received: May 28, 2022
Revised: September 23, 2022
Accepted: December 7, 2022
Published: December 12, 2022

REFERENCES

1. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. *50*, D543–D552. https://doi.org/10.1093/nar/gkab1038.
2. Asano, T., Boisson, B., Onodi, F., Matuozzo, D., Moncada-Velez, M., Maglorius Renkilaraj, M.R.L., Zhang, P., Meertens, L., Bolze, A., Materna, M., et al. (2021). X-linked recessive TLR7 deficiency in ~1% of men under 60 years old with life-threatening COVID-19. Sci. Immunol. *6*, eabl4348. https://doi.org/10.1126/sciimmunol.abl4348.
3. Zhang, Q., Bastard, P., Liu, Z., Le Pen, J., Moncada-Velez, M., Chen, J., Ogishi, M., Sabli, I.K.D., Hodeib, S., Korol, C., et al. (2020). Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. Science *370*, eabd4570. https://doi.org/10.1126/science.abd4570.
4. Zeberg, H., and Pääbo, S. (2021). A genomic region associated with protection against severe COVID-19 is inherited from Neandertals. Proc. Natl. Acad. Sci. USA *118*. e2026309118. https://doi.org/10.1073/pnas.2026309118.
5. Casanova, J.L. (2015). Human genetic basis of interindividual variability in the course of infection. Proc. Natl. Acad. Sci. USA *112*, E7118–E7127. https://doi.org/10.1073/pnas.1521644112.
6. Peng, K., Safonova, Y., Shugay, M., Popejoy, A.B., Rodriguez, O.L., Breden, F., Brodin, P., Burkhardt, A.M., Bustamante, C., Cao-Lormeau, V.M., et al. (2021). Diversity in immunogenomics: the value and the challenge. Nat. Methods *18*, 588–591. https://doi.org/10.1038/s41592-021-01169-5.
7. Watson, C.T., and Breden, F. (2012). The immunoglobulin heavy chain locus: genetic variation, missing data, and implications for human disease. Genes Immun. *13*, 363–373. https://doi.org/10.1038/gene.2012.12.
8. Matsuda, F., Ishii, K., Bourvagnet, P., Kuma, Ki, Hayashida, H., Miyata, T., and Honjo, T. (1998). The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. J. Exp. Med. *188*, 2151–2162. https://doi.org/10.1084/jem.188.11.2151.
9. Rodriguez, O.L., Gibson, W.S., Parks, T., Emery, M., Powell, J., Strahl, M., Deikus, G., Auckland, K., Eichler, E.E., Marasco, W.A., et al. (2020). A novel framework for characterizing genomic haplotype diversity in the human immunoglobulin heavy chain locus. Front. Immunol. *11*, 2136. https://doi.org/10.3389/fimmu.2020.02136.
10. Watson, C.T., Steinberg, K.M., Huddleston, J., Warren, R.L., Malig, M., Schein, J., Willsey, A.J., Joy, J.B., Scott, J.K., Graves, T.A., et al. (2013). 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. *92*, 530–546. https://doi.org/10.1016/j.ajhg.2013.03.004.
11. Corcoran, M.M., Phad, G.E., Vázquez Bernat, N., Stahl-Hennig, C., Sumida, N., Persson, M.A., Martin, M., and Karlsson Hedestam, G.B. (2016). Production of individualized V gene databases reveals high levels of immunoglobulin genetic diversity. Nat. Commun. *7*, 13642. https://doi.org/10.1038/ncomms13642.

12. Gidoni, M., Snir, O., Peres, A., Polak, P., Lindeman, I., Mikocziova, I., Sarna, V.K., Lundin, K.E.A., Clouser, C., Vigneault, F., et al. (2019). Mosaic deletion patterns of the human antibody heavy chain gene locus shown by Bayesian haplotyping. Nat. Commun. *10*, 628. https://doi.org/10.1038/s41467-019-08489-3.

13. Lee, J.H., Toy, L., Kos, J.T., Safonova, Y., Schief, W.R., Havenar-Daughton, C., Watson, C.T., and Crotty, S. (2021). Vaccine genetics of IGHV1-2 VRC01-class broadly neutralizing antibody precursor naïve human B cells. npj Vaccines *6*, 113. https://doi.org/10.1038/s41541-021-00376-7.

14. Avnir, Y., Tallarico, A.S., Zhu, Q., Bennett, A.S., Connelly, G., Sheehan, J., Sui, J., Fahmy, A., Huang, C.Y., Cadwell, G., et al. (2014). Molecular signatures of hemagglutinin stem-directed heterosubtypic human neutralizing antibodies against influenza A viruses. PLoS Pathog. *10*, e1004103. https://doi.org/10.1371/journal.ppat.1004103.

15. Avnir, Y., Watson, C.T., Glanville, J., Peterson, E.C., Tallarico, A.S., Bennett, A.S., Qin, K., Fu, Y., Huang, C.Y., Beigel, J.H., et al. (2016). IGHV1-69 polymorphism modulates anti-influenza antibody repertoires, correlates with IGHV utilization shifts and varies by ethnicity. Sci. Rep. *6*, 20842. https://doi.org/10.1038/srep20842.

16. Claireaux, M.e.a. (2022). A public antibody class recognizes a novel S2 epitope exposed on open conformations of SARS-CoV-2 spike. Preprint at bioRxiv. https://doi.org/10.1101/2021.12.01.470767v1.

17. Robbiani, D.F., Gaebler, C., Muecksch, F., Lorenzi, J.C.C., Wang, Z., Cho, A., Agudelo, M., Barnes, C.O., Gazumyan, A., Finkin, S., et al. (2020). Convergent antibody responses to SARS-CoV-2 in convalescent individuals. Nature *584*, 437–442. https://doi.org/10.1038/s41586-020-2456-9.

18. Sakharkar, M., Rappazzo, C.G., Wieland-Alter, W.F., Hsieh, C.L., Wrapp, D., Esterman, E.S., Kaku, C.I., Wec, A.Z., Geoghegan, J.C., McLellan, J.S., et al. (2021). Prolonged evolution of the human B cell response to SARS-CoV-2 infection. Sci. Immunol. *6*, eabg6916. https://doi.org/10.1126/sciimmunol.abg6916.

