Naive human B cells can neutralize SARS-CoV-2 through recognition of its receptor binding domain

Jared Feldman[1*], Julia Bals[1*] Kerri St. Denis[1], Evan C. Lam[1], Blake M. Hauser[1], Larance Ronsard[1], Maya Sangesland[1], Thalia Bracamonte Moreno[1], Vintus Okonkwo[1], Nathania Hartojo[1] Alejandro B. Balazs[1], Daniel Lingwood[1#] and Aaron G. Schmidt[1,2#]

[1]Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, 02139, USA

[2]Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

*equal contribution

#corresponding

Key words: germline precursors, B cells, immunity, SARS-CoV-2, coronaviruses

Correspondence:

Daniel Lingwood
Tel:857-268-7180; E-mail: dlingwood@mgh.harvard.edu

Aaron G. Schmidt
Tel: 857-268-7118; E-mail: aschmidt@crystal.harvard.edu

Running title: Isolation of antibody germline precursors that bind and neutralize SARS-CoV-2

ABSTRACT

Exposure to a pathogen elicits an adaptive immune response aimed to control and eradicate. This initial exposure is imprinted on the immune system, so that a subsequent encounter to the same pathogen or a variant will result in a memory recall response that is often protective. Interrogating the naive B cell repertoire in terms of both abundance and specificity to said pathogen may contribute to an understanding of how to potentially elicit protective responses. Here, we isolated naive B cells across 8 human donors, targeting the SARS-CoV-2 receptor-binding domain (RBD). Single B cell sorting, and subsequent sequence analysis, showed diverse gene usage and pairing with no apparent restriction on complementarity determining region length in either the heavy or light chains. We show that recombinantly expressed IgGs and Fabs of these germline precursors bind SARS-CoV-2 RBD. Importantly, a subset of these naive antibodies also bind SARS-CoV, an emergent variant (501Y.V2) and a potential pandemic (WIV-1) coronavirus. Furthermore, naive antibodies can also neutralize SARS-CoV-2 pseudoviruses in the absence of any somatic hypermutation, suggesting that protective immunity to coronaviruses, more broadly, may be genetically encoded. Future studies aimed at understanding the naive repertoire to other coronaviruses may ultimately reveal shared specificities that could be leveraged to develop pan-coronavirus vaccines aimed at priming encoded germline responses.

MAIN TEXT

Initial exposure to viral antigens by natural infection or vaccination primes an immune response and often establishes an immune memory which can prevent or control future infections. The naive repertoire contains potential B cell receptor (BCR) rearrangements capable of recognizing these antigens, often the surface-exposed glycoproteins. An early step in generating humoral immunity involves activation of these naive B cells through recognition of a cognate antigen (*1*) which in turn can lead to affinity maturation through somatic hypermutation (SHM) and subsequent differentiation (*2*). The initial engagement of the naive repertoire begins this cascade and often coincides with the eventual generation of a protective or neutralizing antibody response (*3, 4*).

For SARS-CoV-2, the etiological agent of COVID-19, the development of a neutralizing antibody response after primary infection or vaccination is associated with protection against reinfection in non-human primates (*5-8*). In humans, the presence of neutralizing antibodies can predict disease severity and survival after primary SARS-CoV-2 infection (*9*) and correlates with protection from secondary infection (*10, 11*). Further, memory B cells induced by natural infection in humans may persist and, in some cases, continue to affinity mature for greater than 6 months after primary infection (*12-15*) providing potentially durable long-term protection. Comparable levels of neutralizing antibody titers were present in convalescent COVID-19 subjects and vaccine recipients (*16-18*) further supporting the role of adaptive immune responses in helping to control and prevent disease severity. Understanding the origins of these protective responses is critical.

Both infection- and vaccine-elicited antibodies target the major envelope glycoprotein, spike, present on the virion surface (*18*). A substantial component of the neutralizing response engages the receptor binding domain (RBD) (*19-24*) and does so by sterically blocking interactions with the viral receptor ACE2 (*25-30*). Isolated RBD-directed monoclonal antibodies derive from diverse heavy- and light-chain variable gene segments suggesting that multiple biochemical solutions for developing RBD-directed antibodies are encoded within the human B-cell repertoire (*19, 21, 24, 31*). Potential immunogenicity of this antigenic site is based on the human naive B cell repertoire, and the overall frequency of naive BCRs that have some level of intrinsic affinity to stimulate their elicitation (*32-35*); in other words, to reliably amplify a protective SARS-CoV-2 response through vaccination or infection, the appropriate naive B cells must be both present (i.e.,

precursor frequency) and activatable (i.e., above an affinity threshold). However, antigen-specificity of naive B cells is largely undefined. Here, we characterize the human naive B cell repertoire specific for the SARS-CoV-2 RBD from healthy individuals in order to understand their relative abundance, affinity, and potential for providing a protective (i.e., neutralizing) response after natural infection or vaccination.

To measure the reactivity of naive human B cells specific for the SARS-CoV-2 RBD we adapted an *ex vivo* B cell profiling approach used previously to study epitope-specific naive precursors targeting neutralizing sites on HIV (*36-38*) and influenza virus surface glycoproteins (*32*). We first designed a SARS-CoV-2 RBD construct that positions two glycans at residues 475 and 501 to selectively block binding to ACE2 and the receptor-binding motif (RBM)-directed antibody, B38 (**fig. S1**) (*39*). Using this "ΔRBM" probe, in addition to wildtype SARS-CoV-2 spike, and RBD probes, we isolated naive ($CD19^+/IgD^+/IgG^-$) B cells specific to the RBD and, more finely, the RBM from the peripheral blood of 8 SARS-CoV-2 seronegative human donors (**Fig. 1A** and **fig. S2**). We defined RBM-specificity as B cells that bound to fluorescently labeled spike and RBD, but not the ΔRBM probe (**fig. S3A**). Although rare, all 8 donors had detectable populations of RBM-specific naive B cells (**fig. S3B**). The median frequency of RBM-specific B cells among total and naive B cells was 0.0021% and 0.0023%, respectively (**Fig. 1B**). Within spike-reactive, naive cells, the median frequency of RBM-specific B cells was 3.2% (**Fig. 1C**); this potentially suggests that a large proportion of spike epitopes targeted by naive responses reside outside of the RBD. The majority of IgD^+ RBM-specific B cells were $CD27^-$ (mean frequency ~97%), in agreement with the naive B cell phenotype (**fig. S3C**).

To understand in more detail the properties of this naive repertoire, we obtained 163 paired heavy- and light-chain antibody sequences from 5 of the 8 donors **(Fig. 1D** and **fig. S4)**. Sequence analysis showed that all clones were unique with diverse gene usage for both heavy and light chains and minimal gene pairing preferences (**fig. S4**). These data reflect the polyclonal gene usage observed in RBD-specific memory B cells sequenced from COVID-19 convalescent individuals (*21, 24, 31*) and vaccine recipients (*18*), suggesting that a diverse pool of antibody precursors can be activated upon antigen exposure. In comparing this naive repertoire to gene usage distribution from non-SARS-CoV-2-specific repertoires (*40*), we observed an increase in mean repertoire frequency of

~20% for IGHV3-9 in 4 out of 5 sequenced donors (**fig. S5A**). Notably, this enrichment of IGHV3-9 was also observed in isolated memory B cells from convalescent individuals (*41*) and vaccine recipients (*18*), as well as in expanded IgG^+ B cells sequenced from a cohort of COVID-19 subjects during acute infection (*31*). These expanded clones detected shortly after symptom onset displayed low levels of somatic hypermutation (SHM) (*31*), suggesting potential IGHV3-9 usage in an early extrafollicular response in which naive B cells differentiate into short-lived plasma cells (*42*). Additionally, IGHV3-53 and 3-30 gene segments, over-represented in RBD-specific antibodies isolated from convalescent subjects (*22, 30, 43*), were recovered from three sequenced donors (13 total clones; ~8.0% of total). The amino acid length of heavy and light chain third complementarity-determining regions (CDR3) ranged from 8 to 27 (average length ~16) for HCDR3 and 4 to 13 (average length ~10) for LCDR3 (**Fig. 1E**). These lengths are normally distributed relative to both unselected human repertoires (*40, 44*) and RBD-specific memory B cell repertoires (*18, 21, 22, 24*); this is in contrast to antibody precursors targeting the influenza and HIV receptor binding sites which have strict requirements for length (*45*) or gene usage (*46, 47*). These data suggest that overall HCDR3 length does not restrict precursor frequency and there appears no inherent bias for CDR3 length conferring RBM-specificity. The majority of obtained sequences were at germline in both the variable heavy (V_H) and light (V_L) chains. However, despite sorting B cells with a naive phenotype, some sequences were recovered that deviated from germline. Specifically, the V_H ranged from 91.4 to 100% identity to germline, with a median of 99.7%; the V_L ranged from 93.6 to 100%, with a median of 99.3% (**Fig. 1F**).

To obtain affinities of the isolated naive antibodies, we cloned and recombinantly expressed 44 IgGs selected to reflect the polyclonal RBD-specific repertoire with representatives from diverse variable region gene segments (**Table S1**). Additionally, we ensured diversity in terms of HCDR3 length, kappa and lambda usage, as well as representation from all 5 donors. By ELISA, we identified IgGs with detectable binding to SARS-CoV-2 RBD; we summarize these results for all antibodies (**Fig. 2**) and parsed by donor (**fig. S6A**). Across 5 donors, 36 (~81%) bound to monomeric SARS-CoV-2 RBD (**Fig. 2A**) with EC_{50} values ranging from 3.3 to 410 nM and a mean of 62 nM (**Fig. 2B** and **fig. S6A**). These antibodies included 32 unique variable heavy and light chain pairings (**Table S1**). We further defined the epitopic region of these IgGs using the ΔRBM construct and the individual glycan variants, Δ501 and Δ475, both of which independently

block ACE2 cell-surface binding but are on opposite sides of the RBM (**fig. S1E, F**). 11 IgGs had no detectable ΔRBM binding (e.g., ab079, ab119), while 21 IgGs had reduced ELISA binding relative to wild-type RBD (**Fig. 2A, B**), reflected in the reduced ΔRBM median EC_{50} values (**Fig. 2C**). We also identified examples of antibodies sensitive to only Δ475 (e.g., ab185) and only Δ501 (e.g., ab007) (**Fig. 2B** and **fig. S7**). Of the binding population, there is no apparent predisposition for HCDR3 length or light chain pairing (**Fig. 2 D, E**).

