Convergent Antibody Responses to SARS-CoV-2 Infection in Convalescent Individuals

Davide F. Robbiani[1*,#], Christian Gaebler[1*], Frauke Muecksch[2*], Julio C. C. Lorenzi[1*], Zijun Wang[1*], Alice Cho[1*], Marianna Agudelo[1*], Christopher O. Barnes[6*], Shlomo Finkin[1*], Thomas Hagglof[1*], Thiago Y. Oliveira[1*], Charlotte Viant[1], Arlene Hurley[4], Katrina G. Millard[1], Rhonda G. Kost[5], Melissa Cipolla[1], Anna Gazumyan[1], Kristie Gordon[1], Filippo Bianchini[1], Spencer T. Chen[1], Victor Ramos[1], Roshni Patel[1], Juan Dizon[1], Irina Shimeliovich[1], Pilar Mendoza[1], Harald Hartweger[1], Lilian Nogueira[1], Maggi Pack[1], Jill Horowitz[1], Fabian Schmidt[2], Yiska Weisblum[2], Hans-Heinrich Hoffmann[3], Eleftherios Michailidis[3], Alison W. Ashbrook[3], Eric Waltari[7], John E. Pak[7], Kathryn E. Huey-Tubman[6], Nicholas Koranda[6], Pauline R. Hoffman[6], Anthony P. West, Jr.[6], Charles M. Rice[3], Theodora Hatziioannou[2], Pamela J. Bjorkman[6], Paul D. Bieniasz[2,8,#], Marina Caskey[1,#], Michel C. Nussenzweig[1,8,#]

[1]Laboratory of Molecular Immunology, The Rockefeller University, New York, NY 10065, USA
[2]Laboratory of Retrovirology, The Rockefeller University, New York, NY 10065, USA
[3]Laboratory of Virology and Infectious Disease, The Rockefeller University, New York, NY 10065, USA
[4]Hospital Program Direction, The Rockefeller University, New York, NY 10065, USA
[5]Hospital Clinical Research Office, The Rockefeller University, New York, NY 10065, USA
[6]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA
[7]Chan Zuckerberg Biohub, 499 Illinois Street, San Francisco, CA 94158, USA
[8]Howard Hughes Medical Institute
[*]Equal contribution
[#]Send correspondence to Paul D. Bieniasz: pbieniasz@rockefeller.edu, Marina Caskey: mcaskey@rockefeller.edu, Michel C. Nussenzweig: nussen@rockefeller.edu, or Davide F. Robbiani: drobbiani@rockefeller.edu

ABSTRACT

During the COVID-19 pandemic, SARS-CoV-2 infected millions of people and claimed hundreds of thousands of lives. Virus entry into cells depends on the receptor binding domain (RBD) of the SARS-CoV-2 spike protein (S). Although there is no vaccine, it is likely that antibodies will be essential for protection. However, little is known about the human antibody response to SARS-CoV-2[1-5]. Here we report on 68 COVID-19 convalescent individuals. Plasmas collected an average of 30 days after the onset of symptoms had variable half-maximal neutralizing titers ranging from undetectable in 18% to below 1:1000 in 78%, while only 3% showed titers >1:5000. Antibody cloning revealed expanded clones of RBD-specific memory B cells expressing closely related antibodies in different individuals. Despite low plasma titers, antibodies to distinct epitopes on RBD neutralized at half-maximal inhibitory concentrations (IC_{50}s) as low as few ng/mL. Thus, most convalescent plasmas obtained from individuals who recover from COVID-19 without hospitalization do not contain high levels of neutralizing activity. Nevertheless, rare but recurring RBD-specific antibodies with potent antiviral activity were found in all individuals tested, suggesting that a vaccine designed to elicit such antibodies could be broadly effective.

Between April 1 and April 17, 2020, 73 eligible participants enrolled in the study. Of these, 48 (65.8%) were individuals diagnosed with SARS-CoV-2 infection by RT-PCR (cases), and 25 (34.2%) were close contacts of individuals diagnosed with SARS-CoV-2 infection (contacts). While inclusion criteria allowed for enrollment of asymptomatic participants, 5 close contacts that did not develop symptoms were excluded from further analyses. The 68 cases and contacts were free of symptoms suggestive of COVID-19 for at least 14 days at the time of sample collection. Participant demographics and clinical characteristics are shown in Table 1 and Extended Data Tables 1 and 2. Only one individual who tested positive for SARS-CoV-2 infection by RT-PCR remained asymptomatic. The other 67 participants reported symptoms suggestive of COVID-19 with an average onset of approximately 30 days (range 17 to 48 days) before sample collection. In this cohort, symptoms lasted for an average of 10 days (0-28 days), and none of the participants were hospitalized. The most common symptoms were fever (82.4%), cough (64.7%), myalgia (55.9%) and fatigue (54.4%) while baseline comorbidities were infrequent (8.8%) (Table 1 and Extended Data Tables 1 and 2). There were no significant differences in duration or severity of symptoms, or in time from onset of symptoms to sample collection between genders or between cases and contacts (Extended Data Figure 1).

Plasma samples were tested for binding to the SARS-CoV-2 RBD and trimeric spike (S) proteins by ELISA using anti-IgG or -IgM secondary antibodies for detection (Fig. 1, Extended Data Table 1 and Extended Data Figs. 2 and 3). Three independent negative controls and the plasma sample from participant 21 (COV21) were included for normalization of the area under the curve (AUC). Overall, 88% and 66% of the plasma samples tested showed anti-RBD IgG and IgM AUCs that were at least 2 standard deviations above the control (Fig. 1 a, b). In contrast, only 40% and 21%

of the plasma samples showed anti-S IgG and IgM responses that were at least 2 standard deviations above control (Fig. 1 c, d). There was no correlation between IgG or IgM levels and age, the timing of sample collection relative to onset of symptoms, and no significant difference between cases and contacts (Fig. 1g, h, and Extended Data Figs. 2 a-c and 3 c-i). In contrast, IgG binding to RBD and S were directly correlated to duration of symptoms, but IgM binding was not (Fig. 1e, f, and Extended Data Fig. 3). Finally, antibody levels to S were similar between genders, but females had lower titers of IgG-binding antibodies to RBD than males (Extended Data Fig. 2d).

To measure the neutralizing activity in convalescent plasmas we used pseudotyped virus assays that employed HIV-1-based virions carrying a nanoluc luciferase reporter and the SARS-CoV-2 spike (pSARS-CoV2-S_{trunc} see Methods, Fig. 2 and Extended Data Fig. 4). The overall level of neutralizing activity in the cohort, as measured by the half-maximal neutralizing titer (NT_{50}) was generally low, with 18% undetectable and 78% below 1,000 (Fig. 2a, b). The geometric mean NT_{50} was 212 (arithmetic mean = 850), and only 2 individuals reached NT_{50}s above 5,000 (Fig. 2a, b, Extended Data Table 1). Neutralizing activity correlated with the duration of symptoms and symptom severity but did not correlate with the timing of sample collection relative to onset of symptoms, age, sex or case/contact status (Fig. 2c, e, f and Extended Data Fig. 5). Notably, levels of RBD- and S-binding IgG antibodies correlated strongly with NT_{50} (Fig. 2d and Extended Data Fig. 5e).

To determine the nature of the antibodies elicited by SARS-CoV-2 infection we used flow cytometry to isolate individual B lymphocytes with receptors that bound to RBD from the blood

of 6 individuals including the 2 top neutralizers (Fig. 3). The frequency of antigen-specific B cells, identified by their ability to bind to both Phycoerythrin (PE)- and BV711-labeled RBD, ranged from 0.07 to 0.005% of circulating B cells in COVID-19 convalescents but they were undetectable in controls (Fig.3a and Extended Data Fig.6). We obtained 534 paired IgG heavy and light chain (IGH and IGL) sequences by reverse transcription and subsequent PCR from individual RBD-binding B cells from the 6 convalescent individuals (see Methods). When compared to the human antibody repertoire, several IGH and IGL genes were significantly over-represented (Extended Data Fig. 7). The average number of V genes nucleotide mutations for IGH and IGL was 4.2 and 2.8, respectively (Extended Data Fig. 8), which is lower than in antibodies cloned from individuals suffering from chronic infections such as Hepatitis B or HIV-1, and similar to antibodies derived from primary malaria infection or non-antigen-enriched circulating IgG memory cells[6-8] (Wang et al, in press, https://www.biorxiv.org/content/10.1101/2020.03.04.976159v1.full). Among other antibody features, IGH CDR3 length was indistinguishable from the reported norm and hydrophobicity was below average (Extended Data Fig. S8)[9].

As is the case with other human pathogens, there were expanded clones of antigen binding B cells in all COVID-19 individuals tested (see Methods and Fig. 3b, c). Overall, 32.2% of the recovered IGH and IGL sequences were from clonally expanded B cells (range 21.8-57.4% across individuals, Fig. 3b). Antibodies that shared specific combinations of IGHV and IGLV genes in different individuals comprised 14% of all the clonal sequences (colored pie slices in Fig. 3b, c). Remarkably, the amino acid sequences of some antibodies found in different individuals were nearly identical. For example, clonal antibodies with IGHV1-58/IGKV3-20 and IGHV3-30-3/IGKV1-39 found repeatedly in different individuals had amino acid sequence identities of up to

99% and 92%. We conclude that the IgG memory response to the SARS-CoV-2 RBD is highly enriched in recurrent clonally expanded antibody sequences.

To examine the binding properties of anti-SARS-CoV-2 antibodies, we expressed 34 representative antibodies, 24 from clones and 10 from singlets found in 3 individuals. ELISA assays showed that 94% (32 out of 34) of the antibodies bound to the SARS-CoV-2 RBD with an average half-maximal effective concentration (EC_{50}) of 6.6 ng/mL (Fig 4a, b). To determine whether these antibodies have neutralizing activity, we tested them against the SARS-CoV2-S_{trunc} pseudovirus (Fig. 4c, d and Extended Data Table 3). Among 32 RBD binding antibodies tested, we found 20 that neutralized with nanogram per milliliter half-maximal inhibitory concentrations (IC_{50}s) ranging from 4.4 to 709 ng/mL (Fig. 4c, d and Extended Data Table 3). Potent neutralizing antibodies were found in individuals irrespective of their plasma NT_{50}s. For example, C002 and C121 with IC_{50}s of 8.9 and 6.7 ng/mL, respectively, were obtained from individuals COV21 and COV107 whose plasma NT_{50} values were of 5053 and 298 respectively (Fig. 2b and 4). Finally, clones of antibodies with shared IGHV and IGLV genes were among the best neutralizers, e.g., antibody C002 with IGHV3-30/IGKV1-39 that is shared by the 2 donors with the best plasma neutralizing activity (red pie slice in Fig. 3b and Fig. 4). We conclude that even individuals with modest plasma neutralizing activity harbor rare IgG memory B cells that produce potent SARS-CoV-2 neutralizing antibodies.

To determine whether human anti-SARS-CoV-2 monoclonal antibodies with neutralizing activity can bind to distinct domains on the RBD we performed bilayer interferometry experiments in which a preformed antibody-RBD immune complex was exposed to a second monoclonal. The

antibodies tested comprised 2 groups, whereas, C002, and C105 bound to a pre-formed C121-RBD complex, C104, C110 and C119 did not (Fig 4e and f). We conclude that like SARS-CoV-1[10], there are at least 2 distinct neutralizing epitopes on the RBD of SARS-CoV-2.