19. Yuan, M., Liu, H., Wu, N.C., and Wilson, I.A. (2021). Recognition of the SARS-CoV-2 receptor binding domain by neutralizing antibodies. Biochem. Biophys. Res. Commun. *538*, 192–203. https://doi.org/10.1016/j.bbrc.2020.10.012.

20. Luo, S., Yu, J.A., and Song, Y.S. (2016). Estimating copy number and allelic variation at the immunoglobulin heavy chain locus using short reads. PLoS Comput. Biol. *12*, e1005117. https://doi.org/10.1371/journal.pcbi.1005117.

21. Brouwer, P.J.M., Caniels, T.G., van der Straten, K., Snitselaar, J.L., Aldon, Y., Bangaru, S., Torres, J.L., Okba, N.M.A., Claireaux, M., Kerster, G., et al. (2020). Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. Science *369*, 643–650. https://doi.org/10.1126/science.abc5902.

22. Kreer, C., Zehner, M., Weber, T., Ercanoglu, M.S., Gieselmann, L., Rohde, C., Halwe, S., Korenkov, M., Schommers, P., Vanshylla, K., et al. (2020). Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing antibodies from COVID-19 patients. Cell *182*, 843–854.e12. https://doi.org/10.1016/j.cell.2020.06.044.

23. Seydoux, E., Homad, L.J., MacCamy, A.J., Parks, K.R., Hurlburt, N.K., Jennewein, M.F., Akins, N.R., Stuart, A.B., Wan, Y.H., Feng, J., et al. (2020). Analysis of a SARS-CoV-2-infected individual reveals development of potent neutralizing antibodies with limited somatic mutation. Immunity *53*, 98–105.e5. https://doi.org/10.1016/j.immuni.2020.06.001.

24. Feldman, J., Bals, J., Altomare, C.G., St Denis, K., Lam, E.C., Hauser, B.M., Ronsard, L., Sangesland, M., Moreno, T.B., Okonkwo, V., et al. (2021). Naive human B cells engage the receptor binding domain of SARS-CoV-2, variants of concern, and related sarbecoviruses. Sci. Immunol. *6*, eabl5842. https://doi.org/10.1126/sciimmunol.abl5842.

25. Kirik, U., Greiff, L., Levander, F., and Ohlin, M. (2017). Parallel antibody germline gene and haplotype analyses support the validity of immunoglobulin germline gene inference and discovery. Mol. Immunol. *87*, 12–22. https://doi.org/10.1016/j.molimm.2017.03.012.

26. Vázquez Bernat, N., Corcoran, M., Nowak, I., Kaduk, M., Castro Dopico, X., Narang, S., Maisonasse, P., Dereuddre-Bosquet, N., Murrell, B., and Karlsson Hedestam, G.B. (2021). Rhesus and cynomolgus macaque immunoglobulin heavy-chain genotyping yields comprehensive databases of germline VDJ alleles. Immunity *54*, 355–366.e4. https://doi.org/10.1016/j.immuni.2020.12.018.

27. Vázquez Bernat, N., Corcoran, M., Hardt, U., Kaduk, M., Phad, G.E., Martin, M., and Karlsson Hedestam, G.B. (2019). High-quality library preparation for NGS-based immunoglobulin germline gene inference and repertoire expression analysis. Front. Immunol. *10*, 660. https://doi.org/10.3389/fimmu.2019.00660.

28. Tiller, T., Meffre, E., Yurasov, S., Tsuiji, M., Nussenzweig, M.C., and Wardemann, H. (2008). Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. J. Immunol. Methods *329*, 112–124. https://doi.org/10.1016/j.jim.2007.09.017.

29. Barnes, C.O., Jette, C.A., Abernathy, M.E., Dam, K.A., Esswein, S.R., Gristick, H.B., Malyutin, A.G., Sharaf, N.G., Huey-Tubman, K.E., Lee, Y.E., et al. (2020). SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. Nature *588*, 682–687. https://doi.org/10.1038/s41586-020-2852-1.

30. Tan, T.J.C., Yuan, M., Kuzelka, K., Padron, G.C., Beal, J.R., Chen, X., Wang, Y., Rivera-Cardona, J., Zhu, X., Stadtmueller, B.M., et al. (2021). Sequence signatures of two public antibody clonotypes that bind SARS-CoV-2 receptor binding domain. Nat. Commun. *12*, 3815. https://doi.org/10.1038/s41467-021-24123-7.

31. Vanshylla, K., Fan, C., Wunsch, M., Poopalasingam, N., Meijers, M., Kreer, C., Kleipass, F., Ruchnewitz, D., Ercanoglu, M.S., Gruell, H., et al. (2022). Discovery of ultrapotent broadly neutralizing antibodies from SARS-CoV-2 elite neutralizers. Cell Host Microbe *30*, 69–82.e10. https://doi.org/10.1016/j.chom.2021.12.010.

32. Yuan, M., Liu, H., Wu, N.C., Lee, C.D., Zhu, X., Zhao, F., Huang, D., Yu, W., Hua, Y., Tien, H., et al. (2020). Structural basis of a shared antibody response to SARS-CoV-2. Science *369*, 1119–1123. https://doi.org/10.1126/science.abd2321.

33. Zhou, X., Ma, F., Xie, J., Yuan, M., Li, Y., Shaabani, N., Zhao, F., Huang, D., Wu, N.C., Lee, C.D., et al. (2021). Diverse immunoglobulin gene usage and convergent epitope targeting in neutralizing antibody responses to SARS-CoV-2. Cell Rep. *35*, 109109. https://doi.org/10.1016/j.celrep.2021.109109.

34. Zost, S.J., Gilchuk, P., Chen, R.E., Case, J.B., Reidy, J.X., Trivette, A., Nargi, R.S., Sutton, R.E., Suryadevara, N., Chen, E.C., et al. (2020). Rapid isolation and profiling of a diverse panel of human monoclonal antibodies targeting the SARS-CoV-2 spike protein. Nat. Med. *26*, 1422–1427. https://doi.org/10.1038/s41591-020-0998-x.