We next asked whether these antibodies could recognize the SARS CoV-2 variant 501Y.V2 currently circulating in South Africa (*48*). This variant includes a combination of the three RBD mutations: K417N, E484K, and N501Y (*48*). We find that 22 antibodies (50%) had detectable binding for this variant (**Fig. 2B**). Interestingly one antibody, ab033, seemed to be specific to this variant and did not bind SARS-CoV-2. We additionally tested the cross-reactivity of these antibodies to both SARS-CoV and WIV-1, related coronaviruses previously circulating and of pandemic potential (*49*), respectively. While the majority of antibodies (~86%) were specific to the SARS-CoV-2 RBD, 5 antibodies were also cross-reactive to both SARS-CoV and WIV-1. Of these cross-reactive antibodies, ab017 and ab114, derives from the same IGHV3-33 and IGVL2-14 paring but were isolated from different donors, suggesting a shared or public clonotype. Additionally, we found that ab189 was uniquely sensitive to the glycan at residue 475; this was surprising given the lack of conservation between of SARS-CoV, CoV-2 and WIV-1 RBMs, despite using the same ACE2 receptor (*50*). Collectively, these data demonstrate that coronavirus -specific naive antibodies are present in distinct donors, are of unrestricted gene usage and when as expressed as IgGs, have detectable affinity to RBDs from the currently administered vaccines, a circulating variant of concern, and from at least two related viruses.

To obtain binding kinetics independent of avidity effects from bivalent IgGs, 11 antibodies were selected for expression as Fabs to determine monovalent binding affinity (K_{DS}) by biolayer interferometry (BLI). Using monomeric RBD as the analyte, 9 of the 11 Fabs had detectable binding with K_{DS} ranging from ~6.5 to ~75 μM; the other two remaining Fabs (ab177, ab185), gave unreliable affinity measurements (i.e., >100 μM) (**fig. S8**). Notably, all Fabs had characteristically fast off rates (k_{off}). This observation is consistent for germline B cells where fast off-rates are compensated by avidity due to overall BCR surface density (*51*); subsequent affinity

gains via somatic hypermutation often result in slowing of the off-rate and is a canonical mechanism of improved antigen binding (*52-54*). While the biologically relevant antigen affinities of naive B cells are uncertain, recent studies have suggested that B cell activation and affinity maturation is not restricted by immeasurably low affinity BCR interactions (*55-57*), thus the isolated naive antibodies described here, especially those with demonstrable affinity may very well be engaged and elicited upon antigen exposure.

We next used a SARS-CoV-2 pseudovirus assay (*9*) to ask whether any of the isolated antibodies were capable of blocking transduction of target cells. We found that of the 36 RBD-binding antibodies tested in this assay, 5 had detectable levels of neutralization (~14%) (**Fig. 3A** and **fig. S9**). These neutralizing antibodies, obtained from multiple donors, have few commonalities with respect to their gene usages and HCDR3 lengths (**Fig. 3B**). While these naive antibodies were not as potent compared to the control antibody B38, isolated from a memory B cell (*29*), the observation, nevertheless, that the naive repertoire contains antibodies that have, at germline, the potential to neutralize is noteworthy.

Here, we showed that isolated naive antibodies from 5 different donors can bind SARS-CoV-2 RBD and its variant 501Y.V2, a proportion of which are cross-reactive to SARS CoV and WIV-1. These data suggest that RBM-specific precursors may be represented across a substantial fraction of individual human naive repertories. This observation is consistent with longitudinal studies of SARS-CoV-2 infected individuals in which the majority of convalescent individuals seroconverted with detectable RBD serum antibodies and neutralization titers (*14, 58*). A subset of these germline antibodies already shows neutralization activity, albeit weakly, against SARS-CoV-2 pseudoviruses. They therefore have the potential to affinity mature into more potent neutralizing antibodies after vaccination or natural infection. Indeed, recent studies have characterized potent RBD-directed neutralizing antibodies with a limited level of somatic hypermutation (*19, 21, 24, 31, 59*). The naive antibodies characterized here engage epitopes across the RBM with a range of angles of approach as defined by our glycan variant probes; this is consistent with other RBD-targeting antibodies characterized by epitope-mapping, deep mutational scanning and structures (*25, 27, 60*) (**Fig. 4**). Having BCRs within the naive repertoire that recognize distinct or partially overlapping epitopes across the RBM may be advantageous in

eliciting protective responses to variants of SARS-CoV-2 (*61-64*). Collectively, our observations described here for SARS-CoV-2 are notably in contrast to other viruses such as HIV, whose germline precursors targeting its receptor binding site lack demonstrable affinity, are highly restricted in their gene usages and cannot at germline neutralize the virus (*65-67*). Future investigations aimed at understanding the naive repertoire with respect to coronaviruses of pandemic potential may reveal further commonalties in antigen-specific precursors, enabling the development of pan-coronavirus vaccines aimed at priming specific germline responses of protective potential.

METHODS

Donor Samples

PBMCs were isolated from blood donors obtained from the MGH blood donor center (8 donors total). Prior to donating blood, subjects were required to sign a donor attestation/consent statement, as per hospital requirements, stating ‘‘I give permission for my blood to be used for transfusion to patients or for research’’. The gender and age are not recorded, however eligible donors are of at least 16 years old and weigh a minimum of 110lbs. All experiments were conducted with MGH Institutional Biosafety Committee approval (MGH protocol 2014B000035). Isolated PBMCs were used for B cell enrichment and single cell sorting (described below); plasma was aliquoted and stored at -80 ºC until further use. Additionally, the control convalescent sera used for ELISA was obtained under the approved Partners Institutional Review Board (protocol 2020P000895) for use of patient samples for the development and validation of SARS-CoV-2 diagnostic tests (*9*).

Expression and purification of recombinant CoV Antigens

Plasmids encoding the receptor binding domains (RBDs) were designed based on GenBank sequences MN975262.1 (SARS-CoV-2), ABD72970.1 (SARS-CoV), and AGZ48828.1 (WIV-1). Constructs were codon optimized and synthesized by IDT. Glycosylation sites at SARS-CoV-2 RBD residues 501 and/or 475 as well as 501Y.V2 mutations (K417N/E484K/N501Y) were inserted by QuikChange Mutagenesis (Agilent) following the manufacturer’s suggested protocol and were sequence confirmed (Genwiz). SARS-CoV-2 spike contained a C-terminal foldon trimerization domain and HRV 3C-cleavable 6xHis and 2xStrep II tags (*68*). All proteins were

transiently expressed in Expi293F cells (ThermoFisher). 5 to 7 days post-transfection, supernatants were harvested by centrifugation and further purified using immobilized metal affinity chromatography (IMAC) with cobalt-TALON resin (Takara) followed by Superdex 200 Increase 10/300 GL size exclusion column (GE Healthcare).

Expression and purification IgGs and Fabs

IgG and Fab genes for the heavy- and light-chain variable domains were synthesized and codon optimized by IDT and subcloned into pVRC protein expression vectors and sequence confirmed (Genewiz). Fabs and IgGs were similarly expressed and purified as described above for RBDs. IgGs were buffer exchanged into PBS while Fabs were concentrated and further purified by Superdex 200 Increase 10/300 GL size exclusion column.

ELISA

Both sera and monoclonal antibody reactivity to CoV antigens were assayed by ELISA. Briefly, 96-well plates (Corning) were coated with 5 µg/ml of monomeric RBDs diluted in PBS and incubated overnight at 4ºC. Plates were blocked with 1% BSA in PBS containing 1% Tween-20 (PBS-T) for 1hr at room temperature (RT). Blocking solution was discarded and 4-fold serial dilutions of human plasma (1:20 starting dilution) or isolated monoclonal antibodies (150 µg/ml starting concentration) in PBS were added to wells and incubated for 1hr at RT. Plates were then washed three times with PBS-T. Secondary, anti-human IgG-HRP (Abcam), was added to each well at 1:20,000 dilution in PBS-T and incubated for 1hr at RT. Plates were then washed three times with PBS-T and developed with 1-Step ABTS substrate (ThermoFisher) per manufacturer recommendations. Absorbance was measured using a plate reader at 405nm. EC_{50} values were determined for monoclonal antibodies by non-linear regression (sigmoidal) using GraphPad Prism 8.4.3 software. (GraphPad Software). ELISAs against 501Y.V2, SARS-CoV and WIV-1 RBDs were done at a single IgG concentration (150 µg/ml) in replicate. Positive binding was defined by an $OD_{405} \geq 0.30$.

276

277 **ACE-2 cell binding assay**

278 ACE-2 expressing 293T cells were incubated with 200 nM of RBD antigen in PBS for 1hr on ice.
279 Cells were resuspended in 50µL of secondary stain containing streptavidin-PE (Invitrogen) at a
280 1:200 dilution and incubated for 30 min on ice. Cell binding was analyzed by flow cytometry using
281 a Stratedigm S1300Exi Flow Cytometer equipped with a 96 well plate high throughput sampler.
282 Rsulting data were analyzed using FlowJo (10.7.1).

283

284 **Probe Generation**

285 SARS-CoV-2 RBD and ΔRBM constructs were expressed as dimeric murine-Fc (mFc; IgG1)
286 fusion proteins containing a HRV 3C-cleavable C-terminal 8xHis and SBP tags and purified as
287 described above. SBP-tagged RBD- and ΔRBM-mFc dimers were individually mixed with
288 fluorescently labeled streptavidin, SA-BV650 and SA-BV786 (BioLegend), to form RBD-mFc-
289 BV650 and ΔRBM-mFc-BV786 tetramers. SARS-CoV-2 spike with a C-terminal Strep II tag was
290 labeled separately with StrepTactin PE and APC (IBA) to form spike-PE and -APC tetramers,
291 respectively. Both labeling steps were performed for 30 min at 4 ºC prior to sorting.