Human monoclonal antibodies with neutralizing activity against pathogens ranging from viruses to parasites have been obtained from naturally infected individuals by single cell antibody cloning. Several have been shown to be effective in protection and therapy in model organisms and in early phase clinical studies, but only one antiviral monoclonal is currently in clinical use[11]. Antibodies are relatively expensive and more difficult to produce than small molecule drugs. However, they differ from drugs in that they can engage the host immune system through their constant domains that bind to Fc gamma receptors on host immune cells[12]. These interactions can enhance immunity and help clear the pathogen or infected cells, but they can also lead to disease enhancement during Dengue[13] and possibly coronavirus infections[14]. This problem has impeded Dengue vaccine development but would not interfere with the clinical use of potent neutralizing antibodies that can be modified to prevent Fc gamma receptor interactions and remain protective against viral pathogens[15].

Antibodies are essential elements of most vaccines and will likely be crucial component of an effective vaccine against SARS-CoV-2[16]. The observation that plasma neutralizing activity is low in most convalescent individuals, but that recurrent anti-SARS-CoV-2 RBD antibodies with potent neutralizing activity can be found in individuals with unexceptional plasma neutralizing activity suggests that humans are intrinsically capable of generating anti-RBD antibodies that potently

neutralize SARS-CoV-2. Thus, vaccines that specifically and efficiently induce antibodies targeting the SARS-CoV-2 RBD may be especially effective.

Table

Table 1. Cohort characteristics

Gender	n	Average age	Case/Contact	Average duration Sx total	Average duration Sx onset to visit	Average Sx Severity (0-10)	ELISA binding (AUC) RBD IgG	ELISA binding (AUC) RBD IgM	ELISA binding (AUC) S IgG	ELISA binding (AUC) S IgM	Neutralization (NT50)
Male	37	42 (21-68)	27/10	10 (0-21)	31 (21-42)	5 (0-9)	2.31	2.18	4.72	1.52	880
Female	31	41 (19-75)	21/10	9 (1-28)	30 (17-48)	4.4 (2-9)	1.95	2.10	4.62	1.43	814

Figures

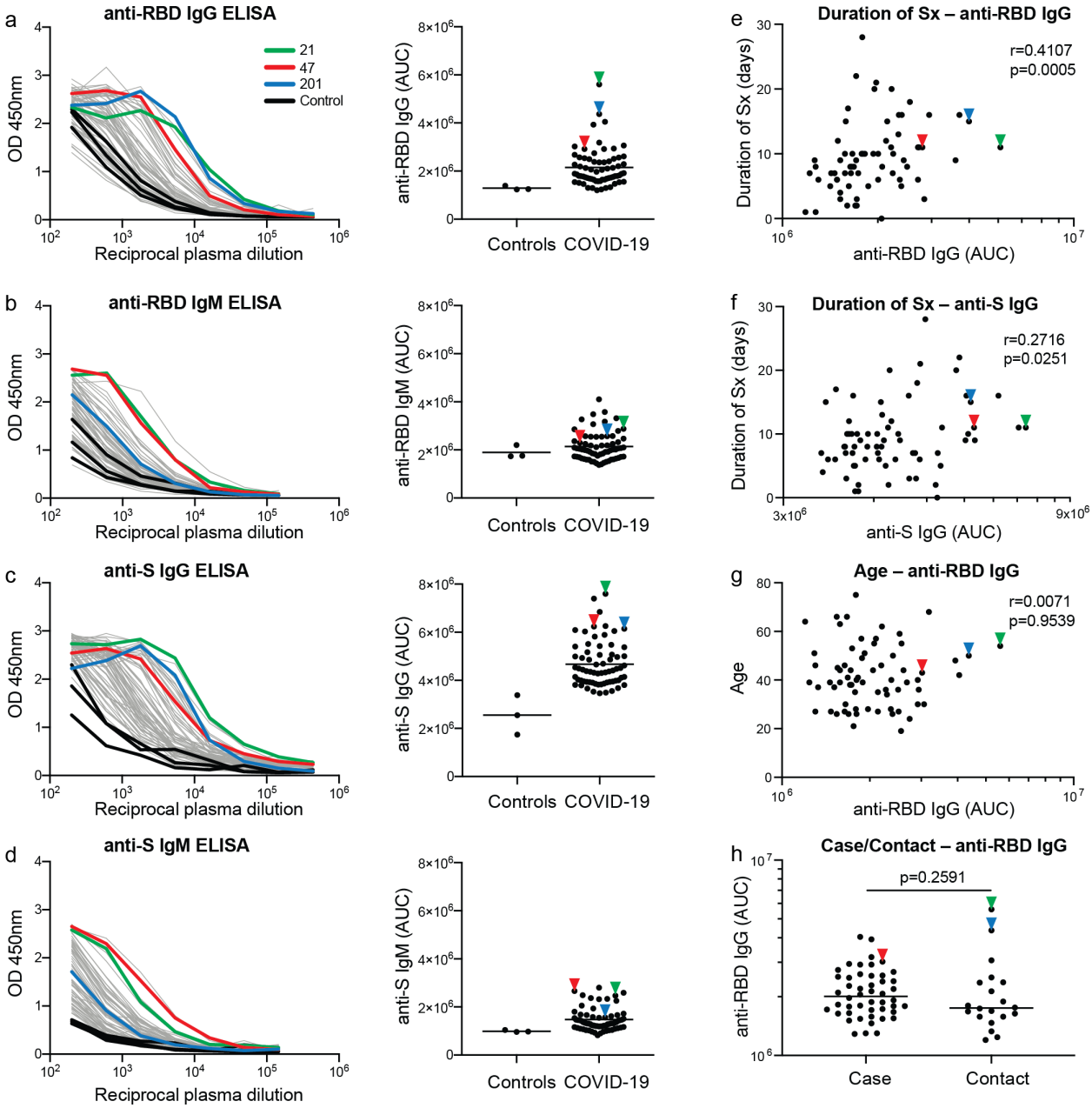


Figure 1. Plasma antibodies against SARS-CoV-2. a-d, Graphs show results of ELISAs measuring plasma reactivity to RBD (**a, b**) and S protein (**c, d**). Left shows optical density units at 450 nm (OD, Y axis) and reciprocal plasma dilutions (X axis). Negative control in black; individuals 21, 47, and 201 in green, red, and blue lines and arrowheads, respectively. Right shows normalized area under the curve (AUC) for controls and each of 68 individuals in the cohort. **e**,

Duration of symptoms (Sx) in days (Y axis) plotted against normalized AUC for IgG binding to RBD (X axis) r=0.4107 and p=0.0005. **f**, Duration of symptoms in days (Y axis) plotted against normalized AUC for IgG binding to S (X axis) r=0.2716 and p=0.0251. **g,** Age (Y axis) plotted against normalized AUC for IgG binding to RBD (X axis) r=0.0071 p=9539. The r and p values for the correlations in e-g were determined by two-tailed Spearman's. **h,** Normalized AUC for IgG binding to RBD (Y axis) for all RT-PCR confirmed cases and close contacts in the cohort. Horizontal bars indicate median values. Statistical significance was determined using two-tailed Mann-Whitney U test (p=0.2591).

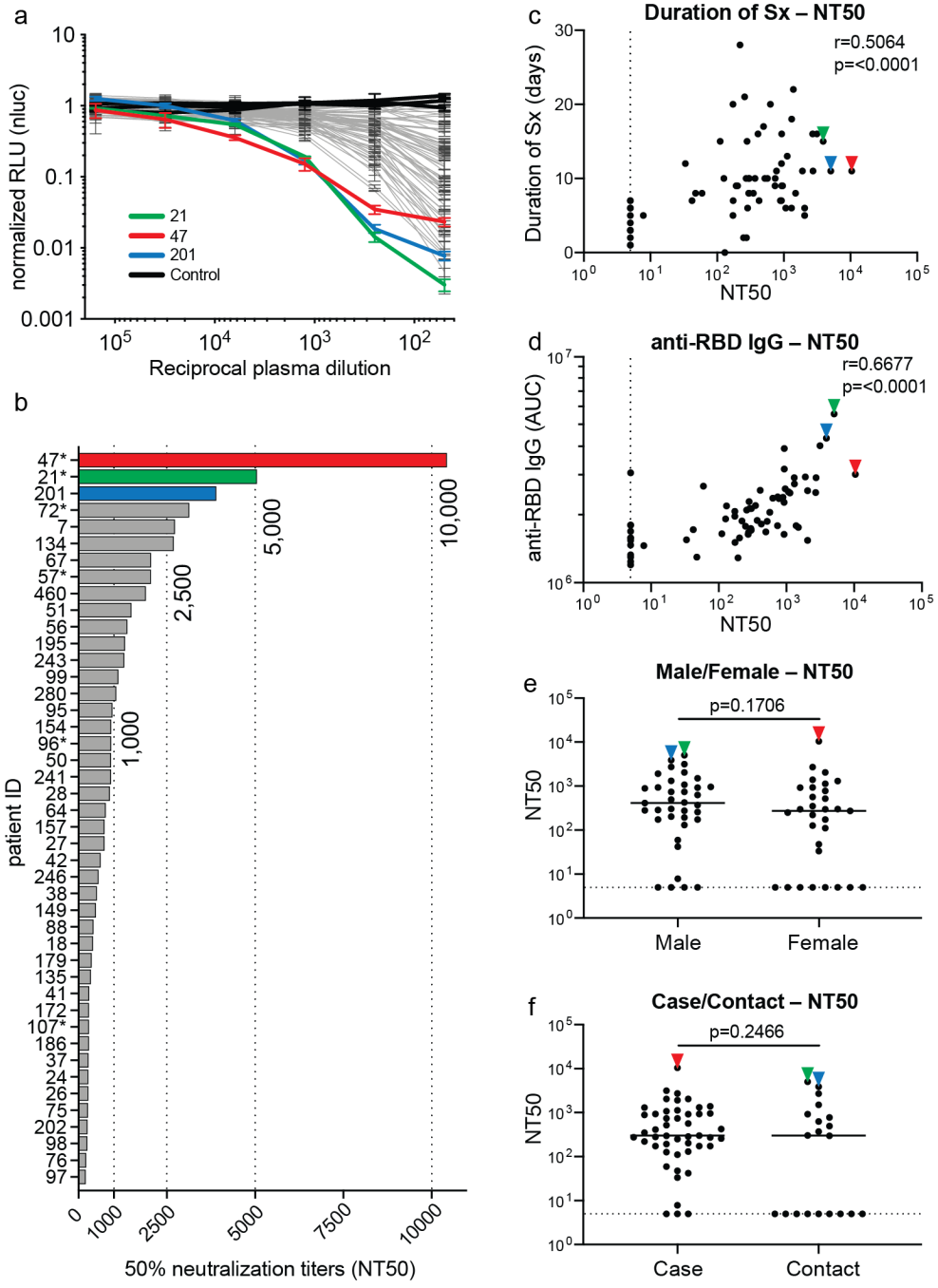


Figure 2. Neutralization of SARS-CoV-2 pseudovirus by plasma. **a,** Graph shows normalized relative luminescence values (RLU, Y axis) in cell lysates of 293T-ACE2 cells 48 hours after infection with nanoluc-expressing SARS-CoV-2 pseudovirus in the presence of increasing concentrations of plasma (X axis) derived from 68 participants (grey, except individuals 21, 47 and 201 in green, red, and blue lines, bars and arrowheads, respectively) and 3 negative controls

(black lines). Standard deviations of duplicates of one representative experiment are shown. **b,** Average half-maximal inhibitory plasma neutralizing titer (NT_{50}) for the 44 of 68 individuals with NT_{50}s >200. See also Extended Data Table 1. **c,** Duration of symptoms in days (Y axis) plotted against NT_{50} (X axis) r=0.5064, p=<0.0001. **d,** AUC for anti-RBD IgG ELISA (Y axis) plotted against NT_{50} (X axis) r=0.6677, p=<0.0001. **e,** NT_{50} (Y axis) for all males and females in the cohort p=0.1706**. f,** NT_{50} (Y axis) for all cases and contacts in the cohort p=0.2466. Dotted line (NT_{50}=5) represents lower limit of detection (LLOD). Samples with undetectable neutralizing titers were plotted at LLOD. Correlations in c and d were determined by two-tailed Spearman's. Horizontal bars indicate median values. Statistical significance in e and f was determined using two-tailed Mann-Whitney U test.