35. Phad, G.E., Pushparaj, P., Tran, K., Dubrovskaya, V., Àdori, M., Martinez-Murillo, P., Vázquez Bernat, N., Singh, S., Dionne, G., O'Dell, S., et al. (2020). Extensive dissemination and intraclonal maturation of HIV Env vaccine-induced B cell responses. J. Exp. Med. *217*, e20191155. https://doi.org/10.1084/jem.20191155.

36. Barnes, C.O., West, A.P., Jr., Huey-Tubman, K.E., Hoffmann, M.A.G., Sharaf, N.G., Hoffman, P.R., Koranda, N., Gristick, H.B., Gaebler, C., Muecksch, F., et al. (2020). Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. Cell *182*, 828–842.e16. https://doi.org/10.1016/j.cell.2020.06.025.

37. Sheward, D.J.P., P., Das, H., Kim, C., Kim, S., Hanke, L., Dyrdak, R., McInerney, G., Albert, A., Murrell, B., et al. (2022). Structural basis of Omicron neutralization by affinity-matured public antibodies. Preprint at bioRxiv. https://doi.org/10.1101/2022.01.03.474825v1.

38. Cho, H., Gonzales-Wartz, K.K., Huang, D., Yuan, M., Peterson, M., Liang, J., Beutler, N., Torres, J.L., Cong, Y., Postnikova, E., et al. (2021). Bispecific antibodies targeting distinct regions of the spike protein potently neutralize SARS-CoV-2 variants of concern. Sci. Transl. Med. *13*, eabj5413. https://doi.org/10.1126/scitranslmed.abj5413.

39. Scheid, J.F., Barnes, C.O., Eraslan, B., Hudak, A., Keeffe, J.R., Cosimi, L.A., Brown, E.M., Muecksch, F., Weisblum, Y., Zhang, S., et al. (2021). B cell genomics behind cross-neutralization of SARS-CoV-2 variants and SARS-CoV. Cell *184*, 3205–3221.e24. https://doi.org/10.1016/j.cell.2021.04.032.

40. Cao, Y., Yisimayi, A., Jian, F., Song, W., Xiao, T., Wang, L., Du, S., Wang, J., Li, Q., Chen, X., et al. (2022). BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. Nature *608*, 593–602. https://doi.org/10.1038/s41586-022-04980-y.

41. Hansen, J., Baum, A., Pascal, K.E., Russo, V., Giordano, S., Wloga, E., Fulton, B.O., Yan, Y., Koon, K., Patel, K., et al. (2020). Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. Science *369*, 1010–1014. https://doi.org/10.1126/science.abd0827.

42. He, B., Liu, S., Wang, Y., Xu, M., Cai, W., Liu, J., Bai, W., Ye, S., Ma, Y., Hu, H., et al. (2021). Rapid isolation and immune profiling of SARS-CoV-2 specific memory B cell in convalescent COVID-19 patients via LIBRA-seq. Signal Transduct. Target. Ther. *6*, 195. https://doi.org/10.1038/s41392-021-00610-7.

43. Cho, A., Muecksch, F., Schaefer-Babajew, D., Wang, Z., Finkin, S., Gaebler, C., Ramos, V., Cipolla, M., Mendoza, P., Agudelo, M., et al. (2021). Anti-SARS-CoV-2 receptor-binding domain antibody evolution after mRNA vaccination. Nature *600*, 517–522. https://doi.org/10.1038/s41586-021-04060-7.

44. Karlsson, E.K., Kwiatkowski, D.P., and Sabeti, P.C. (2014). Natural selection and infectious disease in human populations. Nat. Rev. Genet. *15*, 379–393. https://doi.org/10.1038/nrg3734.

45. Liston, A., Humblet-Baron, S., Duffy, D., and Goris, A. (2021). Human immune diversity: from evolution to modernity. Nat. Immunol. *22*, 1479–1489. https://doi.org/10.1038/s41590-021-01058-1.

46. Venkataraman, T., Valencia, C., Mangino, M., Morgenlander, W., Clipman, S.J., Liechti, T., Valencia, A., Christofidou, P., Spector, T., Roederer, M., et al. (2022). Analysis of antibody binding specificities in twin and SNP-genotyped cohorts reveals that antiviral antibody epitope selection is a heritable trait. Immunity *55*, 174–184.e5. https://doi.org/10.1016/j.immuni.2021.12.004.

47. Mikocziova, I., Greiff, V., and Sollid, L.M. (2021). Immunoglobulin germline gene variation and its impact on human disease. Genes Immun. *22*, 205–217. https://doi.org/10.1038/s41435-021-00145-5.

48. Zhang, J.Y., Roberts, H., Flores, D.S.C., Cutler, A.J., Brown, A.C., Whalley, J.P., Mielczarek, O., Buck, D., Lockstone, H., Xella, B., et al. (2021). Using de novo assembly to identify structural variation of eight complex immune system gene regions. PLoS Comput. Biol. *17*, e1009254. https://doi.org/10.1371/journal.pcbi.1009254.

49. Nadel, B., Tang, A., Lugo, G., Love, V., Escuro, G., and Feeney, A.J. (1998). Decreased frequency of rearrangement due to the synergistic effect of nucleotide changes in the heptamer and nonamer of the recombination signal sequence of the V kappa gene A2b, which is associated with increased susceptibility of Navajos to Haemophilus influenzae type b disease. J. Immunol. *161*, 6068–6073.

50. Sangesland, M., Torrents de la Peña, A., Boyoglu-Barnum, S., Ronsard, L., Mohamed, F.A.N., Moreno, T.B., Barnes, R.M., Rohrer, D., Lonberg, N., Ghebremichael, M., et al. (2022). Allelic polymorphism controls autoreactivity and vaccine elicitation of human broadly neutralizing antibodies against influenza virus. Immunity *55*, 1693–1709.e8. https://doi.org/10.1016/j.immuni.2022.07.006.

51. Jones, B.E., Brown-Augsburger, P.L., Corbett, K.S., Westendorf, K., Davies, J., Cujec, T.P., Wiethoff, C.M., Blackbourne, J.L., Heinz, B.A., Foster, D., et al. (2021). The neutralizing antibody, LY-CoV555, protects against SARS-CoV-2 infection in nonhuman primates. Sci. Transl. Med. *13*, eabf1906. https://doi.org/10.1126/scitranslmed.abf1906.