292

293 **Single B Cell Sorting**

294 Naive B cells were purified from PBMCs using the MACS Human B Cell isolation kit (Miltenyi
295 Biotec) and incubated with 25nM of each SARS-CoV-2 probe (RBD-mFc-BV650, ΔRBM-mFc-
296 BV786, spike-PE, and spike-APC) for 30 min at 4ºC. Cells were stained with anti-human CD19
297 (Alexa-700), CD3 (PerCP-Cy5), IgD (PE-Cy7), IgG (BV711), CD27 (BV510), LiveDead Violet
298 (Invitrogen), and Calcien (Invitrogen) for an additional 30 min. RBM-specific naive B cells,
299 defined as $CD19^{+}/CD3^{-}/IgG^{-}/IgD^{+}$/spike PE^{+}/spike $APC^{+}/RBD^{+}/\Delta RBM^{--}$, were single-cell sorted
300 using BD FACS Aria II (BD Biosciences) into 96-well plates containing lysis buffer supplemented
301 with 1% BME. Within the $CD19^{+}/IgG^{-}/IgD^{+}$ gated cells, we also confirmed that 97% of the events
302 were CD27 negative. Plates were stored at -80 ºC for subsequent analysis. Flow cytometry data
303 was analyzed using FlowJo software version 10.7.1.

304

305 **BCR Sequencing**

BCR Sequencing was carried out as described previously (*32*). Briefly, whole transcriptome amplification (WTA) was performed on the sorted cell-lysates according to the Smart-Seq2 protocol (*69*). We then amplified heavy and light chain sequences from the WTA products utilizing pools of partially degenerate pools of V region specific primers (Qiagen HotStar Taq Plus). Heavy and light chain amplifications were carried out separately, with each pool containing pooled primers against human IGHV and heavy chain constant region genes, or human IGLV, IGKV, and light chain constant region genes. Cellular barcodes and index adapters (based on Nextera XT Index Adapters, Illumina Inc.) were added using a step-out PCR method. Amplicons were then pooled and sequenced using a 250x250 paired end 8x8 index reads on an Illumina Miseq System. The data were then demultiplexed, heavy and light chain reads were paired, and overlapping sequence reads were obtained (Panda-Seq) (*70*) and aligned against the human IMGT database (*71*).

Interferometry binding experiments

Interferometry experiments were performed using a BLItz instrument (ForteBio). Fabs (0.1 mg/ml) were immobilized on Ni-NTA biosensors. The SARS-CoV-2 RBD analyte was titrated (10µM, 5µM, 2.5µM, and 1µM) to acquire binding affinities; the K_D was obtained through global fit of the titration curves by applying a 1:1 binding isotherm using vendor-supplied software.

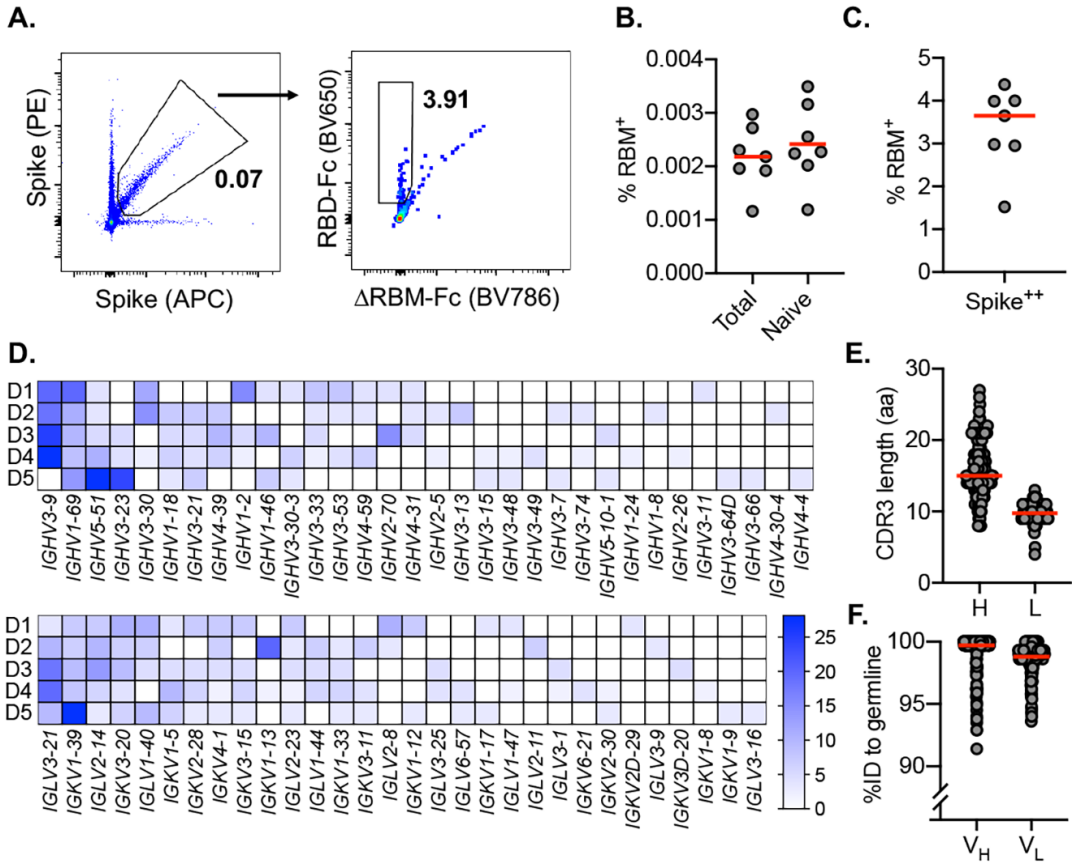
Pseudotyped Neutralization Assay

SARS-CoV-2 neutralization was assessed using lentiviral particles pseudotyped as previously described (*9, 72*). Briefly, lentiviral particles were produced via transient transfection of 293T cells. The titers of viral supernatants were determined via flow cytometry on 293T-ACE2 cells (*72*) and via the HIV-1 $p24^{CA}$ antigen capture assay (Leidos Biomedical Research, Inc.). Assays were performed in 384-well plates (Grenier) using a Fluent Automated Workstation (Tecan). IgGs starting at 150 µg/ml, were serially diluted (3-fold) in 20µL followed by addition of 20 µL of pseudovirus containing 250 infectious units and incubated at room temperature for 1 hr. Finally, 10,000 293T-ACE2 cells (*72*) in 20 µL cell media containing 15 µg/ml polybrene were added to each well and incubated at 37 °C for 60-72 hrs. Following transduction, cells were lysed using a previously described assay buffer (*73*) and shaken for 5 min prior to quantitation of luciferase expression using a Spectramax L luminometer (Molecular Devices). Percent neutralization was

337 determined by subtracting background luminescence measured from cells control wells (cells only)
338 from sample wells and dividing by virus control wells (virus and cells only). Data were analyzed
339 using Graphpad Prism.
340
341 **Acknowledgements**
342 We thank members of the Schmidt and Lingwood Labs for helpful discussions, especially Timothy
343 Caradonna and Daniel Maurer. We thank Goran Bajic for critical reading of the manuscript. We
344 thank Samuel Kazer, James Gatter and Alex Shalek for BCR sequencing advice, Jason McLellan
345 for the SARS-CoV-2 spike plasmid, and Nir Hacohen and Michael Farzan for ACE-2 expressing
346 293T cells. We acknowledge support from NIH (R01AI146779, R01AI124378, R01AI137057,
347 R01AI153098, R01AI155447, DP2DA042422, DP2DA040254, T32 AI007245), a Massachusetts
348 Consortium on Pathogenesis Readiness (MassCPR) grant to A.G.S. and a MGH Transformative
349 Scholars Program and Charles H. Hood Foundation to A.B.B.
350
351 **Author Information**
352 **Author Contributions** J.F., J.B., A.B.B., D.L. A.G.S. designed research; J.F., J.B., K.S.D., E.C.L,
353 B.M.H., L.R. M.S. T.B.M.,V.O., N.H., performed research; J.F., J.B., D.L. A.G.S analyzed data;
354 J.F. and A.G.S. wrote the paper. J.F., J.B., B.M.H, A.B.B., D.L. A.G.S edited and commented on
355 the paper.
356
357 **Correspondence and requests for materials should be addressed to:** Daniel Lingwood
358 (dlingwood@mgh.harvard.edu) or Aaron G. Schmidt (aschmidt@crystal.harvard.edu).
359
360 **Competing financial interest:** No competing financial interests.
361

362 **FIGURES**

363



364

365

366 **Fig. 1. SARS-CoV-2-specific naive B cells isolation and characterization**. **(A)** RBM-specific
367 naive B cells from seronegative human donors were isolated by fluorescence-activated cell sorting
368 gated on $CD19^{+}IgD^{+}IgG^{-}$; representative plot from donors 1 and 2 is shown. ΔRBM is a sorting
369 probe with N-linked glycans at residues 501 and 475. RBM-specific B cell frequency among (**B**)
370 total, naive, and (**C**) spike-positive cells from each donor ($n = 8$). (**D**) Heat map showing variable-
371 gene usage for all paired B cell sequences. Scale indicates percent of total sequences for each
372 donor separately. (**E**) Heavy (H) and light (L) CDR3 amino acid length distribution determined
373 using IMGT numbering. Red bars indicate median amino acid length. (**F**) Divergence from
374 inferred germline gene sequences. Red bars indicate the median percent values.