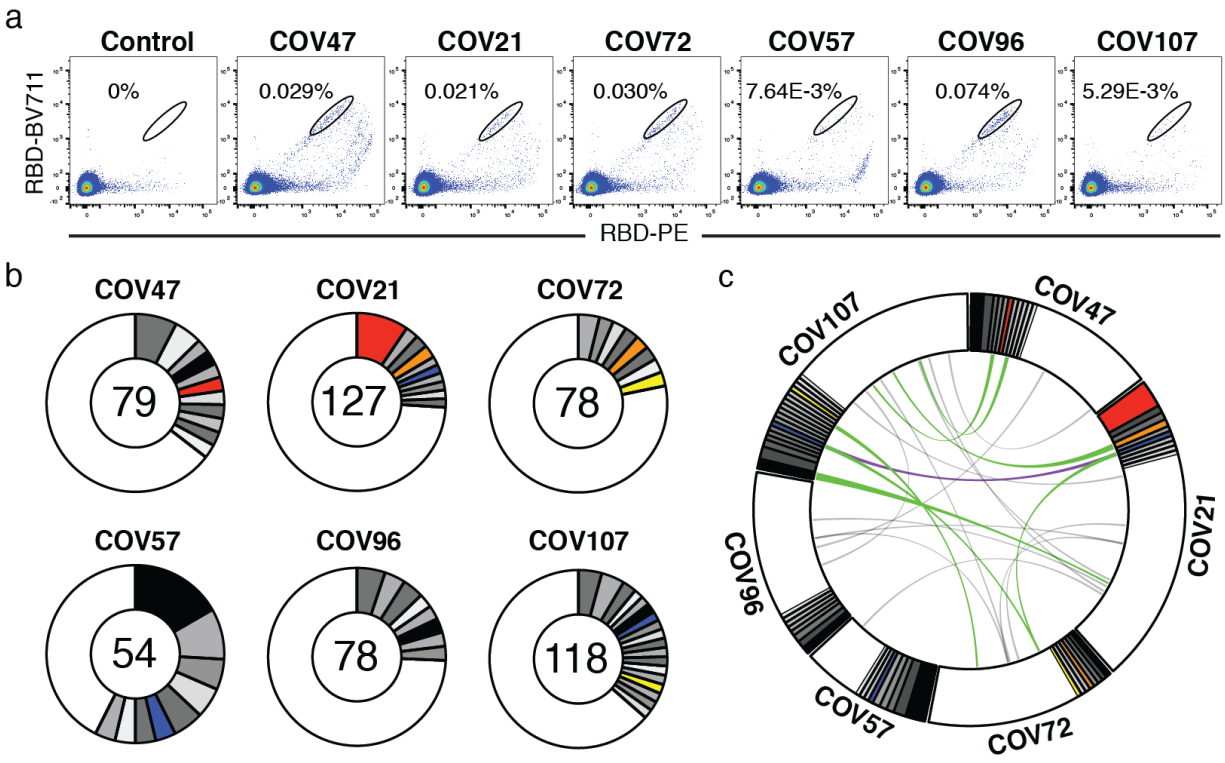


Figure 3. Anti-SARS-CoV-2 RBD antibodies. a. Representative flow cytometry plots showing dual BV711- and PE-RBD binding B cells in control and 6 study individuals (for gating strategy see Extended Data Fig. 6). Percentages of antigen specific B cells are indicated. **b,** Pie charts depicting the distribution of antibody sequences from 6 individuals. The number in the inner circle indicates the number of sequences analyzed for the individual denoted above the circle. White indicates sequences isolated only once, and grey or colored pie slices are proportional to the number of clonally related sequences. Red, blue, orange and yellow pie slices indicate clones that share the same IGHV and IGLV genes. **c,** Circos plot shows sequences from all 6 individuals with clonal relationships depicted as in **b**. Interconnecting lines indicate the relationship between antibodies that share V and J gene segment sequences at both IGH and IGL. Purple, green and gray lines connect related clones, clones and singles, and singles to each other, respectively.

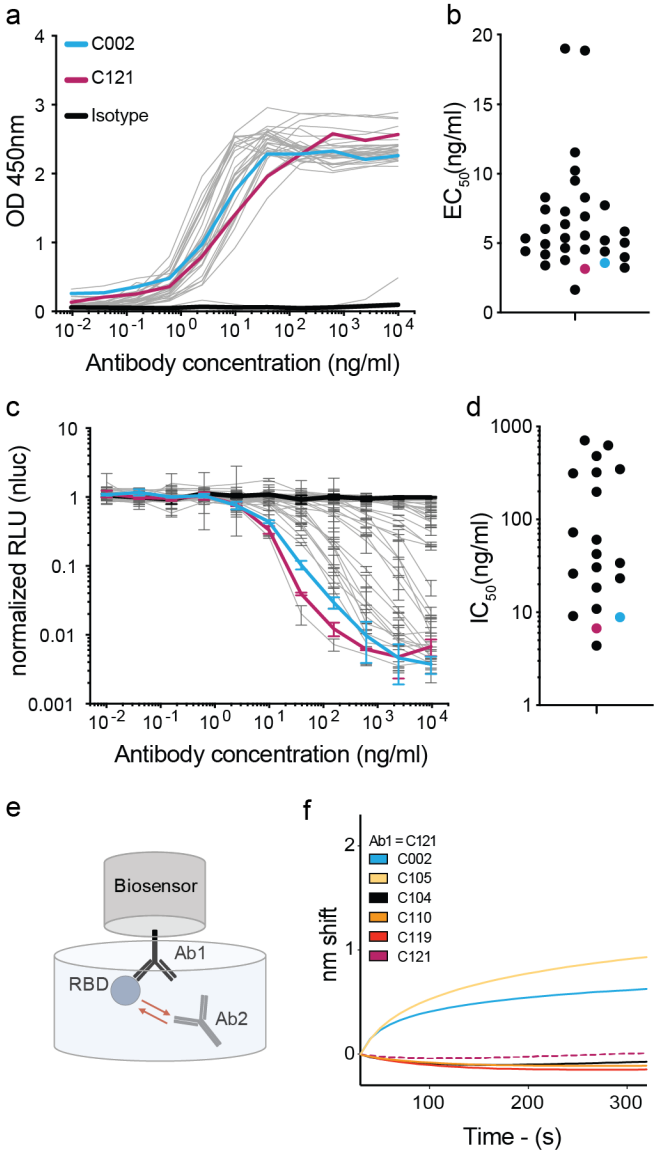


Figure 4. Anti-SARS-CoV-2 RBD antibody reactivity. a, Graph show results of ELISA assays measuring monoclonal antibody reactivity to RBD. Left shows optical density units at 450 nm (OD, Y axis) and reciprocal dilutions (X axis). Representative experiment. **b,** Graph shows mean EC_{50}s of at least two independent experiments. **c,** Graph shows normalized relative luminescence values (RLU, Y axis) in cell lysates of $293T_{ACE2}$ cells 48 hours after infection with SARS-CoV-2 pseudovirus in the presence of increasing concentrations of monoclonal antibodies (X axis). Representative experiment. **d**. IC_{50}s of at least two independent experiments (see Extended Data

Table 3). For a and c panels, isotype control antibody in black, C002 blue, C121 red. **e,** Diagrammatic representation of Biolayer Interferometry competition assay. **f,** Graph shows binding of C002 and C105 but not C104, C110, C119 to preformed C121-RBD immune complexes.

Methods

Study participants. Study participants were recruited at the Rockefeller University Hospital in New York from April 1 through April 17, 2020. Eligible participants were adults aged 18-75 years who were either diagnosed with SARS-CoV-2 infection by RT-PCR and were free of symptoms of COVID-19 for at least 14 days (cases), or who were close contacts (e.g., household, co-workers, members of same religious community) with someone who had been diagnosed with SARS-CoV-2 infection by RT-PCR and were free of symptoms suggestive of COVID-19 for at least 14 days (contacts). Exclusion criteria included presence of symptoms suggestive of active SARS-CoV-2 infection, or hemoglobin < 12 g/dL for males and < 11 g/dL for females.

Study participants were residents of the Greater New York City tri-state region and were enrolled sequentially according to eligibility criteria. Participants were first interviewed by phone to collect information on their clinical presentation, and subsequently presented to the Rockefeller University Hospital for a single blood sample collection. Participants were asked to rate the highest severity of their symptoms on a numeric rating scale ranging from 0 to 10. The score was adapted from the pain scale chart, where 0 was the lack of symptoms, 4 were distressing symptoms (e.g. fatigue, myalgia, fever, cough, shortness of breath) that interfered with daily living activities, 7 were disabling symptoms that prevented the performance of daily living activities, and 10 was unimaginable/unspeakable discomfort (in this case, distress due to shortness of breath). All participants provided written informed consent before participation in the study and the study was conducted in accordance with Good Clinical Practice.

Blood samples processing and storage. Peripheral Blood Mononuclear Cells (PBMCs) were obtained by gradient centrifugation and stored in liquid nitrogen in the presence of FCS and

DMSO. Plasma and serum samples were aliquoted and stored at -20C or less. Prior to experiments, aliquots of plasma samples were heat-inactivated (56C for 1 hour) and then stored at 4C.

Cloning, expression and purification of recombinant SARS-CoV-2 proteins. Codon-optimized nucleotide sequences encoding the SARS-CoV-2 S ectodomain (residues 16-1206) and receptor binding domain (RBD; residues 331-524) were synthesized and subcloned into the mammalian expression pTwist-CMV BetaGlobin vector by Twist Bioscience Technologies based on an early SARS-CoV-2 sequence isolate (GenBank MN985325.1). The SARS-CoV-2 RBD construct included an N-terminal human IL-2 signal peptide and dual C-terminal tags ($(GGGGS)_2$-HHHHHHHH (octa-histidine), and GLNDIFEAQKIEWHE (AviTag)). The SARS-CoV-2 S ectodomain was modified as previously described [4]. Briefly, the S ectodomain construct included an N-terminal mu-phosphatase signal peptide, 2P stabilizing mutations (K986P and V987P), mutations to remove the S1/S2 furin cleavage site ($_{682}$RRAR$_{685}$ to GSAS), a C-terminal extension (IKGSG-RENLYFQG (TEV protease site), GGGSG-YIPEAPRDGQAYVRKDGEWVLLSTFL (foldon trimerization motif), G-HHHHHHHH (octa-histidine tag), and GLNDIFEAQKIEWHE (AviTag)). The SARS-CoV-2 S 2P ectodomain and RBD constructs were produced by transient transfection of 500 mL of Expi293 cells (Thermo Fisher) and purified from clarified transfected cell supernatants four days post-transfection using Ni^{2+}-NTA affinity chromatography (GE Life Sciences). Affinity-purified proteins were concentrated and further purified by size-exclusion chromatography (SEC) using a Superdex200 16/60 column (GE Life Sciences) running in 1x TBS (20 mM Tris-HCl pH 8.0, 150 mM NaCl, and 0.02% NaN_3). Peak fractions were analyzed by SDS-PAGE, and fractions corresponding to soluble S 2P trimers or monomeric RBD proteins were pooled and stored at 4˚C.