52. Cao, Y., Su, B., Guo, X., Sun, W., Deng, Y., Bao, L., Zhu, Q., Zhang, X., Zheng, Y., Geng, C., et al. (2020). Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients' B cells. Cell *182*, 73–84.e16. https://doi.org/10.1016/j.cell.2020.05.025.

53. Tzarum, N., Giang, E., Kong, L., He, L., Prentoe, J., Augestad, E., Hua, Y., Castillo, S., Lauer, G.M., Bukh, J., et al. (2019). Genetic and structural insights into broad neutralization of hepatitis C virus by human VH1-69 antibodies. Sci. Adv. *5*, eaav1882. https://doi.org/10.1126/sciadv.aav1882.

54. Weber, T., Potthoff, J., Bizu, S., Labuhn, M., Dold, L., Schoofs, T., Horning, M., Ercanoglu, M.S., Kreer, C., Gieselmann, L., et al. (2022). Analysis of antibodies from HCV elite neutralizers identifies genetic determinants of broad neutralization. Immunity *55*, 341–354.e7. e347. https://doi.org/10.1016/j.immuni.2021.12.003.

55. Gilman, M.S., Castellanos, C.A., Chen, M., Ngwuta, J.O., Goodwin, E., Moin, S.M., Mas, V., Melero, J.A., Wright, P.F., Graham, B.S., et al. (2016). Rapid profiling of RSV antibody repertoires from the memory B cells of naturally infected adult donors. Sci. Immunol. *1*, eaaj1879. https://doi.org/10.1126/sciimmunol.aaj1879.

56. Peres, A., Gidoni, M., Polak, P., and Yaari, G. (2019). RAbHIT: R antibody haplotype inference tool. Bioinformatics *35*, 4840–4842. https://doi.org/10.1093/bioinformatics/btz481.

57. Hanke, L., Vidakovics Perez, L., Sheward, D.J., Das, H., Schulte, T., Moliner-Morro, A., Corcoran, M., Achour, A., Karlsson Hedestam, G.B., Hällberg, B.M., et al. (2020). An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. Nat. Commun. *11*, 4420. https://doi.org/10.1038/s41467-020-18174-5.

58. Stewart, S.A., Dykxhoorn, D.M., Palliser, D., Mizuno, H., Yu, E.Y., An, D.S., Sabatini, D.M., Chen, I.S., Hahn, W.C., Sharp, P.A., et al. (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. RNA *9*, 493–501. https://doi.org/10.1261/rna.2192803.

59. Rogers, T.F., Zhao, F., Huang, D., Beutler, N., Burns, A., He, W.T., Limbo, O., Smith, C., Song, G., Woehl, J., et al. (2020). Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. Science *369*, 956–963. https://doi.org/10.1126/science.abc7520.

60. Tegunov, D., and Cramer, P. (2019). Real-time cryo-electron microscopy data preprocessing with Warp. Nat. Methods *16*, 1146–1152. https://doi.org/10.1038/s41592-019-0580-y.

61. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296. https://doi.org/10.1038/nmeth.4169.

62. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132. https://doi.org/10.1107/S0907444904019158.

63. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. *30*, 70–82. https://doi.org/10.1002/pro.3943.

64. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). Phenix: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221. https://doi.org/10.1107/S0907444909052925.

65. Hsieh, C.L., Goldsmith, J.A., Schaub, J.M., DiVenere, A.M., Kuo, H.C., Javanmardi, K., Le, K.C., Wrapp, D., Lee, A.G., Liu, Y., et al. (2020). Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. Science *369*, 1501–1505. https://doi.org/10.1126/science.abd0826.

66. Sheward, D.J., Mandolesi, M., Urgard, E., Kim, C., Hanke, L., Perez Vidakovics, L., Pankow, A., Smith, N.L., Castro Dopico, X., McInerney, G.M., et al. (2021). Beta RBD boost broadens antibody-mediated protection against SARS-CoV-2 variants in animal models. Cell Rep. Med. *2*, 100450. https://doi.org/10.1016/j.xcrm.2021.100450.

67. Hanke, L., Sheward, D.J., Pankow, A., Vidakovics, L.P., Karl, V., Kim, C., Urgard, E., Smith, N.L., Astorga-Wells, J., Ekström, S., et al. (2022). Multivariate mining of an alpaca immune repertoire identifies potent cross-neutralizing SARS-CoV-2 nanobodies. Sci. Adv. *8*, eabm0220. https://doi.org/10.1126/sciadv.abm0220.

68. Engen, J.R., and Wales, T.E. (2015). Analytical aspects of hydrogen exchange mass spectrometry. Annu. Rev. Anal. Chem. (Palo Alto. Calif) *8*, 127–148. https://doi.org/10.1146/annurev-anchem-062011-143113.

69. Masson, G.R., Burke, J.E., Ahn, N.G., Anand, G.S., Borchers, C., Brier, S., Bou-Assaf, G.M., Engen, J.R., Englander, S.W., Faber, J., et al. (2019). Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. Nat. Methods *16*, 595–602. https://doi.org/10.1038/s41592-019-0459-y.

70. Custódio, T.F., Das, H., Sheward, D.J., Hanke, L., Pazicky, S., Pieprzyk, J., Sorgenfrei, M., Schroer, M.A., Gruzinov, A.Y., Jeffries, C.M., et al. (2020). Selection, biophysical and structural analysis of synthetic nanobodies that effectively neutralize SARS-CoV-2. Nat. Commun. *11*, 5588. https://doi.org/10.1038/s41467-020-19204-y.

71. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. *46*, W296–W303. https://doi.org/10.1093/nar/gky427.

72. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501. https://doi.org/10.1107/S0907444910007493.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat Anti-Human IgG(H+L)-HRP	Southern Biotech	Cat#2015-05; RRID:AB_2795588
BD Pharmingen™ FITC Mouse Anti-Human CD3ε	BD Bioscience	Cat#556611; RRID:AB_396484
BD Pharmingen™ FITC Mouse Anti-Human CD14	BD Bioscience	Cat#557153; RRID:AB_396589
BD Horizon™ BV421 Mouse Anti-Human CD20	BD Bioscience	Cat#562873; RRID:AB_2737857
BD Pharmingen™ PE-Cy™7 Mouse Anti-Human CD27	BD Bioscience	Cat#560609; RRID:AB_1727456
BD Horizon™ PE-CF594 Mouse Anti-Human IgG	BD Bioscience	Cat#562538; RRID:AB_2737640
Fc Receptor Binding Inhibitor Polyclonal Antibody	Thermo Fisher	Cat#14-9161-73; RRID:AB_468582
Bacterial and virus strains		
E. coli XL10-Gold Ultracompetent Cells	Agilent Technologies	Cat#200315;
Biological samples		
Human PBMCs and Plasma	This paper	N/A
Chemicals, peptides, and recombinant proteins		
Stabilized Soluble Spike	Hsieh et al.[56]	N/A
Stabilized Soluble RBD	Hsieh et al.[56]	N/A
Biotinylated Spike	Hsieh et al.[56]	N/A
Agarose	Sigma-Aldrich	Cat#A9539; CAS:9012-36-6
Skim Milk Powder	Sigma-Aldrich	Cat#70166
Dimethyl Sulfoxide	Sigma-Aldrich	Cat#D2650
Ficoll-Paque PLUS	GE Healthcare	Cat#17-1440-03
Trypan Blue Stain (0.4%)	Thermo Fisher	Cat#T10282
SeeBlue Plus2 Pre-stained Protein Standard	Thermo Fisher	Cat#LC5925
NuPAGE MES SDS Running Buffer (20X)	Thermo Fisher	Cat#NP0002
NuPAGE™ LDS Sample Buffer (4X)	Thermo Fisher	Cat#NP0007
GeneRuler 1 kb Plus DNA Ladder	Thermo Fisher	Cat#SM1333
E-Gel™ Sample Loading Buffer	Thermo Fisher	Cat#10482055
GelRed Nucleic Acid Stain, 10,000x	Biotium	Cat#41003-1
Ampicillin, sodium salt, irradiated	Thermo Fisher	Cat#11593027
Antibiotic Antimycotic Solution (100×), Stabilized	Sigma-Aldrich	Cat#A5955
Dulbecco's Phosphate Buffered Saline	Sigma-Aldrich	Cat#D8537
ELISA TMB Stabilized Chromogen	Thermo Fisher	Cat#SB02
RPMI 1640 Medium	Cytiva	Cat#SH30096.01
FreeStyle™ 293 Expression Medium	Thermo Fisher	Cat#12338018
FreeStyle™ MAX Reagent	Thermo Fisher	Cat#16447100
Opti-MEM™ Reduced Serum Medium	Thermo Fisher	Cat#31985062
AMPure XP, 60 mL	Beckman Coulter	Cat#A63881
Streptavidin, Allophycocyanin Conjugate	Thermo Fisher	Cat#S32362
RiboLock RNase Inhibitor (40 U/μL)	Thermo Fisher	Cat#EO0382
KAPA HiFi HotStart ReadyMix	Roche Molecular Systems	Cat#7958935001
Lipofectamine 3000	Invitrogen	Cat# L3000075
Deuterium oxide 99.9%	Sigma-Aldrich	Cat#151882
Tris(2-carboxyethyl)phosphine hydrochloride (TCEP)	Sigma-Aldrich	Cat#75259
Critical commercial assays		
Pierce Fab Preparation Kit	Thermo Fisher	Cat#44985
Qubit dsDNA BR Assay Kits	Thermo Fisher	Cat#Q32853