375

376

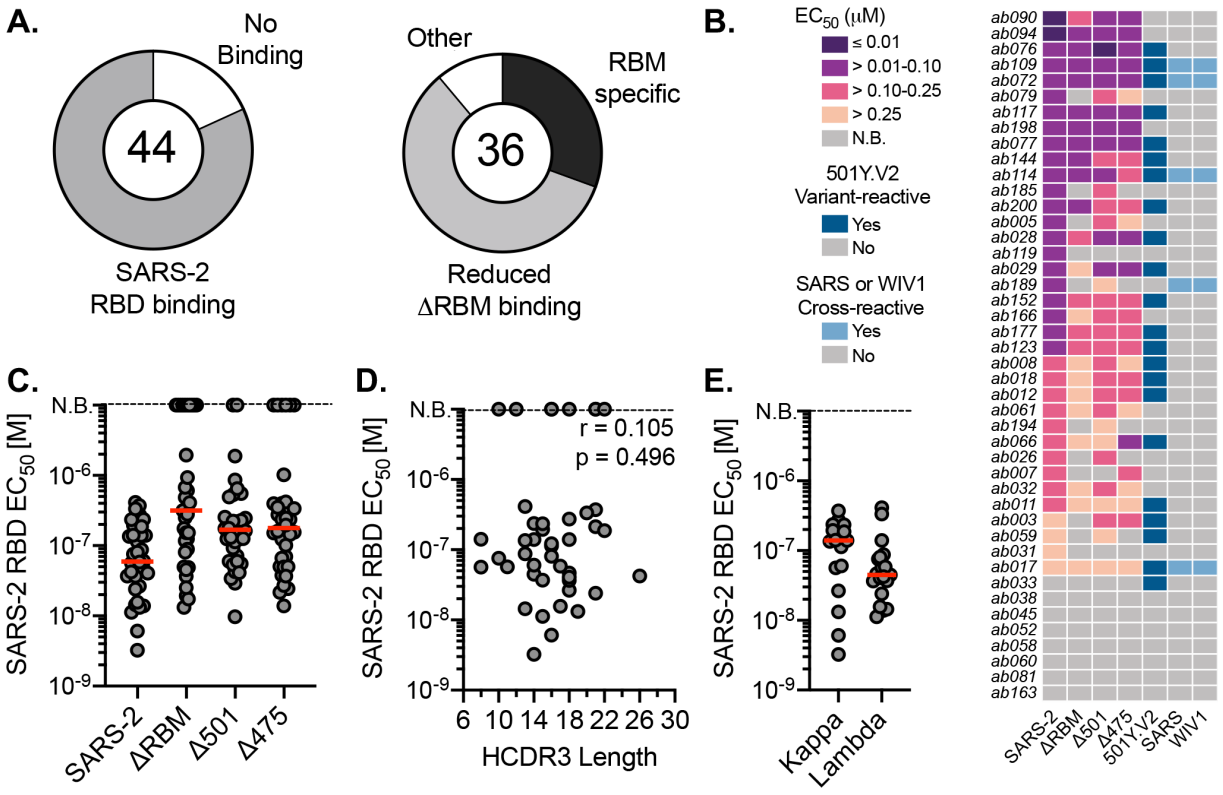
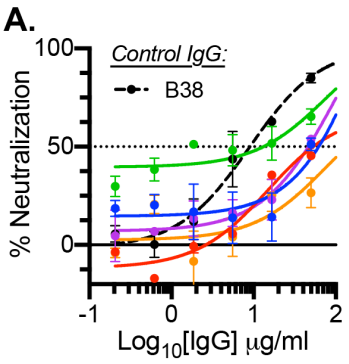


Fig. 2. Binding properties and specificity of isolated naive antibodies. (**A**) ELISA binding of 44 IgGs to SARS-CoV-2 RBD and ΔRBM binding (left and right panels, respectively). 36 IgGs with reduced ΔRBM binding were defined by >1.5 fold-change in ELISA EC_{50} relative to wild-type RBD. (**B**) IgG binding to SARS-CoV-2 (wildtype and 501Y.V2), CoV-1, WIV-1 RBDs and individual RBD glycan variants. (**C**) ELISA EC_{50} values for IgGs with detectable SARS-CoV-2 RBD binding (n = 36). Red bars indicate the mean EC_{50} values. (**D**) Pearson correlation analysis of SARS-CoV-2 RBD affinities and HCDR3 length. (**E**) ELISA EC_{50}s for IgGs based on kappa or lambda gene usage.

391

392



B.

	Heavy		Light	
ab ID	IGHV	HCDR3	IGLV	LCDR3
ab008	IGHV1-69	ARTSYASFAGLGWFDP	IGKV3-20	QQYGSSPGLT
ab029	IGHV5-51	ARSDGYSYGYY	IGLV2-14	SSYTSSRTPWV
ab066	IGHV1-69	ARQGKRTYYDILTGSPAGAFDI	IGKV4-1	QQYYSTPLIT
ab072	IGHV3-23	AKPPRGYYDSSGYYYEGWYFDL	IGLV2-14	SSYTSSSTLVV
ab090	IGHV1-2	ARHYGDYGEDAFDI	IGKV3-15	QQYNNWPLT

393

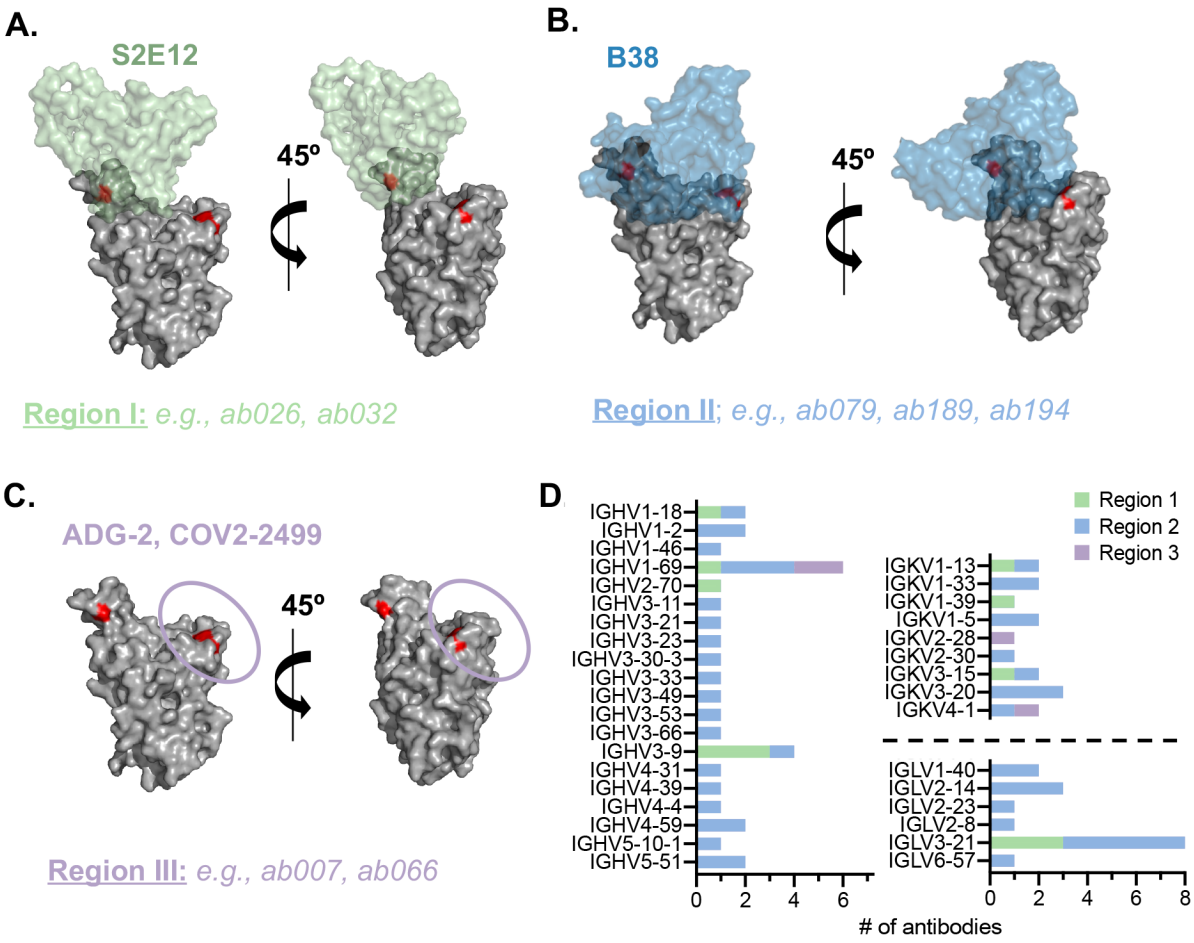
394

395 **Fig. 3. Naive antibodies neutralize SARS-CoV-2** (**A**) SARS-CoV-2 pseudovirus neutralization
396 assay for 5 IgGs. The neutralizing monoclonal antibody, B38, was used as a positive control.
397 Dashed lines indicate IC_{50} values and data represent means ± SD of two technical replicates. (**B**)
398 Variable region gene usage and CDR3 sequences for the IgGs.

399

400

401



402

403

404 **Fig. 4. SARS-CoV-2 RBD-specific naive antibodies and their epitope specificity.** (**A**) Epitopic
405 Region I defined by loss of binding to Δ475; structure of RBM-directed S2E12 (PDB 7K4N) and
406 representative naive IgGs (italics) in this study shown. (**B**) Epitopic Region II, defined by loss of
407 binding to ΔRBM but neither single glycan variant alone; structure of RBM-directed B38 (PDB
408 7BZ5) and representative naive IgGs. (**C**) Epitopic Region III, defined by loss of binding to Δ501
409 with representative naive IgGs (no structure is presently available to define this region). (**D**)
410 Summary of gene usage of isolated naive antibodies from each Region (**A**-**C**); bars represent total
411 number of antibodies and color-coded by each Region.