ELISAs. ELISAs to evaluate antibodies binding to SARS-CoV-2 RBD and trimeric spike proteins were performed by coating of high binding 96 half well plates (Corning #3690) with 50 µL per well of a 1µg/mL protein solution in PBS overnight at 4°C. Plates were washed 6 times with washing buffer (1xPBS with 0.05% Tween 20 (Sigma-Aldrich)) and incubated with 170 µL per well blocking buffer (1xPBS with 2% BSA and 0.05% Tween20 (Sigma)) for 1 hour at room temperature (RT). Immediately after blocking, monoclonal antibodies or plasma samples were added in PBS and incubated for 1 hr at RT. Plasma samples were assayed at a 1:200 starting dilution and seven additional 3-fold serial dilutions. Monoclonal antibodies were tested at 10µg/ml starting concentration and 10 additional 4-fold serial dilutions. Plates were washed 6 times with washing buffer and then incubated with anti-human IgG or IgM secondary antibody conjugated to horseradish peroxidase (HRP) (Jackson Laboratories) in blocking buffer at a 1:5000 dilution. Plates were developed by addition of the HRP substrate, TMB (ThermoFisher) for 10 minutes, then the developing reaction was stopped by adding 50µl 1M H_2SO_4 and absorbance was measured at 450nm with an ELISA microplate reader (FluoStar Omega, BMG Labtech). For plasma and monoclonal antibodies, a positive control plasma sample (COV21) was added in duplicate to every individual assay plate for normalization. The positive sample optical density measured was set to 100 and AUC measurements were normalized according to the factor defined for each plate's positive control. For monoclonal antibodies, the half-maximal effective concentration (EC_{50}) was determined using 4-parameter nonlinear regression (GraphPad Prism).

$293T_{ACE2}$ cells. For constitutive expression of ACE2 in 293T cells, a cDNA encoding ACE2, carrying two inactivating mutations in the catalytic site (H374N & H378N), was inserted into CSIB

3' to the SFFV promoter[18]. 293T$_{ACE2}$ cells were generated by transduction with CSIB based virus followed by selection with 5 µg/ml Blasticidin.

SARS-CoV-2 pseudotyped reporter virus. A plasmid expressing a C-terminally truncated SARS-CoV-2 S protein (pSARS-CoV2-S$_{trunc}$) was generated by insertion of a human-codon optimized cDNA encoding SARS-CoV-2 S lacking the C-terminal 19 codons (Geneart) into pCR3.1. An *env*-inactivated HIV-1 reporter construct (pNL4-3ΔEnv-NanoLuc) was generated from pNL4-3 [19] by introducing a 940 bp deletion 3' to the *vpu* stop-codon, resulting in a frameshift in *env*. The human codon-optimized NanoLuc Luciferase reporter gene (*Nluc,* Promega) was inserted in place of nucleotides 1-100 of the *nef*-gene. To generate pseudotyped viral stocks, 293T cells were transfected with pNL4-3ΔEnv-NanoLuc and pSARS-CoV2-S$_{trunc}$ using polyethyleneimine. Co-transfection of pNL4-3ΔEnv-NanoLuc and pSARS-CoV2-S$_{trunc}$ leads to production of HIV-1-based virions carrying the SARS-CoV-2 Spike protein on the surface. Eight hours after transfection, cells were washed twice with PBS and fresh media was added. Supernatants containing virions were harvested 48 hours post transfection, filtered and stored at -80 °C. Infectivity of virions was determined by titration on 293T$_{ACE2}$ cells.

Pseudotyped virus neutralization assay. Five-fold serially diluted plasma from COVID-19 convalescent individuals and healthy donors was incubated with the SARS-CoV-2 pseudotyped virus for 1 hour at 37 degrees. The mixture was subsequently incubated with 293T$_{ACE2}$ cells for 48 hours after which cells were washed twice with PBS and lysed with Luciferase Cell Culture Lysis 5x reagent (Promega). NanoLuc Luciferase activity in lysates was measured using the Nano-Glo Luciferase Assay System (Promega). Relative luminescence units obtained were normalized to

those derived from cells infected with SARS-CoV-2 pseudotyped virus in the absence of plasma. The half-maximal inhibitory concentration for plasma (NT_{50}) or monoclonal antibodies (IC_{50}) was determined using 4-parameter nonlinear regression (GraphPad Prism). For experiments testing antibody combinations, equal amounts of each antibody were added at the same concentration of when tested alone (for example, if antibody A and B alone are tested at 10ng/mL, the combination is 10ng/mL of each, resulting in 20ng/mL total concentration and this is the concentration that is plotted).

Biotinylation of viral protein for use in flow cytometry. Purified and Avi-tagged SARS-CoV-2 RBD was biotinylated using the Biotin-Protein Ligase-BIRA kit according to manufacturer's instructions (Avidity). Ovalbumin (Sigma, A5503-1G) was biotinylated using the EZ-Link Sulfo-NHS-LC-Biotinylation kit according to the manufacturer's instructions (Thermo Scientific). Biotinylated Ovalbumin was conjugated to streptavidin Alexa Fluor 647 (Biolegend, 405237) and RBD to streptavidin PE (BD biosciences, 554061) and streptavidin BV711 (BD biosciences, 563262) respectively [20].

Single cell sorting by flow cytometry. PBMCs were enriched for B cells by negative selection using a pan B cell isolation kit according to the manufacturer's instructions (Miltenyi Biotec, 130-101-638). The enriched B cells were incubated in FACS buffer (1 X Phosphate-buffered Saline (PBS), 2% calf serum, 1 mM EDTA) with the following anti-human antibodies: anti-CD20-PECy7 (BD Biosciences, 335793), anti-CD3-APC-eFluro 780 (Invitrogen, 47-0037-41), anti-CD8-APC-eFluro 780 (Invitrogen, 47-0086-42), anti-CD16-APC-eFluro 780 (Invitrogen, 47-0168-41), anti-CD14-APC-eFluro 780 (Invitrogen, 47-0149-42), as well as Zombie NIR (BioLegend, 423105),

and fluorophore-labeled RBD and Ovalbumin for 30 minutes on ice [20]. Single $CD3^{-}CD8^{-}CD16^{-}CD20^{+}Ova^{-}RBD\text{-}PE^{+}RBD\text{-}APC^{+}$ B cells were sorted into individual wells of a 96-well plates containing 4 µl of lysis buffer (0.5 X PBS, 10mM DTT, 3000 units/mL RNasin Ribonuclease Inhibitors (Promega, N2615) per well using a FACS Aria III (Becton Dickinson). The sorted cells were frozen on dry ice, and then stored at −80°C or immediately used for subsequent RNA reverse transcription.

Antibody sequencing, cloning and expression. Antibodies were identified and sequenced as described previously [21-23]. Briefly, RNA from single cells was reverse-transcribed (SuperScript III Reverse Transcriptase, Invitrogen, 18080-044) and the cDNA stored at -20°C or used for subsequent amplification of the variable IGH, IGL and IGK genes by nested PCR and Sanger sequencing [21]. Amplicons from the first PCR reaction were used as templates for Sequence- and Ligation-Independent Cloning (SLIC) into antibody expression vectors. Recombinant monoclonal antibodies were produced and purified as previously described [24].

Biolayer interferometry.

BLI assays were performed on the Octet Red instrument (ForteBio) at 30 °C with shaking at 1,000 r.p.m. Epitope binding assays were performed with protein A biosensor (ForteBio 18-5010), following the manufacture protocol "classical sandwich assay". (1) Sensor check: sensors immersed 30sec in buffer alone (buffer ForteBio 18-1105). (2) Capture 1st Ab: sensors immersed 10min with Ab1 at 40µg/mL. (3) Baseline: sensors immersed 30sec in buffer alone. (4) Blocking: sensors immersed 5min with IgG isotype control at 50µg/mL. (6) Antigen association: sensors immersed 5min with RBD at 100µg/mL. (7) Baseline: sensors immersed 30sec in buffer alone. (8)

Association Ab2: sensors immersed 5min with Ab2 at 40µg/mL. Curve fitting was performed using the Data analysis software (ForteBio).

Computational analyses of antibody sequences. Antibody sequences were trimmed based on quality and annotated using Igblastn v1.14.0 [25] with IMGT domain delineation system. Annotation was performed systematically using Change-O toolkit v.0.4.5 [26]. Heavy and light chains derived from the same cell were paired, and clonotypes were assigned based on their V and J genes using in-house R and Perl scripts (Fig. 3b, c).

The frequency distributions of human V genes in anti-SARS-CoV-2 antibodies from this study was compared to Sequence Read Archive SRP010970 [17]. The V(D)J assignments were done using IMGT/High V-Quest and the frequencies of heavy and light chain V genes were calculated for 14 and 13 individuals, respectively, using sequences with unique CDR3s. The two-tailed t test with unequal variances was used to determine statistical significance (Extended Data Figure 7).

Nucleotide somatic hypermutation and CDR3 length were determined using in-house R and Perl scripts. To calculate the GRAVY scores of hydrophobicity [27] we used Guy H.R. Hydrophobicity scale based on free energy of transfer (kcal/mole)[28] implemented by the R package Peptides available in the Comprehensive R Archive Network repository (https://journal.r-project.org/archive/2015/RJ-2015-001/RJ-2015-001.pdf). We used 534 heavy chain CDR3 amino acid sequences from this study and 22,654,256 IGH CDR3 sequences from the public database of memory B-cell receptor sequences [29]. The Shapiro-Wilk test was used to determine whether the GRAVY scores are normally distributed. The GRAVY scores from all 534 IGH CDR3 amino acid sequences from this study were used to perform the test and 5000 GRAVY scores of the sequences from the public database were randomly selected. The Shapiro-Wilk p-values were 6.896×10^{-3}

and 2.217 x 10^{-6} for sequences from this study and the public database, respectively, indicating the data are not normally distributed. Therefore, we used the Wilcoxon non-parametric test to compare the samples, which indicated a difference in hydrophobicity distribution ($p = 5 \times 10^{-6}$; Extended Data Figure 7).