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Qubit Protein BR Assay Kits	Thermo Fisher	Cat#A50668
MinElute Gel Extraction Kit	Qiagen	Cat#28606
QIAGEN Plasmid Plus Midi Kit	Qiagen	Cat#12945
QIAGEN Plasmid Plus Maxi Kit	Qiagen	Cat#12965
Gibson Assembly Master Mix	New England Biolabs	Cat#E2611L
LIVE/DEAD™ Fixable Aqua Dead Cell Stain Kit	Thermo Fisher	Cat#L34957
SuperScript™ IV Reverse Transcriptase	Thermo Fisher	Cat#18090010
HotStarTaq Plus DNA Polymerase	Qiagen	Cat#203603
Bright-Glo Luciferase Assay System	Promega	Cat#E2620
Series S Sensor Chip SA	GE Healthcare	Cat#BR-1005-31
MiSeq Reagent Kit v3	Illumina	Cat#MS-102-3001
PhiX control V3 kit	Illumina	Cat#FC-110-3001
Deposited data		
HC (VDJ) and LC (VJ) sequences of neutralizing mAbs	This paper	GenBank: OP497961-OP497964, ON086918-ON086947.
Repertoire sequence data	This paper	SciLifeLab Data Repository: http://doi.org/10.17044/scilifelab.19317512.
Mass spectrometry raw files and HDX analysis, deposited on ProteomeXchange Consortium via PRIDE	This paper	ProteomeXchange: PXD031945
Cryo-EM structures of CAB-I47 Fabs in complex with the spike trimer	This paper	PDB: 8A99, 8A94, and 8A96, and EMDB: EMD-15273, EMD-15269, and EMD-15271
Supplemental tables	This paper	Mendeley Data: https://doi.org/10.17632/mg7p5msrfs.1
Experimental models: Cell lines		
FreeStyle™ 293-F Cells	Thermo Fisher	Cat#R79007
HEK293T-ACE2	Hanke et al.[57]	N/A
Oligonucleotides		
Random Hexamers (50 μM)	Thermo Fisher	Cat#N8080127
Illumina index primers	Vázquez Bernat et al.[27]	N/A
Nested PCR primers	Vázquez Bernat et al.[27]	N/A
Recombinant DNA		
AbVec2.0-IGHG1	Tiller et al.[28]	Addgene 80795; RRID: Addgene_80795
AbVec1.1-IGKC	Tiller et al.[28]	Addgene 80796; RRID:Addgene_80796
AbVec1.1-IGLC2-XhoI	Tiller et al.[28]	Addgene 99575; RRID:Addgene_99575
pCMV-dR8.2 dvpr	Stewart et al.[58]	Addgene 8455; RRID:Addgene_8455
pBOBI-FLuc	Rogers et al.[59]	Addgene 170674; RRID:Addgene_170674
Software and algorithms		
HDExaminer, version 3.1.1	Sierra Analytics	http://massspec.com/hdexaminer
PyMOL	Schroedinger	http://pymol.org
Inkscape	Inkscape Project	https://inkscape.org
FlowJo v10	BD Bioscience	https://www.flowjo.com
R	R Foundation for Statistical Computing	https://www.R-project.org/
Python 3	Python Software Foundation	https://www.python.org/
IgDiscover, version 0.12.4.dev22+g0bc3365	Corcoran et al.[11]	http://docs.igdiscover.se/en/stable/
BIAevaluation	Cytiva	https://www.cytivalifesciences.com
GraphPad Prism, version 9	GraphPad Software	https://www.graphpad.com
RAbHIT v0.2.0	Peres et al.[56]	https://cran.r-project.org/web/packages/rabhit/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
EPU	Thermo Fisher	https://www.thermofisher.com/se/en/home/electron-microscopy/products/software-em-3d-vis/epu-software.html
WARP	Tegunov and Cramer[60]	http://www.warpem.com/warp/?page_id=65
cryoSPARCv3.2.0 and v.3.3.1	Punjani et al.[61]	https://cryosparc.com/
Coot	Emsley and Cowtan[62]	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
ChimeraX	Pettersen et al.[63]	https://www.rbvi.ucsf.edu/chimerax/
PHENIX 1.19.2	Adams et al.[64]	https://phenix-online.org/documentation/index.html
Other		
Countess Cell Counting Chamber Slides	Thermo Fisher	Cat#C10228
Qubit 3 Fluorometer	Thermo Fisher	Cat#Q33216
Countess II Automated Cell Counter	Thermo Fisher	Cat#AMQAX1000
NuPAGE 4 to 12%, Bis-Tris, 1.0 mm, Mini Protein Gel, 12-well	Thermo Fisher	Cat#NP0322BOX
BD FACSAria Fusion Cell Sorter	BD Bioscience	N/A
SepMate™-15	Stemcell Technologies	Cat#85415
E-Gel™ 96 Agarose Gels, 2%	Thermo Fisher	Cat#G700802
Mother E-Base™ Integrated Power Supply	Thermo Fisher	Cat#EBM03EU
Minisart NML Standard Syringe Filter, Pore Size 0.45 μm	Sartorius	Cat#S16555
Amicon Ultra-15 Centrifugal Filters	Merck Millipore	Cat#UFC903024
Amicon Ultra-4 Centrifugal Filters	Merck Millipore	Cat#UFC8050
Amicon Ultra-0.5mL Centrifugal Filters	Merck Millipore	Cat#UFC5100
HDX Workflow Solution	Trajan Automation	https://www.leaptec.com/products/hdx-workflow-solution
GloMax Navigator Microplate Luminometer	Promega	Cat#GM2000
GE Biacore 2000 SPR Analyzer	Biacore	N/A
MiSeq V3 2 x 300 cycle system	Illumina	https://www.illumina.com/
CryoMatrix® holey grids (R 2/1 geometry)	Zhenjiang Lehua Technology Co., Ltd	http://www.emasian.com/en/index.php?c=product&cid=3
UltrAuFoil® (R1.2/1.3, Gold 300 mesh)	Quantifoil Micro Tools	https://www.emsdiasum.com/microscopy/products/grids/UltrAuFoil.aspx

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Gunilla B. Karlsson Hedestam.

Materials availability

The reagents generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

- The spike-specific single HC and LC sequences of the mAbs have been deposited in GenBank and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Repertoire data are available from SciLifeLab Data Repository: http://doi.org/10.17044/scilifelab.19317512. The HDX data have been deposited to the ProteomeXchange Consortium via the PRIDE[1]partner repository with the dataset identifier ProteomeXchange: PXD031945. The cryo-EM structures are available at PDB and EMBD under the accession codes PDB: 8A99, 8A94, and 8A96, and EMDB: EMD-15273, EMD-15269, and EMD-15271.

- 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.
- The Supplemental Tables are available at Mendeley Data: https://doi.org/10.17632/mg7p5msrfs.1.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human samples

Employees at the Karolinska University Hospital in Stockholm, Sweden, were invited to participate in a study aimed at understanding memory B cell and antibody responses following primary SARS-CoV-2 infection. The study was approved by the National Ethical Review Agency of Sweden (decisions 2020-01620, 2020-02881 and 2020-05630) and approved consent was given by all participants. In brief, participants (gender and age not known) who tested positive in a SARS-CoV-2 RT-PCR test in May 2020 were invited to provide a blood sample in December 2020, approximately seven months after the RT-PCR test. Blood was drawn into EDTA tubes for subsequent isolation of PBMCs using density-gradient centrifugation. The isolated PBMCs were cryopreserved in fetal bovine serum with 10% dimethyl sulfoxide (Sigma) until use.

METHOD DETAILS

Individualized IGHV genotyping and haplotype analysis

To determine the germline IGHV, IGKV and IGLV allele content of the study participants, we generated full-length HC VDJ and LC VJ amplicons as previously described.[27] For the analysis, we used the IgDiscover software (http://docs.igdiscover.se/en/stable/). In brief, total RNA was isolated from approximately 5 million PBMCs using the Qiagen RNeasy mini kit, and 400 ng was reverse transcribed using Superscript IV (Invitrogen) and gene-specific primers for the IgM, IgK and IgL constant regions to generate three separate libraries. In addition to a molecular identifier (UMI), each primer contained a universal reverse primer sequence that enabled amplification when used in a PCR reaction containing multiplex forward primer sets for VH, VK and VL genes, respectively. Paired-end reads were generated using the Illumina MiSeq 2 x 300 bp kit, which resulted in library sizes of 737,111 (IgM), 893,975 (IgK) and 1,116,529 (IgL) from SP14 and a comparable IgM library size from SP13. These were analyzed using IgDiscover to produce individualized germline databases for heavy and light chain alleles. The IgG library used to assess IGHV gene usage in the total IgG repertoire was prepared as previously described.[27] The IgDiscover program contains a module named *plotalleles*, which allows inferred haplotype analysis. This is based on the principle that heterozygous IGHJ or IGHD alleles can be used to map associated IGHV alleles to the chromosome where they are present. Since V(D)J recombination occurs locally on a given chromosome, heterozygous IGHJ or IGHD alleles will only recombine with IGHV alleles present on the same chromosome and not those present on the other chromosome. Hence, maternally derived IGHV alleles will associate with the maternally derived IGHJ or IGHD allele, while paternally derived IGHV alleles will only associate with the paternally derived IGHJ and IGHD alleles. Thus, *plotalleles* creates a phased map of IGHV alleles in given study participants.