412

413

414

REFERENCES

1. N. E. Harwood, F. D. Batista, Early events in B cell activation. *Annual review of immunology* **28**, 185-210 (2009).
2. G. D. Victora, M. C. Nussenzweig, Germinal centers. *Annual review of immunology* **30**, 429-457 (2012).
3. C. Havenar-Daughton, R. K. Abbott, W. R. Schief, S. Crotty, When designing vaccines, consider the starting material: the human B cell repertoire. *Current opinion in immunology* **53**, 209-216 (2018).
4. F. Sallusto, A. Lanzavecchia, K. Araki, R. Ahmed, From vaccines to memory and back. *Immunity* **33**, 451-463 (2010).
5. W. Deng *et al.*, Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. *Science* **369**, 818-823 (2020).
6. A. Chandrashekar *et al.*, SARS-CoV-2 infection protects against rechallenge in rhesus macaques. *Science*, (2020).
7. N. B. Mercado *et al.*, Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature* **586**, 583-588 (2020).
8. Q. Gao *et al.*, Development of an inactivated vaccine candidate for SARS-CoV-2. *Science* **369**, 77-81 (2020).
9. W. F. Garcia-Beltran *et al.*, COVID-19 neutralizing antibodies predict disease severity and survival. *Cell*, (2020).
10. A. Addetia *et al.*, Neutralizing Antibodies Correlate with Protection from SARS-CoV-2 in Humans during a Fishery Vessel Outbreak with a High Attack Rate. *Journal of Clinical Microbiology* **58**, e02107-02120 (2020).
11. S. F. Lumley *et al.*, Antibodies to SARS-CoV-2 are associated with protection against reinfection. *medRxiv*, 2020.2011.2018.20234369 (2020).
12. A. Sokal *et al.*, Maturation and persistence of the anti-SARS-CoV-2 memory B cell response. *bioRxiv*, 2020.2011.2017.385252 (2020).
13. C. Gaebler *et al.*, Evolution of antibody immunity to SARS-CoV-2. *Nature*, (2021).
14. J. M. Dan *et al.*, Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science*, eabf4063 (2021).
15. L. B. Rodda *et al.*, Functional SARS-CoV-2-specific immune memory persists after mild COVID-19. *Cell*, (2020).
16. E. J. Anderson *et al.*, Safety and immunogenicity of SARS-CoV-2 mRNA-1273 vaccine in older adults. *New England Journal of Medicine*, (2020).
17. E. E. Walsh *et al.*, Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates. *New England Journal of Medicine* **383**, 2439-2450 (2020).
18. Z. Wang *et al.*, mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *bioRxiv*, 2021.2001.2015.426911 (2021).
19. P. J. M. Brouwer *et al.*, Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **369**, 643-650 (2020).
20. B. Ju *et al.*, Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115-119 (2020).
21. C. Kreer *et al.*, Longitudinal isolation of potent near-germline SARS-CoV-2-neutralizing antibodies from COVID-19 patients. *Cell* **182**, 843-854. e812 (2020).
22. D. F. Robbiani *et al.*, Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437-442 (2020).
23. A. Z. Wec *et al.*, Broad neutralization of SARS-related viruses by human monoclonal antibodies. *Science* **369**, 731-736 (2020).
24. S. J. Zost *et al.*, Rapid isolation and profiling of a diverse panel of human monoclonal antibodies targeting the SARS-CoV-2 spike protein. *Nature Medicine* **26**, 1422-1427 (2020).

25. C. O. Barnes *et al.*, SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature*, (2020).
26. C. O. Barnes *et al.*, Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **182**, 828-842.e816 (2020).
27. L. Piccoli *et al.*, Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* **183**, 1024-1042.e1021 (2020).
28. R. Shi *et al.*, A human neutralizing antibody targets the receptor binding site of SARS-CoV-2. *Nature*, 1-8 (2020).
29. Y. Wu *et al.*, A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science* **368**, 1274-1278 (2020).
30. M. Yuan *et al.*, Structural basis of a shared antibody response to SARS-CoV-2. *Science* **369**, 1119-1123 (2020).
31. S. C. A. Nielsen *et al.*, Human B Cell Clonal Expansion and Convergent Antibody Responses to SARS-CoV-2. *Cell Host & Microbe* **28**, 516-525.e515 (2020).
32. M. Sangesland *et al.*, Germline-encoded affinity for cognate antigen enables vaccine amplification of a human broadly neutralizing response against influenza virus. *Immunity* **51**, 735-749. e738 (2019).
33. P. Dosenovic *et al.*, Anti–HIV-1 B cell responses are dependent on B cell precursor frequency and antigen-binding affinity. *Proceedings of the National Academy of Sciences*, 201803457 (2018).
34. R. K. Abbott *et al.*, Precursor frequency and affinity determine B cell competitive fitness in germinal centers, tested with germline-targeting HIV vaccine immunogens. *Immunity* **48**, 133-146. e136 (2018).
35. A. Amitai *et al.*, Defining and Manipulating B cell Immunodominance Hierarchies to Elicit Broadly Neutralizing Antibody Responses Against Influenza Virus. *Cell Systems* **11**, 573-588. e579 (2020).
36. C. Havenar-Daughton *et al.*, The human naive B cell repertoire contains distinct subclasses for a germline-targeting HIV-1 vaccine immunogen. *Science translational medicine* **10**, eaat0381 (2018).
37. J. G. Jardine *et al.*, HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. *Science* **351**, 1458-1463 (2016).
38. J. M. Steichen *et al.*, A generalized HIV vaccine design strategy for priming of broadly neutralizing antibody responses. *Science* **366**, (2019).
39. B. M. Hauser *et al.*, Engineered receptor binding domain immunogens elicit pan-coronavirus neutralizing antibodies. *bioRxiv*, 2020.2012.2007.415216 (2020).
40. B. Briney, A. Inderbitzin, C. Joyce, D. R. Burton, Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature* **566**, 393-397 (2019).
41. B. Ju *et al.*, Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115-119 (2020).
42. I. C. MacLennan *et al.*, Extrafollicular antibody responses. *Immunological reviews* **194**, 8-18 (2003).
43. Y. Cao *et al.*, Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* **182**, 73-84.e16 (2020).
44. C. Soto *et al.*, High frequency of shared clonotypes in human B cell receptor repertoires. *Nature* **566**, 398-402 (2019).
45. A. G. Schmidt *et al.*, Viral receptor-binding site antibodies with diverse germline origins. *Cell* **161**, 1026-1034 (2015).
46. T. Zhou *et al.*, Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity* **39**, 245-258 (2013).

47. A. P. West, R. Diskin, M. C. Nussenzweig, P. J. Bjorkman, Structural basis for germ-line gene usage of a potent class of antibodies targeting the CD4-binding site of HIV-1 gp120. *Proceedings of the National Academy of Sciences* **109**, E2083-E2090 (2012).
48. H. Tegally *et al.*, Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv*, 2020.2012.2021.20248640 (2020).
49. V. D. Menachery *et al.*, SARS-like WIV1-CoV poised for human emergence. *Proceedings of the National Academy of Sciences* **113**, 3048-3053 (2016).
50. C. G. Rappazzo *et al.*, Broad and potent activity against SARS-like viruses by an engineered human monoclonal antibody. *Science*, eabf4830 (2021).
51. D. Lingwood *et al.*, Structural and genetic basis for development of broadly neutralizing influenza antibodies. *Nature* **489**, 566 (2012).
52. A. G. Schmidt *et al.*, Immunogenic stimulus for germline precursors of antibodies that engage the influenza hemagglutinin receptor-binding site. *Cell reports* **13**, 2842-2850 (2015).
53. F. D. Batista, M. S. Neuberger, Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity* **8**, 751-759 (1998).
54. J. M. Tas *et al.*, Visualizing antibody affinity maturation in germinal centers. *Science* **351**, 1048-1054 (2016).
55. C. Viant *et al.*, Antibody Affinity Shapes the Choice between Memory and Germinal Center B Cell Fates. *Cell* **183**, 1298-1311. e1211 (2020).
56. R. Di Niro *et al.*, Salmonella infection drives promiscuous B cell activation followed by extrafollicular affinity maturation. *Immunity* **43**, 120-131 (2015).
57. M. Kuraoka *et al.*, Complex antigens drive permissive clonal selection in germinal centers. *Immunity* **44**, 542-552 (2016).
58. D. Stadlbauer *et al.*, Repeated cross-sectional sero-monitoring of SARS-CoV-2 in New York City. *Nature*, 1-5 (2020).
59. S. I. Kim *et al.*, Stereotypic neutralizing VH antibodies against SARS-CoV-2 spike protein receptor binding domain in COVID-19 patients and healthy individuals. *Science Translational Medicine*, (2021).
60. A. J. Greaney *et al.*, Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host & Microbe*, (2020).
61. A. J. Greaney *et al.*, Comprehensive mapping of mutations to the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human serum antibodies. *bioRxiv*, 2020.2012. 2031.425021 (2021).
62. A. S. Lauring, E. B. Hodcroft, Genetic Variants of SARS-CoV-2—What Do They Mean? *JAMA*, (2021).
63. N. G. Davies *et al.*, Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/01 in England. *medRxiv*, (2020).
64. H. Tegally *et al.*, Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv*, (2020).
65. S. Hoot *et al.*, Recombinant HIV envelope proteins fail to engage germline versions of anti-CD4bs bNAbs. *PLoS Pathog* **9**, e1003106 (2013).
66. J. F. Scheid *et al.*, Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333**, 1633-1637 (2011).
67. T. Zhou *et al.*, Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329**, 811-817 (2010).
68. D. Wrapp *et al.*, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).
69. J. J. Trombetta *et al.*, Preparation of single-cell RNA-seq libraries for next generation sequencing. *Current protocols in molecular biology* **107**, 4.22. 21-24.22. 17 (2014).

70. A. P. Masella, A. K. Bartram, J. M. Truszkowski, D. G. Brown, J. D. Neufeld, PANDAseq: paired-end assembler for illumina sequences. *BMC bioinformatics* **13**, 1-7 (2012).
71. B. Shi *et al.*, Comparative analysis of human and mouse immunoglobulin variable heavy regions from IMGT/LIGM-DB with IMGT/HighV-QUEST. *Theoretical Biology and Medical Modelling* **11**, 1-11 (2014).
72. M. J. Moore *et al.*, Retroviruses pseudotyped with the severe acute respiratory syndrome coronavirus spike protein efficiently infect cells expressing angiotensin-converting enzyme 2. *Journal of virology* **78**, 10628-10635 (2004).
73. E. Siebring-van Olst *et al.*, Affordable luciferase reporter assay for cell-based high-throughput screening. *Journal of biomolecular screening* **18**, 453-461 (2013).