Acknowledgements: We thank all study participants who devoted time to our research; Drs. Barry Coller and Sarah Schlesinger, the Rockefeller University Hospital Clinical Research Support Office and nursing staff. Dr. Joseph L. DeRisi for facilitating interactions with the Chan Zuckerberg BioHub. All members of the M.C.N. laboratory for helpful discussions and Masa Jankovic for laboratory support. This work was supported by NIH grant P01-AI138398-S1 and 2U19AI111825 to M.C.N. and C.M.R.; George Mason University Fast Grant to D.F.R.; 3 R01-AI091707-10S1 to C.M.R.; R37-AI64003 to P.D.B.; R01AI78788 to T.H.; The G. Harold and Leila Y. Mathers Charitable Foundation to C.M.R.. C.G. was supported by the Robert S. Wennett Post-Doctoral Fellowship, in part by the National Center for Advancing Translational Sciences (National Institutes of Health Clinical and Translational Science Award program, grant UL1 TR001866), and by the Shapiro-Silverberg Fund for the Advancement of Translational Research. P.D.B. and M.C.N. are Howard Hughes Medical Institute Investigators.

Contributions: D.F.R., P.D.B., P.J.B., T.H., C.R. and M.C.N. conceived, designed and analyzed the experiments. D.F.R., M.C. and C.G. designed clinical protocols. F.M., J.C.C.L., Z.W., A.C., M.A., C.B.O., S.F., T.H., C.V., K.G., F.B., S.C., P.M.D., H.H., L.N., F.S., Y.W., H.-H.H., E.M., A.A., K.E.H.T., N.K. and P.R.H. carried out all experiments. A.G. and M.C. produced antibodies. C.O.B., J.P. and E.W. produced SARS-CoV-2 proteins. A.H., R.K., J.H., K.G.M., C.G. and M.C. recruited participants and executed clinical protocols. R.P., J.D., M.P. and I.S. processed clinical samples. T.Y.O., A.P.W. and V.R. performed bioinformatic analysis. D.F.R., P.D.B., P.J.B., T.H., C.M.R. and M.C.N. wrote the manuscript with input from all co-authors.

Extended Data Tables

Extended Data Table 1. Individual participant demographics and clinical characteristics

						Duration (days)		Sx	ELISA binding (AUC)				
						Sx	Sx onset	Severity	RBD		S		Neutralization
ID	Age	Gender	Race	Ethnicity	Case/Contact	total	to visit	(0-10)	IgG	IgM	IgG	IgM	(NT50)
7	40	M	White	Non-Hispanic	Case	11	30	6	2.92	2.54	7.39	2.37	2730.4
12	27	F	White	Non-Hispanic	Contact	7	24	3	1.47	1.71	4.31	1.23	5
13	28	M	White	Non-Hispanic	Case	5	25	3	1.97	4.10	3.95	1.38	173.2
18	55	M	White	Non-Hispanic	Case	16	28	6	2.57	1.75	3.92	1.31	410.3
20	26	F	White	Non-Hispanic	Case	2	17	5	1.80	2.00	4.49	1.68	5
21	54	M	White	Hispanic	Contact	11	27	7	5.60	2.88	7.60	2.47	5052.7
24	34	M	White	Non-Hispanic	Case	15	30	4	2.29	1.97	4.36	1.45	280.7
26	66	M	White	Non-Hispanic	Case	2	35	3	1.67	1.86	4.00	1.58	276.1
27	28	M	White	Non-Hispanic	Case	9	32	4	1.79	1.64	4.41	1.30	739.3
28	26	M	White	Non-Hispanic	Case	7	21	4	2.39	2.25	4.93	1.66	888.9
29	26	F	White	Non-Hispanic	Contact	5	35	4	1.69	1.54	3.93	1.97	5
30	30	M	White	Non-Hispanic	Contact	3	35	4	3.07	2.55	4.86	1.34	5
31	51	M	White	Non-Hispanic	Case	9	33	3	1.29	1.60	4.69	1.20	192.3
32	46	F	White	Non-Hispanic	Case	8	32	3	1.30	2.18	3.79	0.83	47.4
37	27	M	White	Non-Hispanic	Case	6	28	6	2.13	1.52	4.11	1.04	286.3
38	57	F	White	Non-Hispanic	Case	10	38	4	1.87	2.02	4.30	1.12	518.9
40	44	M	White	Non-Hispanic	Case	7	23	5	1.72	1.50	3.81	1.47	42.1
41	35	M	White	Non-Hispanic	Contact	10	29	8	2.13	1.67	4.15	0.98	302.5
42	40	M	White	Non-Hispanic	Contact	20	36	8	2.37	1.50	5.82	1.43	627.1
46	39	M	White	Non-Hispanic	Case	8	30	2	2.68	1.45	4.14	1.29	59.2
47	43	F	White	Non-Hispanic	Case	11	33	5	3.02	2.30	6.23	2.66	10433.3
48	37	F	White	Non-Hispanic	Case	7	21	5	1.51	1.85	3.47	1.51	173.4
50	27	F	White	Non-Hispanic	Contact	7	28	4	1.64	2.32	5.42	1.45	924.7
51	21	M	White	Non-Hispanic	Contact	8	31	5	1.76	1.69	3.94	1.03	1499.2
54	40	F	White	Non-Hispanic	Contact	3	24	3	1.80	2.10	4.99	1.42	5
56	75	F	White	Non-Hispanic	Case	22	40	3	1.79	1.81	5.89	1.47	1388.4
57	66	M	White	Non-Hispanic	Case	6	21	5	1.54	2.10	4.33	1.00	2048.9
58	64	F	White	Non-Hispanic	Contact	1	32	2	1.20	1.71	3.95	1.21	5
64	28	F	White	Non-Hispanic	Contact	11	32	6	2.36	2.06	4.48	1.66	776.7
67	19	F	N/A	Hispanic	Case	5	29	6	2.56	1.98	5.48	1.32	2052.9
71	45	F	White	Non-Hispanic	Case	12	48	7	1.55	2.04	3.86	1.53	33.3
72	42	M	White	Non-Hispanic	Case	16	35	8	4.05	3.58	6.05	2.59	3138.2
75	46	F	White	Non-Hispanic	Case	10	36	4	1.64	2.37	3.98	1.20	271.5
76	49	F	White	Non-Hispanic	Case	28	34	4	1.88	1.65	5.17	0.97	219.8
77	37	M	White	Non-Hispanic	Contact	6	33	4	1.33	1.83	3.56	1.16	5
81	44	F	White	Non-Hispanic	Contact	3	35	2	1.58	1.73	3.82	1.10	5
82	46	M	N/A	Non-Hispanic	Case	0	-	0	2.19	1.92	5.41	1.81	130.7
88	41	M	White	Non-Hispanic	Case	7	23	4	1.82	3.32	4.97	1.37	424.7
95	44	M	White	Non-Hispanic	Case	9	36	6	2.61	1.62	6.03	1.62	961.9
96	48	F	White	Non-Hispanic	Case	9	30	3	3.93	1.93	6.25	2.26	927.7
97	39	M	White	Non-Hispanic	Case	9	31	3	1.58	1.61	4.03	2.33	202.7
98	35	F	White	Non-Hispanic	Case	2	24	4	1.78	1.47	5.38	1.22	249.0
99	36	F	White	Non-Hispanic	Case	13	29	5	2.50	3.27	4.38	2.49	1127.6
107	53	F	White	Non-Hispanic	Contact	10	29	4	1.74	1.41	4.66	0.90	297.5
110	27	M	White	Non-Hispanic	Case	1	25	1	1.30	1.60	4.00	0.93	5
114	30	F	White	Non-Hispanic	Case	15	36	7	1.65	1.92	3.53	1.55	110.9
124	63	F	Asian	Non-Hispanic	Contact	4	37	3	1.58	2.06	3.49	1.32	5
125	51	F	White	Non-Hispanic	Case	10	26	3	1.92	3.49	3.86	1.24	126.5
130	39	M	White	Non-Hispanic	Contact	7	28	5	1.24	1.72	3.91	1.34	5
131	39	M	White	Non-Hispanic	Case	5	25	4	1.46	1.38	4.44	1.03	7.8
134	27	F	White	Non-Hispanic	Contact	16	22	5	2.51	2.18	6.84	1.94	2700.6
135	62	F	White	Non-Hispanic	Case	8	31	6	2.20	2.02	3.80	1.13	350.0
149	41	M	White	Non-Hispanic	Contact	17	28	6	1.68	2.02	3.67	1.09	494.9
154	68	M	Asian	Non-Hispanic	Case	16	30	9	3.19	2.19	4.85	1.29	928.2
157	50	M	White	Non-Hispanic	Case	10	32	8	2.40	2.86	3.90	2.06	741.7
172	38	F	White	Non-Hispanic	Case	8	22	9	1.71	2.58	4.29	1.20	301.1
178	26	F	White	Non-Hispanic	Case	6	24	4	1.54	1.59	3.66	1.02	5.0
179	39	M	White	Non-Hispanic	Contact	10	37	3	1.89	2.25	3.83	1.73	370.1
186	38	F	N/A	N/A	Case	8	26	2	1.73	2.47	4.37	1.23	296.9
195	24	M	White	Non-Hispanic	Case	18	42	5	2.74	2.55	5.01	2.33	1315.1
201	50	M	White	Non-Hispanic	Contact	15	33	6	4.37	2.57	6.15	1.57	3897.4
202	57	M	White	Non-Hispanic	Case	21	34	7	2.10	2.08	5.07	1.17	257.9
233	55	M	White	Non-Hispanic	Case	20	41	3	2.07	2.11	4.51	1.07	173.2
241	36	M	White	Non-Hispanic	Case	12	30	7	2.27	2.66	4.54	1.46	923.1
243	30	F	Asian	Non-Hispanic	Case	6	26	5	2.92	2.57	5.06	1.14	1300.2
246	44	F	White	Non-Hispanic	Case	10	38	7	2.05	2.79	6.09	1.32	566.0
280	59	M	White	Non-Hispanic	Case	6	32	7	2.53	3.07	4.61	1.19	1072.1
460	36	M	White	Non-Hispanic	Case	11	39	6	2.94	3.18	5.51	2.80	1906.7

Extended Data Table 2. Frequency of symptoms and comorbidities reported by participants

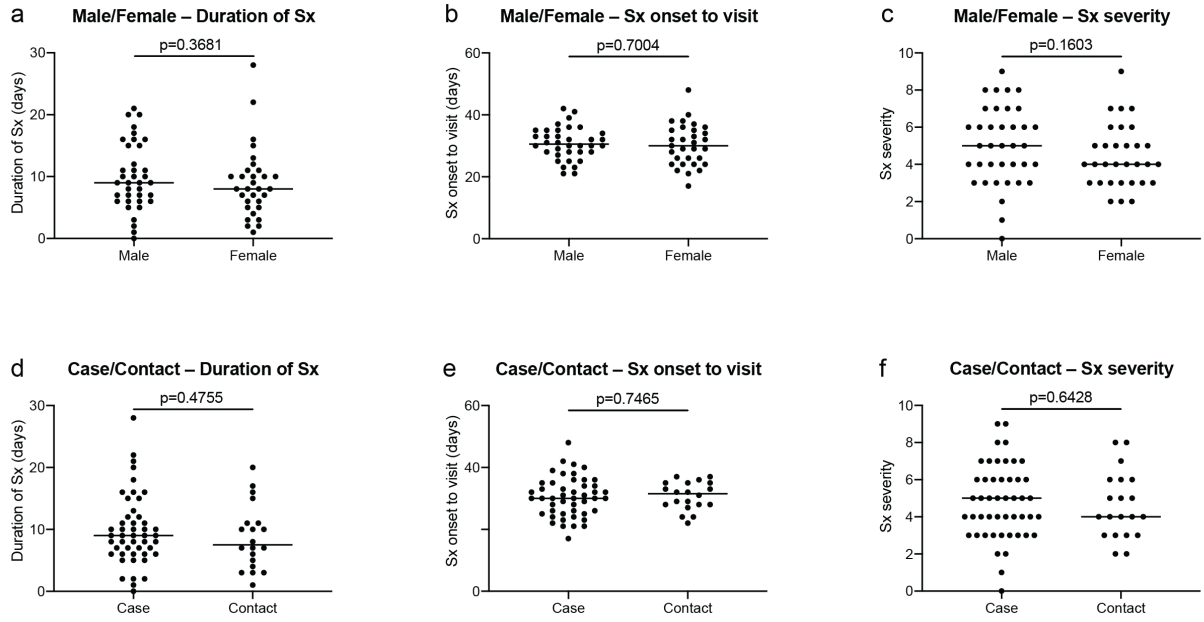
Symptom	Participants (n=68)	%
Fever	56	82.4
Cough	44	64.7
Myalgia	38	55.9
Fatigue	37	54.4
Shortness of breath	29	42.6
Headache	26	38.2
Sore throat	23	33.8
Loss of smell/taste	18	26.5
Diarrhea	15	22.1
Presence of comorbidities (HTN, CAD, DM, COPD, asthma, cancer)	6	8.8