Analysis of IgM libraries from Gidoni et al.

Production of individual VDJ databases from bulk IgM was performed using IgDiscover[11] v0.12.4.dev22+g0bc3365 with the IMGT release 202209–1 (28 February 2022) as the starting database. After genotyping with IgDiscover, we haplotyped nine individuals from the Gidoni et al. study[12] by IGHJ6 heterozygosity with RAbHIT v0.2.0.[56] The gene usage boxplots in Figure S1B were generated from sequences with zero mutations in the V and J loci, which were clonally collapsed by identical V assignment, J assignment, and CDR3 nucleotide sequence.

Cell culture

HEK293F cells were cultured in FreeStyle™ 293 Expression Medium (Thermo Fisher)while HEK293T and HEK293T-ACE2 cells (stably expressing human ACE2)[57] were cultured in Dulbecco's Modified Eagle Medium (glucose, with sodium pyruvate) supplemented with 10% FBS, 100 U/ml penicillin, and 100 μg/ml streptomycin. Cultures were maintained in a humidified incubator (5% CO_2) at 37°C. The HEK293F cells were used for mAb expression and the HEK293T-ACE2 cells for neutralization assays.

Single-cell sorting of SARS-CoV-2 spike-specific memory B cells by flow cytometry

PBMCs were thawed and stained with fluorochrome-conjugated mAbs (and the S probe) before S-specific memory B cells were sorted using a BD FACSAria Fusion. The trimeric S probe was produced by conjugating biotinylated HexaPro spike[65] to streptavidin-allophycocyanin (SA-APC) (Invitrogen). The biotinylated spike was produced using NHS-Chromalink-Biotin (Solulink) with an average ratio of 3:1 (biotin:spike). The labeled spike was purified on PD-10 desalting columns. For the conjugation, 10 μg of spike protein was incubated with 1 μl of SA-APC (1mg/ml) for 20 minutes on a shaker at 4°C. The process was repeated until a total of 5 μl SA-APC had been added. All fluorescently labeled antibodies, as well as the conjugated probe were titrated before sorting. The antigen-specific population was defined as live $CD3^-CD14^-CD20^+CD27^+IgG^+spike^+$ and single cells were sorted into different

wells of 96-well PCR plates containing 4 μl of lysis buffer composed of 10x PBS, 100mM DTT and 40 U/μl RNAase inhibitor (Invitrogen). Plates with sorted cells were centrifuged, sealed, and immediately frozen on dry ice before storage at -80°C.

Single B cell RT-PCR

The 96-well plates containing lysed single B cells were thawed on ice and cDNA was generated by reverse transcription using random hexamers, dNTPs and SuperScript IV reverse transcriptase (Invitrogen). Heavy and light chain V(D)J sequences were amplified separately in 25 μl nested PCR reactions using 3 μl of cDNA in the 1st round PCR and 1.5 μl PCR product in the 2nd round PCR.[27] The HotStarTaq Plus Kit (Qiagen) and 5' leader sequence-specific and 3' IgG-specific primers were used. PCR products from positive wells were purified, Sanger sequenced (Genewiz) and analyzed.

Isolation of monoclonal antibodies

HC and LC V(D)J sequences were cloned into expression vectors containing the human IgG1, Igκ1, or Igλ2 constant regions[28] using Gibson assembly. Briefly, V(D)J sequences were synthesized with overhangs matching the ends of the linearized vectors. Gibson assembly was performed using a Gibson master mix (New England Biolabs), 50 ng of vector, 30 ng of insert in a 20 μl reaction, incubated at 50°C for 1 hour, and transformed into XL10-Gold ultracompetent cells by heat shock at 42°C for 30 seconds (Agilent Technologies). Colonies were screened by PCR and confirmed by Sanger sequencing (Genewiz). Positive colonies were expanded, and plasmids purified using a plasmid Midi-prep kit (Qiagen).

Expression and purification

mAbs were expressed by using 15 μg of each HC and LC plasmid and co-transfecting these into 30 ml of FreeStyle 293-F cells (1 million cells/ml ≥ 90% viability) cultured in FreeStyle 293-F medium (Life Technologies) using 30 μl of Max reagent (Life Technologies). mAbs were purified 7 days post-transfection using Protein G Sepharose columns (GE Healthcare) and analyzed under reducing conditions by SDS-PAGE using NuPAGE Novex 4-12% Bis-Tris polyacrylamide gels and NuPAGE reducing agent (Life Technologies). Germline-reverted versions of CAB-I47, CAB-M77, and CAB-N86 were generated by co-transfecting the respective germline HCs with the matching mature LCs.

ELISA

The mAbs were tested for binding to SARS-CoV-2 S and RBD. MaxiSorp 96-well plates (Nalgene Nunc International) were coated with S or RBD at 1 μg/ml in PBS overnight at 4°C. After incubation with blocking buffer (PBS containing 5% non-fat milk) or serum, mAbs (at a starting concentration of 5μg/ml) were added at 5-fold serial dilutions and incubated for 1.5 h at 37°C. Binding was detected using secondary HRP-conjugated anti-human Fcγ Ab (Jackson ImmunoResearch) at 1:10000 for 45 minutes at room temperature. The signal was developed through the addition of TMB substrate (Invitrogen) for 5 minutes, reactions were terminated with 1 M sulfuric acid, and the optical density (OD) was read at 450 nm. Between each incubation step, the plates were washed six times with PBS containing 0.05% Tween-20 (Sigma).

Preparation of Fab fragments

Fab fragments were prepared by digesting IgG with immobilized papain. The separation of Fab and Fc fragments was achieved with using Protein A column and the Pierce Fab Preparation Kit (ThermoFisher Scientific).