Supplementary Information

Naive human B cells can neutralize SARS-CoV-2 through recognition of its receptor binding domain

Jared Feldman[1*], Julia Bals[1*] Kerri St. Denis[1], Evan C. Lam[1], Blake M. Hauser[1], Larance Ronsard[1], Maya Sangesland[1], Thalia Bracamonte Moreno[1], Vintus Okonkowo[1], Nathania Hartojo[1] Alejandro B. Balazs[1], Daniel Lingwood[1#] and Aaron G. Schmidt[1,2#]

[1]Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, 02139, USA

[2]Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA

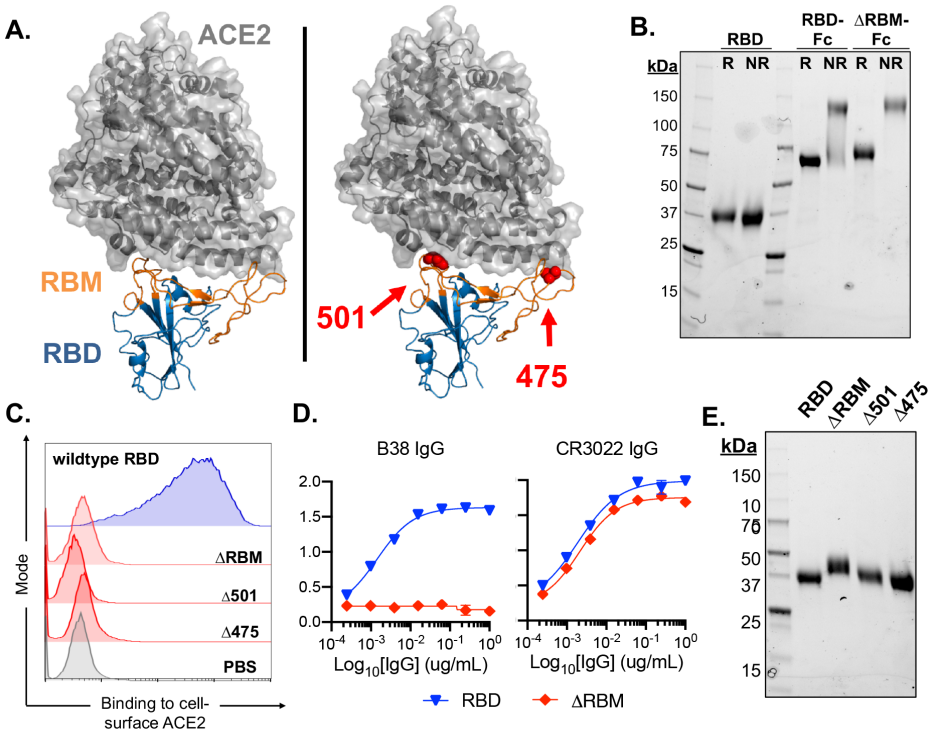


fig. S1. Design and characterization of SARS-CoV-2 antigens. (A) SARS-CoV-2 RBD in complex with viral receptor, ACE2 shown in blue and grey, respectively (PDB 6M0J). Wild-type RBD with, the receptor binding motif (RBM), shown in orange (left panel). Structural model of the ΔRBM probe designed to abrogate binding to ACE2 (right panel). Putative N-linked glycosylation sites engineered onto the RBM are shown in red spheres at amino acid positions 501 and 475. (**B**) SDS-PAGE gel under reducing (R) and non-reducing (NR) conditions for monomeric RBD, RBD-Fc and ΔRBM-Fc. (**C**) Wildtype RBD, ΔRBM and single glycan variant binding to ACE2-expressing 293T cells by flow cytometry. Wild-type RBD binding shown in blue, glycan variant binding shown in red. Streptavidin-PE was used to detect the relative intensity of antigen binding to cell-surface ACE2. A PBS control (gray) indicates secondary-only staining. (**D**) Control antibody ELISA binding to RBD and ΔRBM antigens. RBM-specific antibody, B38 (left). Non-RBM-specific control antibody, CR3022 (right). (**E**) ΔRBM and Δ501 and Δ475 variants analyzed by SDS-PAGE gel under reducing conditions; wildtype RBD is shown for comparison.

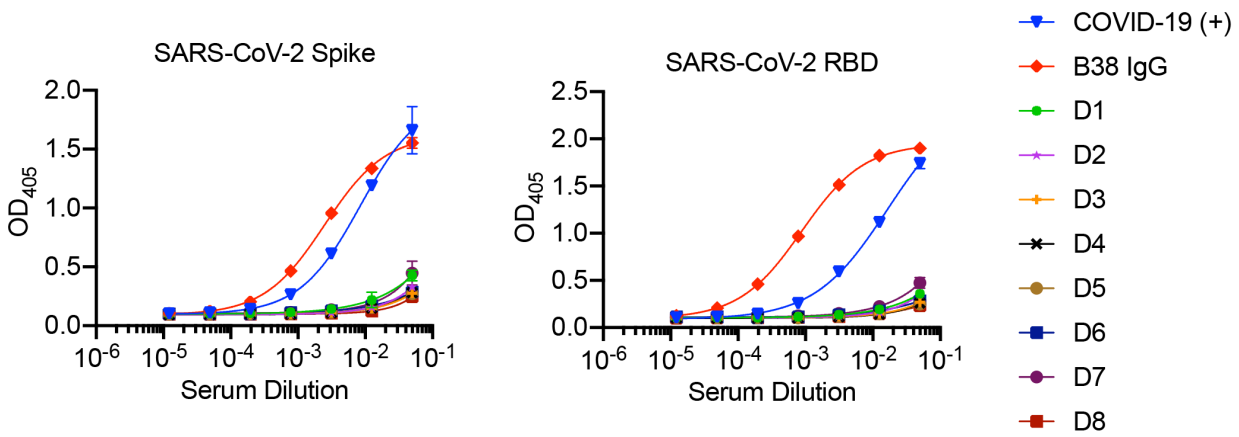


fig. S2. ELISA binding reactivity of healthy donor sera to SARS-CoV-2 spike and RBD. Seronegative donors 1-8 were single-cell B sorted to isolate RBD-specific naive B cells. Sera from a COVID-19 convalescent patient and control antibody, B38, were included as positive controls.

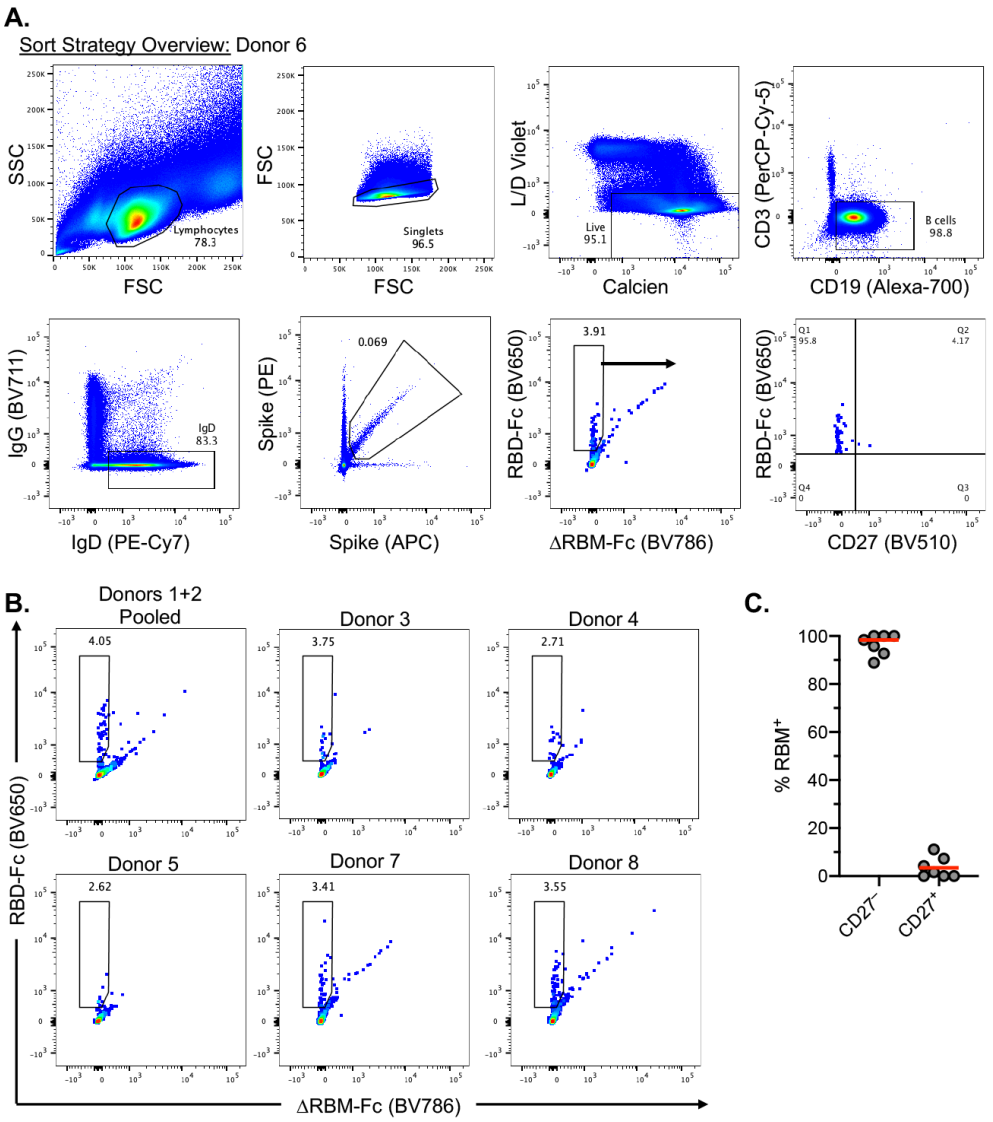


fig. S3. **PBMC flow cytometry analyses.** (**A**) Representative gating strategy used for FACS of PBMCs pooled from donors 1 and 2. Gating was on naive B cells defined by single living lymphocytes that were $CD19^{+}CD3^{-}IgD^{+}IgG^{-}$. Sorted cells were RBM-specific as defined by spike-PE^{+}/spike-APC^{+}/RBD-Fc-$BV650^{+}$/ΔRBM-Fc-BC650-. Sort gate is denoted by the blue arrow. The bottom right plot shows CD27 staining of sorted RBM-specific naive B cells. (**B**) Flow cytometry showing the sort gate and percentage of RBM-specific B cells for the remaining 6 healthy human donors. (**C**) RBM-specific B cell frequency among $CD27^{+}$ and $CD27^{-}$ cells. Each symbol represents a different donor (n = 8).