HTN (hypertension), CAD (coronary artery disease), DM (diabetes mellitus), COPD (chronic obstructive pulmonary disease)

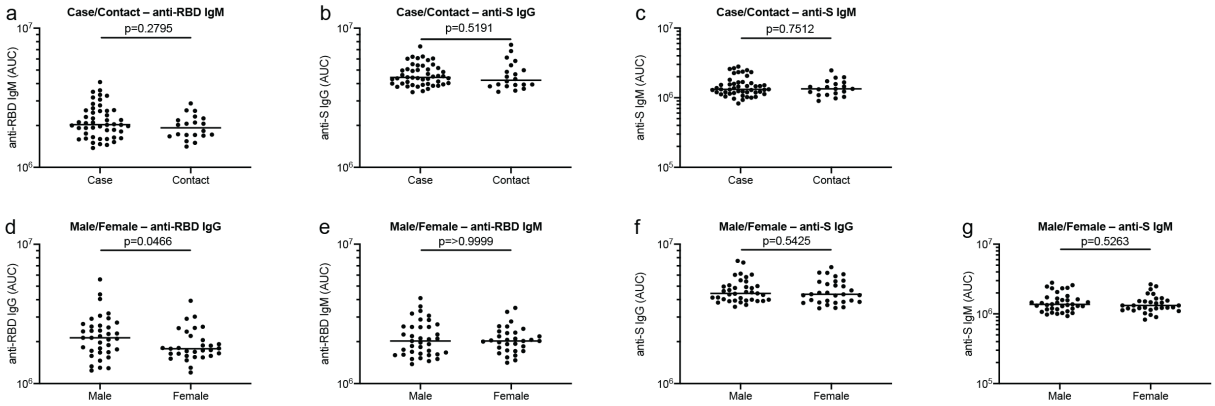
Extended Data Table 3. IC50, IC80 and IC90 of monoclonal antibodies

Participant ID	Antibody ID	IC50 ng/ml	IC80 ng/ml	IC90 ng/ml
COV21	C002	8.88	21.95	37.61
COV21	C003	313.79	992.62	>1000
COV21	C004	10.91	34.63	68.04
COV21	C005	60.49	130.65	205.20
COV21	C006	321.51	>1000	>1000
COV21	C008	625.46	>1000	>1000
COV21	C010	>1000	>1000	>1000
COV21	C013	42.48	360.59	1285.05
COV21	C016	>1000	>1000	>1000
COV21	C017	72.67	256.18	543.87
COV21	C018	>1000	>1000	>1000
COV21	C019	>1000	>1000	>1000
COV21	C021	>1000	>1000	>1000
COV107	C102	34.03	84.21	143.23
COV107	C103	4.38	12.58	23.59
COV107	C104	23.31	72.12	140.28
COV107	C105	26.09	72.24	133.70
COV107	C106	>1000	>1000	>1000
COV107	C107	>1000	>1000	>1000
COV107	C108	480.69	>1000	>1000
COV107	C110	18.44	45.11	77.28
COV107	C113	>1000	>1000	>1000
COV107	C114	>1000	>1000	>1000
COV107	C115	198.33	958.18	>1000
COV107	C117	348.00	>1000	>1000
COV107	C119	9.12	39.45	97.78
COV107	C121	6.73	14.31	22.33
COV072	C130	>1000	>1000	>1000
COV072	C131	30.52	178.90	759.11
COV072	C132	708.67	>1000	>1000
COV072	C133	>1000	>1000	>1000
COV072	C134	>1000	>1000	>1000

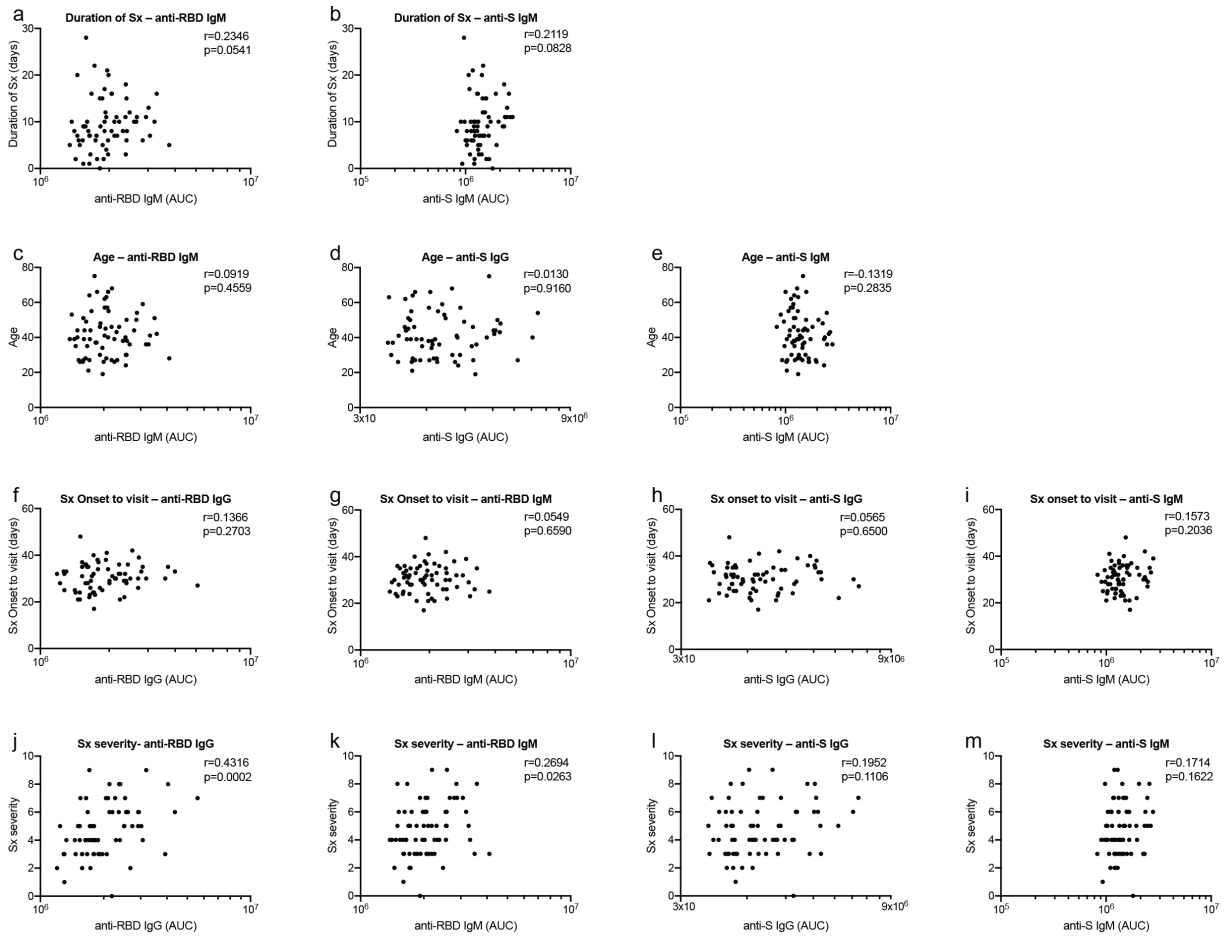
Extended Data Figures



Extended Data Figure 1. Clinical correlates. a, Duration of symptoms in days (Y axis) for all males and females in the cohort p=0.3681. **b,** Time between symptom onset and plasma collection (Y axis) for all males and females in the cohort p=0.7004. **c,** Subjective symptom severity on a scale of 0-10 (Y axis) for all males and females in the cohort p=0.1603. **d,** Duration of symptoms in days (Y axis) for all cases and contacts p=0.4755. **e,** Time between symptom onset and plasma collection in days (Y axis) for all cases and contacts p=0.7465. **f,** Symptom severity (Y axis) for all cases and contacts in the cohort p=0.6428. Horizontal bars indicate median values. Statistical significance was determined using two-tailed Mann-Whitney U test.

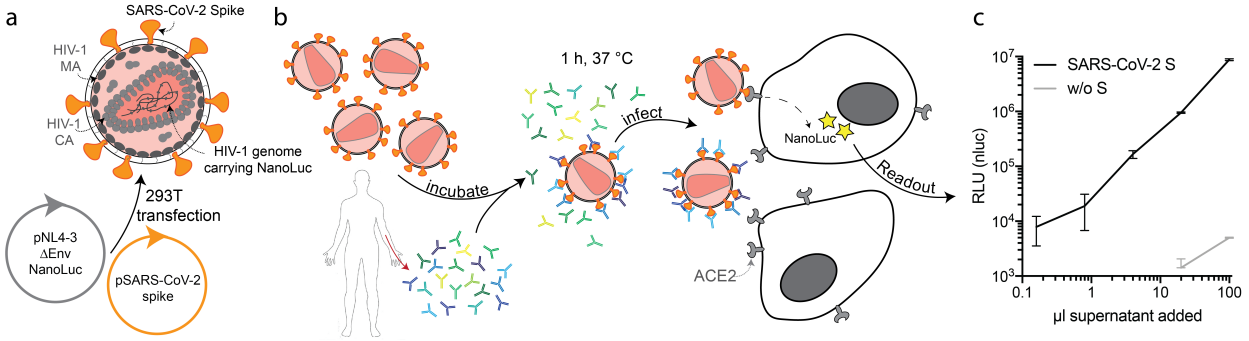


Extended Data Figure 2. Clinical correlates of plasma antibody titers. a, AUC for IgM anti-RBD (Y axis) for all cases and contacts in the cohort p=0.2795. **b,** AUC for IgG anti-S (Y axis) for all cases and contacts in the cohort p=0.5191. **c,** AUC for IgM anti-S (Y axis) for all cases and contacts in the cohort p=0.7512. **d,** AUC for IgG anti-RBD (Y axis) for all males and females in the cohort p=0.0466. **e,** AUC for IgM anti-RBD (Y axis) for all males and females in the cohort p=0.0466. **f,** AUC for IgG anti-S (Y axis) for all males and females in the cohort p=0.5425. **g,** AUC for IgM anti-S (Y axis) for all males and females in the cohort p=0.5263. Horizontal bars indicate median values. Statistical significance was determined using two-tailed Mann-Whitney U test.