Pseudovirus neutralization assay

Expression vectors encoding SARS-CoV-2 spike proteins, B.1 (D614G), B.1.1.7 (Alpha), B.1.351 (Beta), P.1 (Gamma), B.1.617.2 (Delta), B.1.621 (Mu), and B.1.1.529 (Omicron) for pseudovirus neutralization assays were obtained from the G2P-UK National Virology consortium and Barclay Lab (Imperial College London, UK). Spike-pseudotyped lentivirus particles were generated by co-transfection of HEK293T cells with a spike-encoding plasmid, an HIV gag-pol packaging plasmid (Addgene #8455),[58] and a lentiviral transfer plasmid encoding firefly luciferase (Addgene #170674)[59] using polyethylenimine (PEI). Pseudovirus neutralization assays were performed using HEK293T-ACE2 cells, as described.[66] Briefly, pseudoviruses were incubated with serially diluted mAbs for 60 minutes at 37°C in black-walled 96-well plates. 10,000 HEK293T-ACE2 cells were then added to each well, and plates were incubated for 44-48 hours. Luminescence was measured using Bright-Glo (Promega) on a GloMax Navigator Luminometer (Promega). Neutralizing activity was calculated relative to the average of 8 control wells infected in the absence of serum. All neutralization experiments were repeated in two independent experiments.

Surface plasmon resonance

Kinetics for dilution series of the indicated Fabs were measured. Binding kinetics were determined using a Biacore 2000 instrument. All experiments were performed at 25°C in a running buffer composed of 10 mM HEPES, 150 mM NaCl, pH 7.4, and 0.005% Tween-20 (v/v). Site-specific biotinylated RBD was generated using Sortase A and amine-PEG3-biotin.[67] The probe was immobilized on streptavidin sensor chips (Series S sensor Chip SA, GE Healthcare) to ~100 resonance units (RU). A 2-fold dilution series of the Fab was injected at a flow rate of 30 μl/min (association 180s, dissociation 900s), and the immobilized RBD was regenerated using 10 mM glycine-HCL buffer pH 1.5 for 2 x 15 seconds. Data were analyzed using BIAevaluation Software and fitted using the 1:1 Langmuir model with mass transfer.

Hydrogen–deuterium exchange mass spectrometry (HDX-MS)

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 Fabs obtained from the digestion of CAB-F52 and CAB-I47, 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_2O. 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 μL/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 HDExaminer 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 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.[68] 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.[69] A summary of the HDX experimental detail can be found in Table S3. In addition, the mass spectrometry raw data and HDExaminer analysis files have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository[1] with the dataset identifier ProteomeXchange: PXD031945.

Cryo-EM sample preparation and imaging

Spike trimer (1.8 mg/ml) S2-P and the CAB-I47 Fab were mixed in a 1:6 molar ratio followed by incubation on ice for 10 min. Prior to cryo-EM grid preparation, grids were glow-discharged with 25 mA for 2 min using an EMS ×100 (Electron Microscopy Sciences) glow-discharge unit. CryoMatrix® holey grids with amorphous alloy film (R 2/1 geometry; Zhenjiang Lehua Technology Co., Ltd) were used. Three microliters of sample solutions were applied to the grids and the grids with sample were then vitrified in a Vitrobot Mk IV (Thermo Fisher Scientific) at 4 °C and 100% humidity (blot 10 s, blot force 3595 filter papers (Ted Pella Inc.)). Cryo-EM data collection was performed with EPU (Thermo Fisher Scientific) using a Krios G3i transmission-electron microscope (Thermo Fisher Scientific) operated at 300 kV in the Karolinska Institutet's 3D-EM facility (https://ki.se/cmb/3d-em). Images were acquired in 165 kx nanoprobe EFTEM SA mode with a slit width of 10 eV using a K3 Bioquantum. Exposure time was 2.0 s during which 60 movie frames were collected with a fluency of 0.80 e-/Å^2 per frame (Table S4). Motion correction, was performed on the fly using Warp.[60] Due to preferential orientation, we collected a part of the dataset with a 10- and 20-degree stage tilt.

A total of 14,460 movies were selected based on an estimated resolution cutoff of 4 Å and defocus below 2 microns as estimated by Warp. Motion corrected micrographs were imported to cryoSPARCv3.3.2[61] for CTF estimation, particle picking and further processing. In total 980,464 particles were picked, and these were processed for 2D classification and particles with good classes were selected for ab-initio models (two) generation. This was followed by heterogeneous refinement, and in the next step all picked particles were used for two consecutive rounds of heterogeneous refinement. This resulted in one class having 547,209 'good' particles that was used for non-uniform 3D refinement (C1 symm). This resulted in a reconstruction with an overall resolution 2.5Å (FSC 0.143). To separate different structural states, we performed 3DVA with four modes and an 8Å resolution limit. To speed up the 3DVA step particles were downsampled to 256 px box (2.02 Å/px). And resulted 4 modes were displayed as 5 clusters with 8Å resolution limit. Each cluster was then individually refined that resulted in two major states, 1-up with three Fabs and 2-up with 3-Fabs (other classes showing partial occupancy of the Fab). The particles belonging to these two clusters were then re-extracted in a 512 px box (1.01 Å/px) and non-uniform refinement was performed, resulting in maps with an overall resolution 2.5Å (FSC 0.143) for the 1-up state and 2.4Å for the 2-up state, respectively. To obtain high resolution information from the RBD-Fab47 interface we performed particle subtraction (leaving only one RBD-Fab) followed by local refinement (non-uniform). This resulted in a map with an overall resolution of 2.6Å. See image processing scheme for further details (Figure S5B).

Cryo-EM model building and structure refinement

The structure of the spike protein trimer 1-up PDB: 7A25 and 2-up PDB: 7A29[70] was used as starting models for model building. A homology model for the CAB-I47 Fab was generated by SWISS-MODEL.[71] Structure refinement and manual model building were performed using Coot and PHENIX[64] in interspersed cycles with secondary structure, Ramachandran, rotamers, and bond geometry restrains. Structure figures and EM density-map figures were generated with UCSF ChimeraX[63] and COOT,[62,72] respectively. The refinement and validation statistics are described in Table S5.

QUANTIFICATION AND STATISTICAL ANALYSIS

Neutralizing IC_{50} 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 IC50 correlations were determined using Pearson's R (GraphPad Software). The neutralization experiments are from two independent experiments.