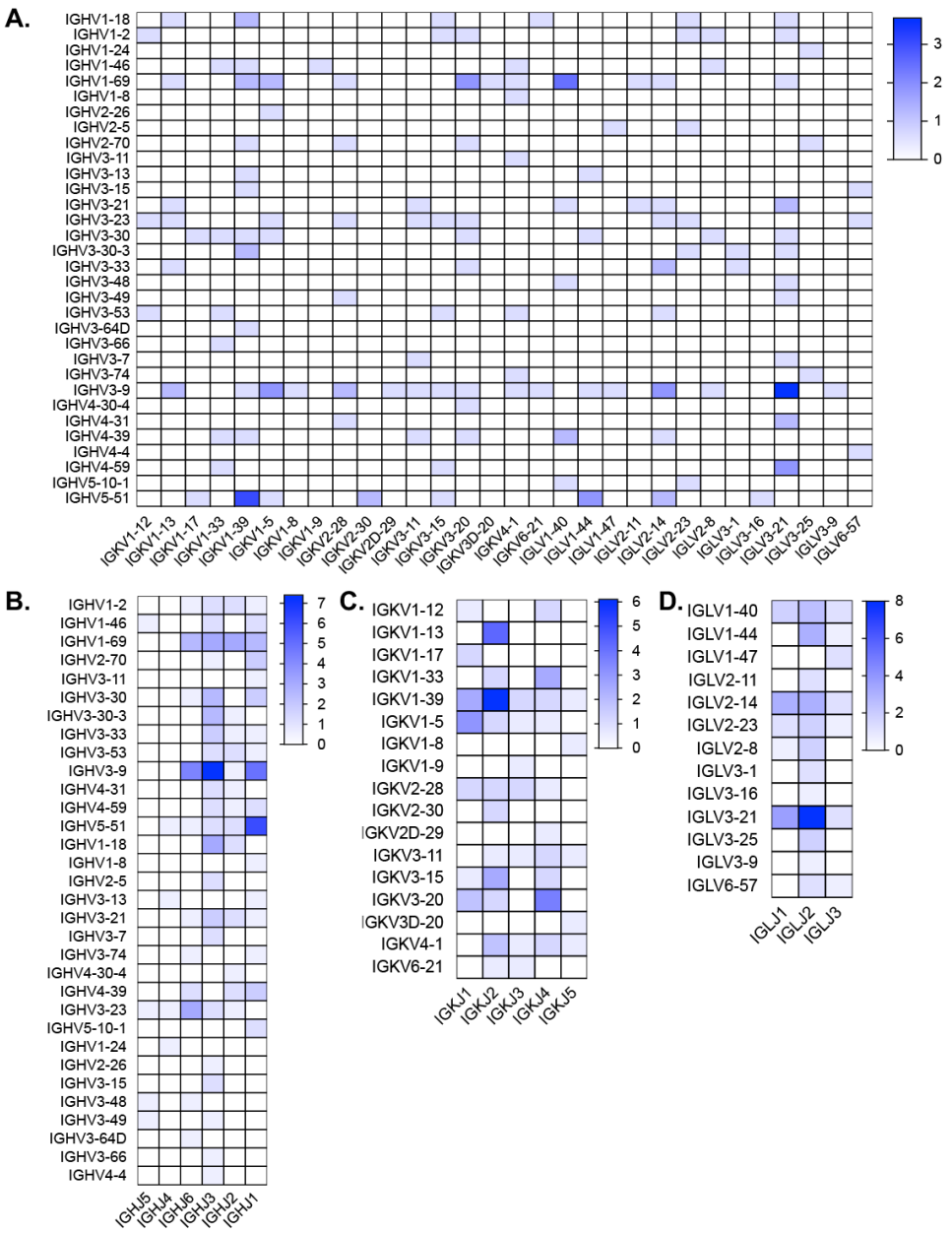
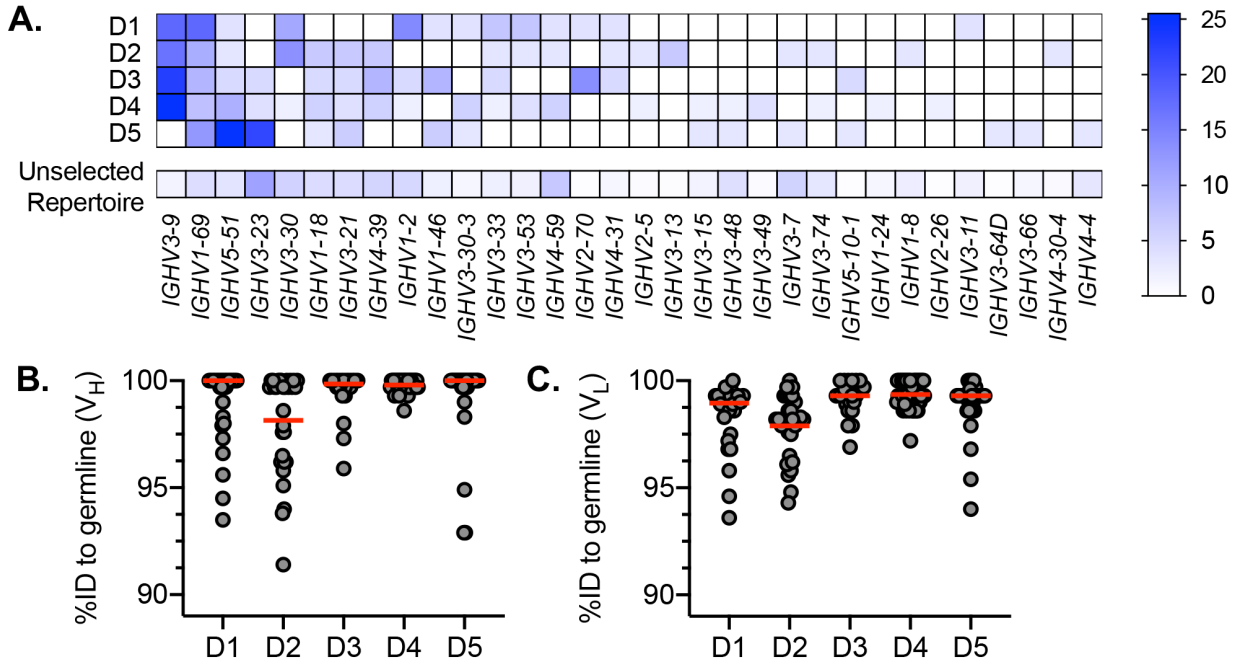


fig. S4. Variable-gene usage of isolated naive antibodies. (**A**) Heatmap showing frequency of V_H:V_L pairing. Overall frequency of variable (V) and joining (J) genes for (**B**) heavy chains, (**C**) kappa light chains, and (**D**) lambda light chains. Heatmap scales represent percent of total paired sequences.



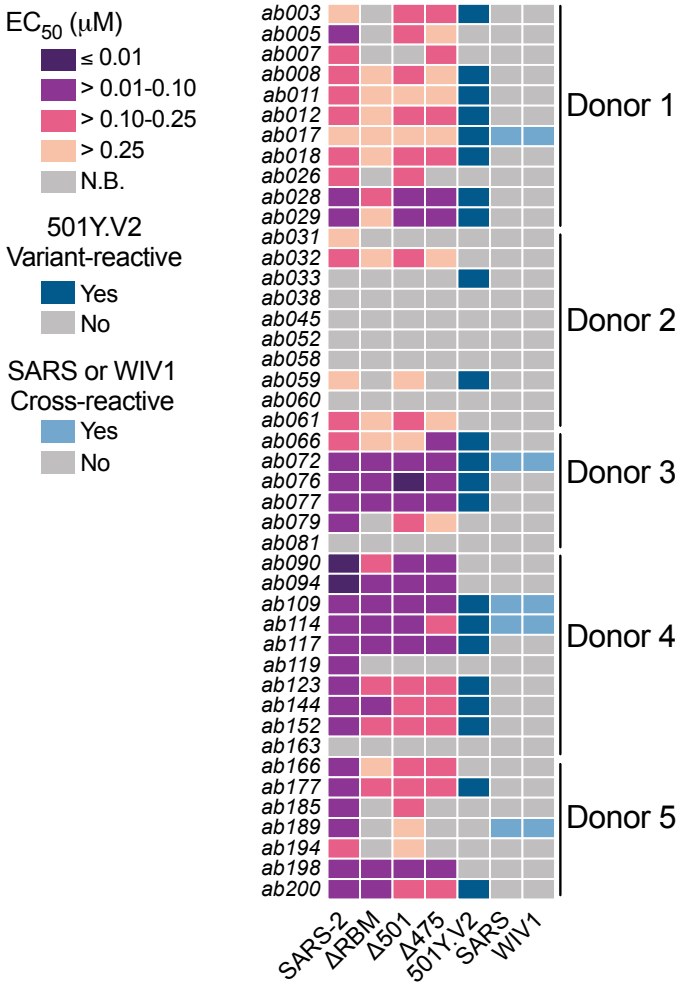
644

645

646 **fig. S5**. **Repertoire comparison and germline identity by individual donor**. (**A**) Heatmap
647 showing V_H-gene usage of isolated antibodies derived from donors 1-5. Unselected repertoire gene
648 usage derive from a high-throughput sequencing data set of circulating B cells across 10 human
649 subjects (*1*). Heatmap scale represent percent of total paired sequence from each donor.
650 Divergence from inferred germline gene sequences separated by individual donor for (**B**) V_H and
651 (**C**) V_L. Red bars indicate the median percent values, and each dot represents an individual paired
652 sequence.