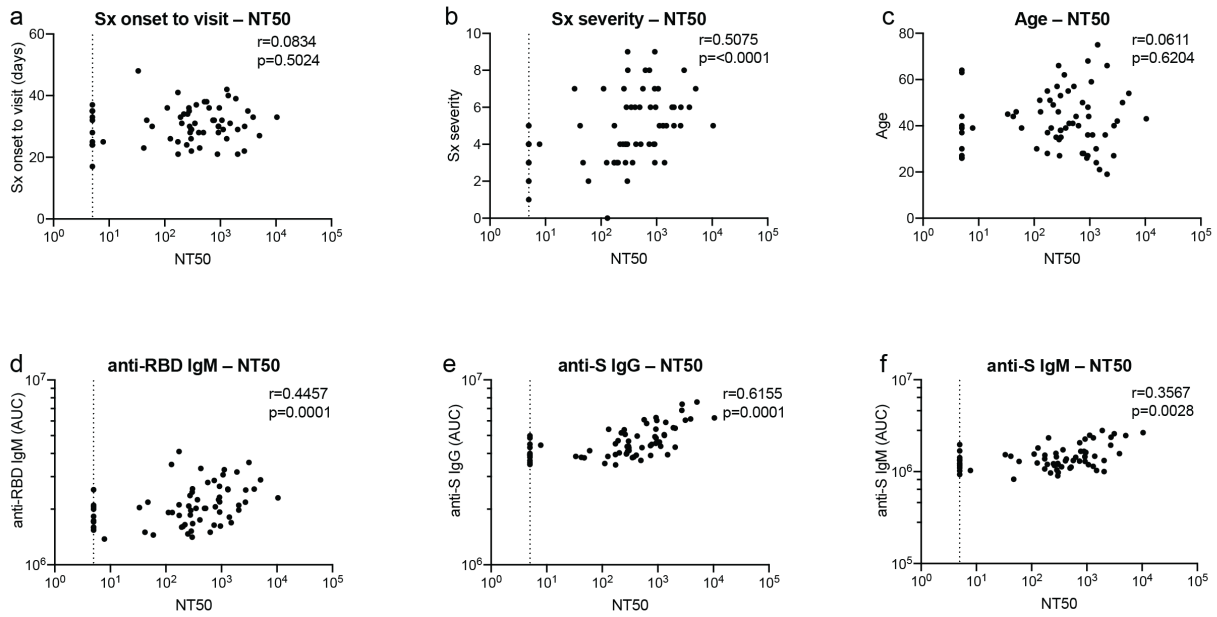


Extended Data Figure 3. Additional clinical correlates of plasma antibody titers. a, Duration of symptoms in days (Y axis) plotted against AUC for IgM anti-RBD (X axis) r=0.2346, p=0.0541. **b,** Duration of symptoms in days (Y axis) plotted against AUC for IgM anti-S (X axis) r=0.2119, p=0.0828. **c,** Age (Y axis) plotted against AUC for IgM anti-RBD (X axis) r=0.0919 p=0.4559. **d,** Age (Y axis) plotted against AUC for IgG anti-S (X axis) r=0.0130 p=0.9160. **e,** Age (Y axis) plotted against AUC for IgM anti-S (X axis) r=0.1319 p=0.2835. **f,** Time between symptom onset and plasma collection in days (Y axis) plotted against AUC for IgG anti-RBD (X axis) r=0.1366 p=0.2703. **g,** Time between symptom onset and plasma collection in days (Y axis) plotted against AUC for IgM anti-RBD (X axis) r=0.0549 p=0.6590. **h,** Time between symptom onset and plasma collection in days (Y axis) plotted against AUC for IgG anti-S (X axis) r=0.0565 p=0.6500. **i,** Time

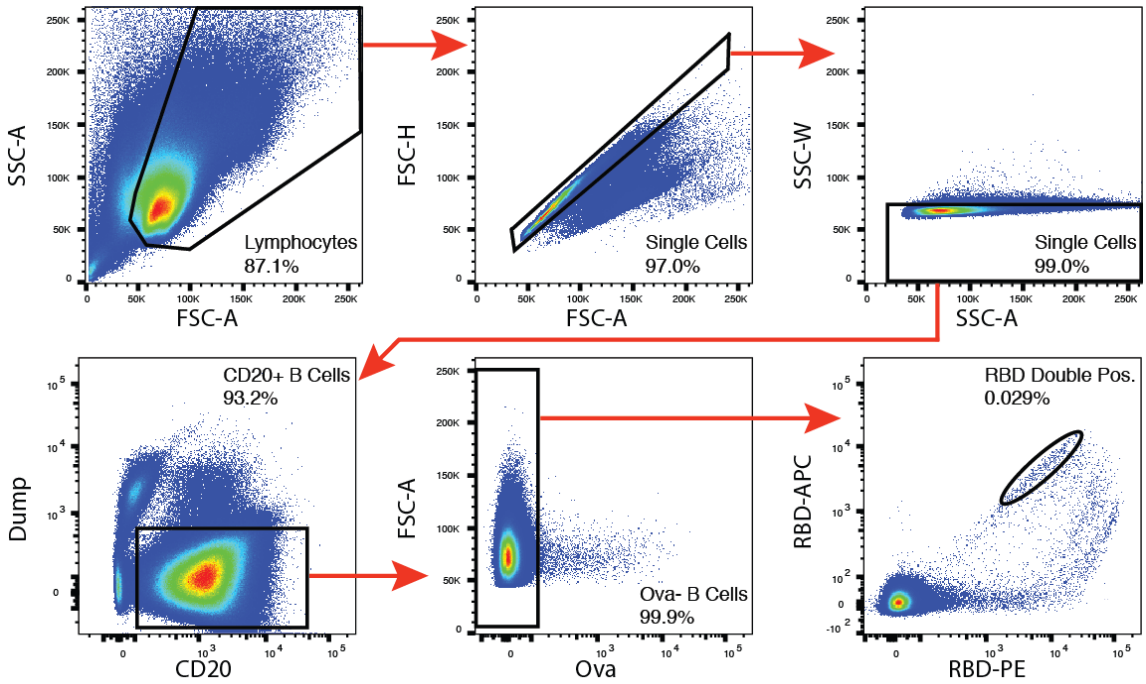
between symptom onset and plasma collection in days (Y axis) plotted against AUC for IgM anti-S (X axis) r=0.1573 p=0.2036. **j,** Severity of symptoms (Y axis) plotted against AUC for IgG anti-RBD (X axis) r=0.4316 p=0.0002. **k,** Severity of symptoms (Y axis) plotted against AUC for IgM anti-RBD (X axis) r=0.2694 p=0.0263. **l,** Severity of symptoms (Y axis) plotted against AUC for IgG anti-S (X axis) r=0.1952 p=0.1106. **m,** Severity of symptoms (Y axis) plotted against AUC for IgM anti-S (X axis) r=0.1714 p=0.1622. All correlations were analyzed by two-tailed Spearman's.



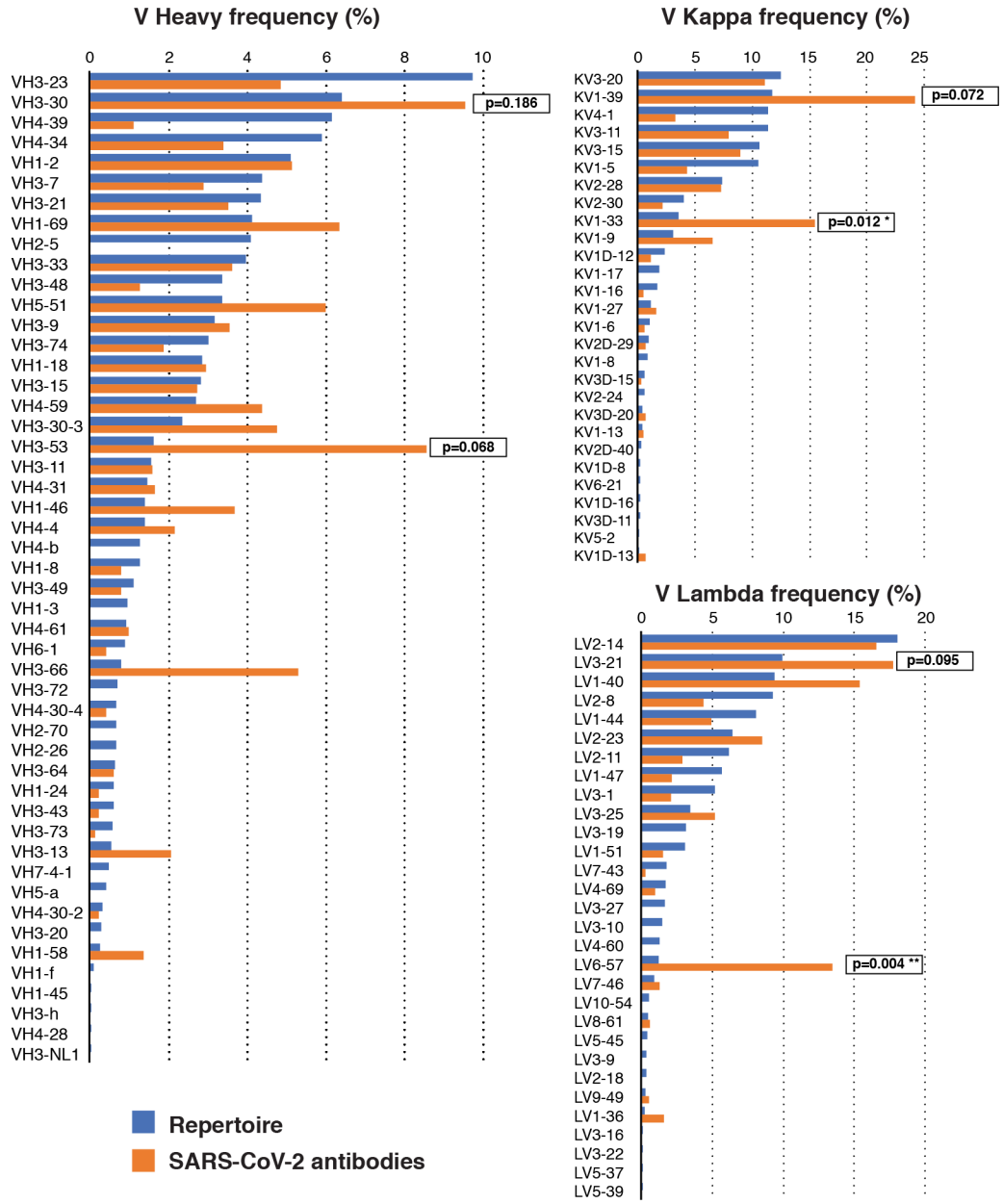
Extended Data Figure 4. Diagrammatic representation of the SARS-CoV2-S_{trunc} luciferase assay. a, Co-transfection of pNL4-3ΔEnv-NanoLuc and pSARS-CoV-2 spike vectors into 293T cells leads to production of SARS-CoV-2 Spike-pseudotyped HIV-1 particles (SARS-CoV-2 pseudovirus) carrying the *NanoLuc* gene. **b,** SARS-CoV-2 pseudovirus is incubated for 1 h at 37 °C with plasma dilutions from COVID-19 patients or monoclonal antibodies. The virus-antibody mixture is used to infect ACE2-expressing 293T cells, which will express NanoLuc Luciferase upon infection. **c,** Relative luminescence units (RLU) reads from lysates of ACE2-expressing 293T cells infected with increasing amounts of SARS-CoV-2 pseudovirus (SARS-CoV-2 S) or non-pseudotyped control virus (w/o S).