653

654



655

656

657

658 **fig. S6. ELISA specificity for isolated naive antibodies related to Figure 2**. Heatmap showing

659 IgG binding to RBDs (n = 44) sorted by donor.

660

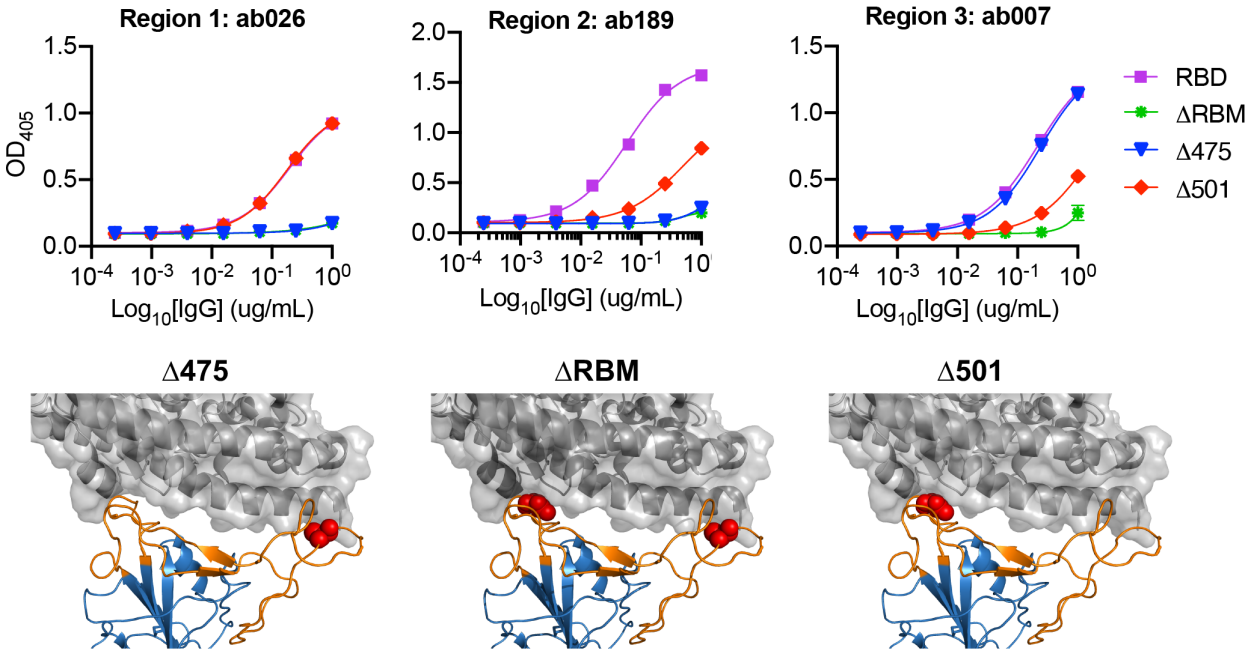
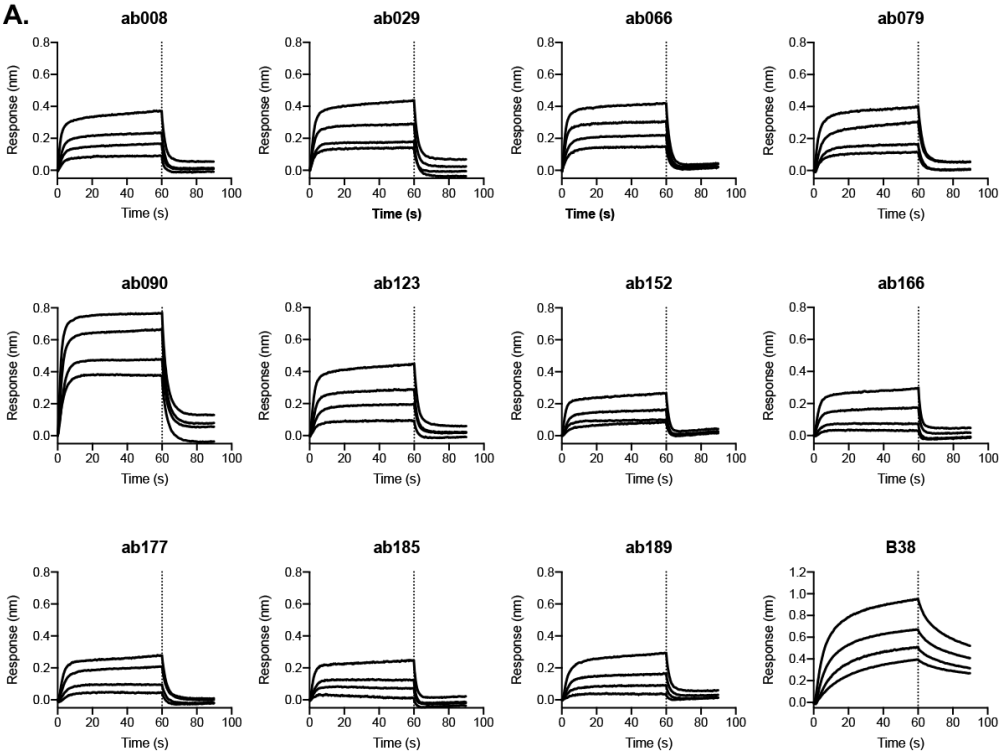


fig. S7. RBD glycan variant binding by ELISA. Representative examples of epitopic regions as defined by patterns of RBD glycan variant binding by ELISA. Region 1 refers to IgGs with reduced binding to Δ475, region 2 refers to IgGs with reduced binding to both glycan variants, and region 3 refers to IgGs with reduced binding to Δ501.

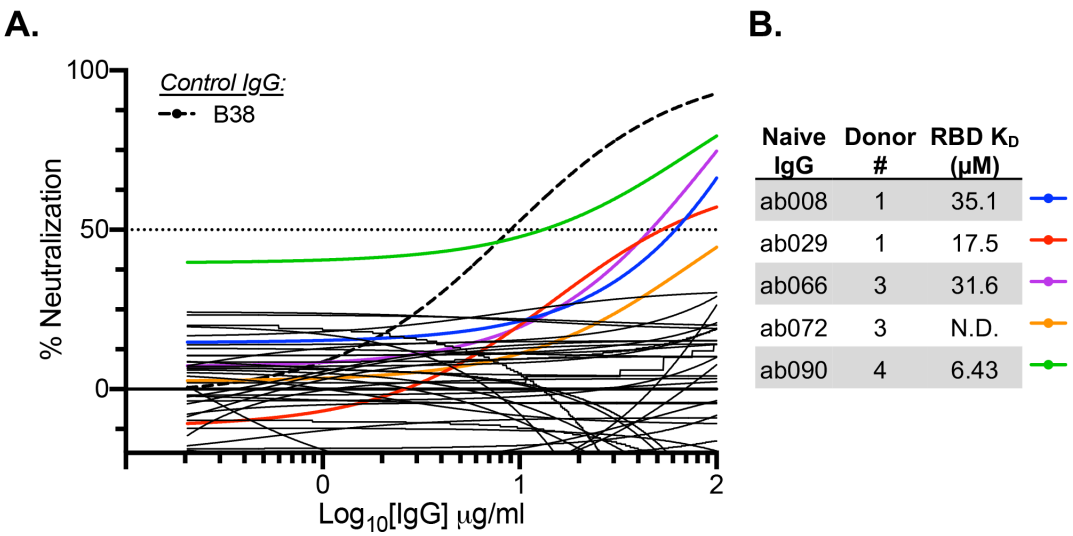
A.



B.

Naive IgG	IGHV	IGLV	K_D (μM)	K_a (10^3) ($M^{-1}s^{-1}$)	K_{off} (10^{-1}) (s^{-1})
B38	*IGHV3-53*	*IGKV1-9*	*2.1*	*10.21*	*0.21*
ab090	IGHV1-2	IGKV3-15	6.4	31.27	2.01
ab029	IGHV5-51	IGLV2-14	17.5	13.91	2.44
ab066	IGHV1-69	IGKV4-1	31.6	9.23	2.92
ab008	IGHV1-69	IGKV3-20	35.1	7.21	2.53
ab123	IGHV3-9	IGLV3-21	40.2	5.86	2.36
ab079	IGHV4-31	IGLV3-21	42.6	4.30	1.83
ab152	IGHV5-51	IGKV2-30	45.1	6.39	2.88
ab166	IGHV1-69	IGLV1-40	55	4.52	2.49
ab189	IGHV3-30-3	IGLV3-21	73.7	2.37	1.75
ab177	IGHV3-21	IGLV2-14	>100	0.06	2.39
ab185	IGHV3-23	IGLV2-23	>100	0.13	3.79

fig. S8. SARS-CoV-2 RBD-binding kinetics of isolated naive antibodies. (**A**) Biolayer interferometry (BLI) binding kinetic analysis of titrated SARS-CoV-2 RBD to immobilized Fabs. Dotted line at 60 s denotes the start of the dissociation phase. (**B**) Kinetic and equilibrium constants for binding to RBD calculated from a 1:1 binding model using a global fit to all curves for each Fab using vendor supplied software. B38 Fab is used as a positive control.



678

679

680 **fig. S9. SARS-CoV-2 pseudotyped neutralization assay**. (**A**) Neutralization data for 44 purified
681 IgGs. Curves in color highlighted antibodies with neutralizing activity with (**B**) donor and
682 monovalent RBD affinity listed for this subset of antibodies. Dashed lines indicate IC_{50} values and
683 data represent means ± SD of two technical replicates.

684

685 **SI REFERENCES**

686

687 1. B. Briney, A. Inderbitzin, C. Joyce, D. R. Burton, Commonality despite exceptional diversity
688 in the baseline human antibody repertoire. *Nature* **566**, 393–397 (2019).
689

690