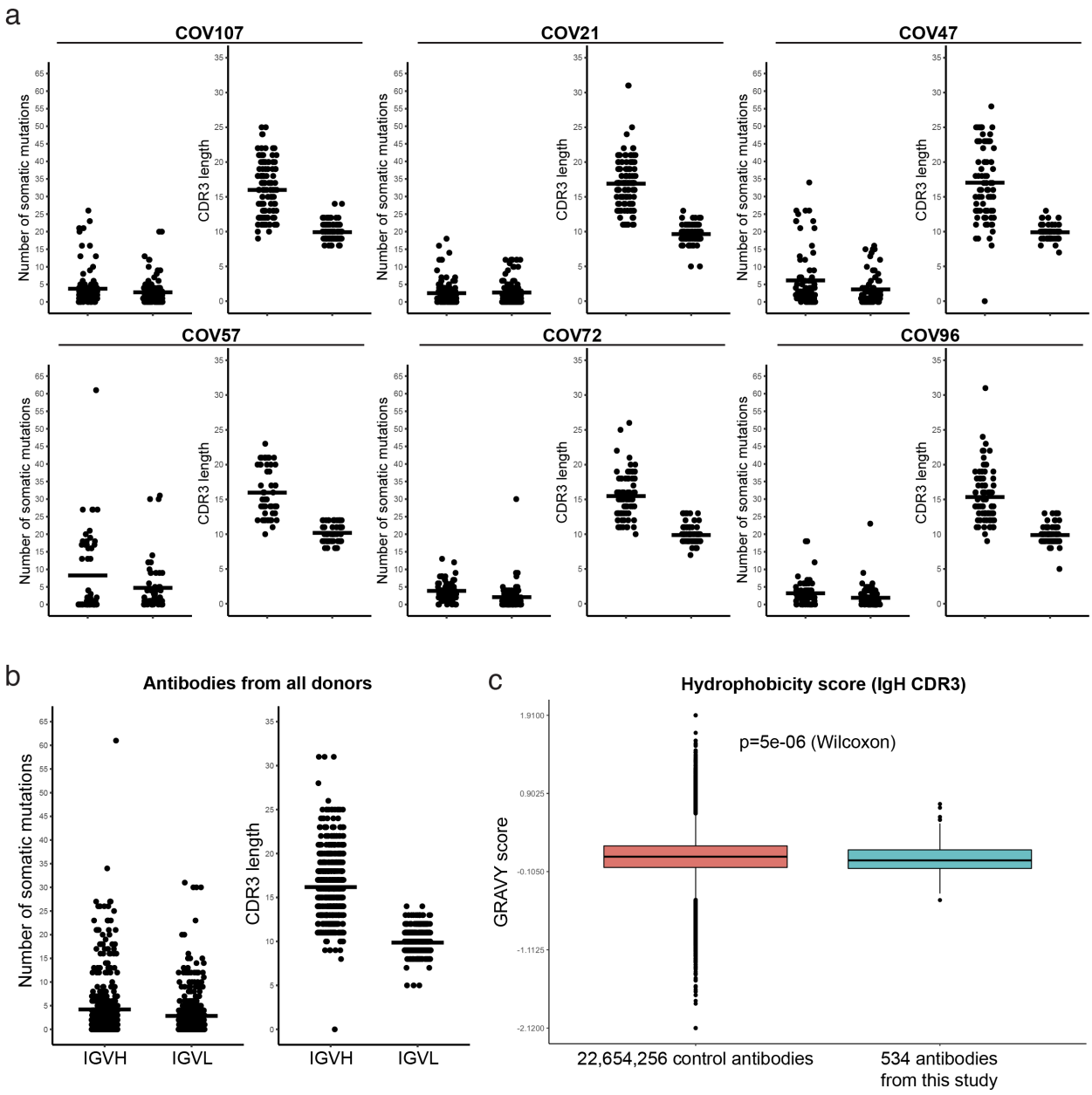
Extended Data Figure 5. Clinical correlates of neutralization. a, Time between symptom onset and plasma collection in days (Y axis) plotted against NT_{50} (X axis) r=0.0834, p=0.5024. **b,** Symptom severity (Y axis) plotted against NT_{50} (X axis) r=0.5075, p=<0.0001. **c,** Age (Y axis) plotted against NT_{50} (X axis) r=0.0611, p=0.6204. **d,** Anti-RBD IgM AUC (Y axis) plotted against NT_{50} (X axis) r=0.4457, p=0.0001. **e,** Anti-S IgG AUC (Y axis) plotted against NT_{50} (X axis) r=0.6155, p=0.0001. **f,** Anti-S IgM AUC (Y axis) plotted against NT_{50} (X axis) r=0.3567, p=0.0028. All correlations were analyzed by two-tailed Spearman's. Dotted line (NT_{50}=5) represents lower limit of detection (LLOD) of pseudovirus neutralization assay. Samples with undetectable neutralizing titers were plotted at LLOD.



Extended Data Figure 6. Flow cytometry. Gating strategy used for cell sorting. Gating was on singlets that were $CD20^{+}$ and $CD3^{-}CD8^{-}CD16^{-}Ova^{-}$. Sorted cells were RBD-$PE^{+}$ and RBD-$BV711^{+}$.



Extended Data Figure 7. Frequency distributions of human V genes. The two-tailed t test with unequal variance was used to compare the frequency distributions of human V genes of anti-SARS-CoV-2 antibodies from this study to Sequence Read Archive SRP010970[17].



Extended Data Figure 8. Analysis of antibody somatic hypermutation and CDR3 length. a, For each individual, the number of somatic nucleotide mutations (Y axis) at the IGVH and IGVL are shown on the left panel, and the amino acid length of the CDR3s (Y axis) are shown on the right panel. The horizontal bar indicated the mean. **b,** same as in a but for all antibodies combined. **c**, Distribution of the hydrophobicity GRAVY scores at the IGH CDR3 in antibody sequences from this study compared to a public database (see Methods).

References

1 Graham, R. L., Donaldson, E. F. & Baric, R. S. A decade after SARS: strategies for controlling emerging coronaviruses. *Nat Rev Microbiol* **11**, 836-848, doi:10.1038/nrmicro3143 (2013).

2 Gralinski, L. E. & Baric, R. S. Molecular pathology of emerging coronavirus infections. *J Pathol* **235**, 185-195, doi:10.1002/path.4454 (2015).

3 Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278, doi:10.1016/j.cell.2020.02.052 (2020).

4 Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292 e286, doi:10.1016/j.cell.2020.02.058 (2020).

5 Jiang, S., Hillyer, C. & Du, L. Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends Immunol*, doi:10.1016/j.it.2020.03.007 (2020).

6 Scheid, J. F. *et al.* Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458**, 636-640, doi:10.1038/nature07930 (2009).

7 Tiller, T. *et al.* Autoreactivity in human IgG+ memory B cells. *Immunity* **26**, 205-213, doi:10.1016/j.immuni.2007.01.009 (2007).

8 Murugan, R. *et al.* Clonal selection drives protective memory B cell responses in controlled human malaria infection. *Sci Immunol* **3**, doi:10.1126/sciimmunol.aap8029 (2018).

9 Briney, B., Inderbitzin, A., Joyce, C. & Burton, D. R. Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature* **566**, 393-397, doi:10.1038/s41586-019-0879-y (2019).

10 ter Meulen, J. *et al.* Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. *PLoS Med* **3**, e237, doi:10.1371/journal.pmed.0030237 (2006).

11 Salazar, G., Zhang, N., Fu, T. M. & An, Z. Antibody therapies for the prevention and treatment of viral infections. *NPJ Vaccines* **2**, 19, doi:10.1038/s41541-017-0019-3 (2017).

12 Bournazos, S. & Ravetch, J. V. Anti-retroviral antibody FcgammaR-mediated effector functions. *Immunol Rev* **275**, 285-295, doi:10.1111/imr.12482 (2017).

13 Feinberg, M. B. & Ahmed, R. Advancing dengue vaccine development. *Science* **358**, 865-866, doi:10.1126/science.aaq0215 (2017).

14 Iwasaki, A. & Yang, Y. The potential danger of suboptimal antibody responses in COVID-19. *Nat Rev Immunol*, doi:10.1038/s41577-020-0321-6 (2020).

15 Van Rompay, K. K. A. *et al.* A combination of two human monoclonal antibodies limits fetal damage by Zika virus in macaques. *Proc Natl Acad Sci U S A* **117**, 7981-7989, doi:10.1073/pnas.2000414117 (2020).

16 Plotkin, S. A. Correlates of protection induced by vaccination. *Clin Vaccine Immunol* **17**, 1055-1065, doi:10.1128/CVI.00131-10 (2010).

17 Rubelt, F. *et al.* Onset of immune senescence defined by unbiased pyrosequencing of human immunoglobulin mRNA repertoires. *PLoS One* **7**, e49774, doi:10.1371/journal.pone.0049774 (2012).

18 Kane, M. *et al.* Identification of Interferon-Stimulated Genes with Antiretroviral Activity. *Cell Host Microbe* **20**, 392-405, doi:10.1016/j.chom.2016.08.005 (2016).

19 Adachi, A. *et al.* Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol* **59**, 284-291 (1986).

20 Wang, Z. *et al.* Isolation of single HIV-1 Envelope specific B cells and antibody cloning from immunized rhesus macaques. *J Immunol Methods* **478**, 112734, doi:10.1016/j.jim.2019.112734 (2020).

21 Tiller, T. *et al.* 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, doi:10.1016/j.jim.2007.09.017 (2008).

22 von Boehmer, L. *et al.* Sequencing and cloning of antigen-specific antibodies from mouse memory B cells. *Nat Protoc* **11**, 1908-1923, doi:10.1038/nprot.2016.102 (2016).

23 Robbiani, D. F. *et al.* Recurrent Potent Human Neutralizing Antibodies to Zika Virus in Brazil and Mexico. *Cell* **169**, 597-609 e511, doi:10.1016/j.cell.2017.04.024 (2017).

24 Klein, F. *et al.* Enhanced HIV-1 immunotherapy by commonly arising antibodies that target virus escape variants. *J Exp Med* **211**, 2361-2372, doi:10.1084/jem.20141050 (2014).

25 Ye, J., Ma, N., Madden, T. L. & Ostell, J. M. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res* **41**, W34-40, doi:10.1093/nar/gkt382 (2013).

26 Gupta, N. T. *et al.* Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356-3358, doi:10.1093/bioinformatics/btv359 (2015).

27 Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132, doi:10.1016/0022-2836(82)90515-0 (1982).

28 Guy, H. R. Amino acid side-chain partition energies and distribution of residues in soluble proteins. *Biophys J* **47**, 61-70, doi:10.1016/S0006-3495(85)83877-7 (1985).

29 DeWitt, W. S. *et al.* A Public Database of Memory and Naive B-Cell Receptor Sequences. *PLoS One* **11**, e0160853, doi:10.1371/journal.pone.0160853 (2016).