Mapping SARS-CoV-2 Antibody Epitopes in COVID-19 Patients with a Multi-Coronavirus Protein Microarray

David Camerini,[a,b] Arlo Z. Randall,[a] Krista Trappl-Kimmons,[a] Amit Oberai,[a] Christopher Hung,[a] Joshua Edgar,[a] Adam Shandling,[a] Vu Huynh,[a] Andy A. Teng,[a] Gary Hermanson,[a] Jozelyn V. Pablo,[a] Megan M. Stumpf,[c] Sandra N. Lester,[c] Jennifer Harcourt,[c] Azaibi Tamin,[c] Mohammed Rasheed,[c] Natalie J. Thornburg,[c] Panayampalli S. Satheshkumar,[c] Xiaowu Liang,[a] Richard B. Kennedy,[d] Angela Yee,[a] Michael Townsend,[c] Joseph J. Campo[a]

[a]Antigen Discovery Incorporated (ADI), Irvine, California, USA
[b]University of California, Irvine, California, USA
[c]Centers for Disease Control and Prevention, Atlanta, Georgia, USA
[d]Mayo Clinic, Rochester, Minnesota, USA

Michael Townsend and Joseph J. Campo co-senior authors. The author order was determined based on the contributions of each author.

ABSTRACT The rapid worldwide spread of SARS-CoV-2 has accelerated research and development for controlling the COVID-19 pandemic. A multi-coronavirus protein microarray was created containing full-length proteins, overlapping protein fragments of various lengths, and peptide libraries from SARS-CoV-2 and four other human coronaviruses. Sera from confirmed COVID-19 patients as well as unexposed individuals were applied to multicoronavirus arrays to identify specific antibody reactivity. High-level IgG, IgM, and IgA reactivity to structural proteins S, M, and N of SARS-CoV-2, as well as accessory proteins such as ORF3a and ORF7a, were observed that were specific to COVID-19 patients. Antibody reactivity against overlapping 100-, 50-, and 30-amino acid fragments of SARS-CoV-2 proteins was used to identify antigenic regions. Numerous proteins of SARS-CoV, Middle East respiratory syndrome coronavirus (MERS-CoV), and the endemic human coronaviruses HCoV-NL63 and HCoV-OC43 were also more reactive with IgG, IgM, and IgA in COVID-19 patient sera than in unexposed control sera, providing further evidence of immunologic cross-reactivity between these viruses. Whereas unexposed individuals had minimal reactivity against SARS-CoV-2 proteins that poorly correlated with reactivity against HCoV-NL63 and HCoV-OC43 S2 and N proteins, COVID-19 patient sera had higher correlation between SARS-CoV-2 and HCoV responses, suggesting that *de novo* antibodies against SARS-CoV-2 cross-react with HCoV epitopes. Array responses were compared with validated spike protein-specific IgG enzyme-linked immunosorbent assays (ELISAs), showing agreement between orthologous methods. SARS-CoV-2 microneutralization titers were low in the COVID-19 patient sera but correlated with array responses against S and N proteins. The multi-coronavirus protein microarray is a useful tool for mapping antibody reactivity in COVID-19 patients.

IMPORTANCE With novel mutant SARS-CoV-2 variants of concern on the rise, knowledge of immune specificities against SARS-CoV-2 proteins is increasingly important for understanding the impact of structural changes in antibody-reactive protein epitopes on naturally acquired and vaccine-induced immunity, as well as broader topics of cross-reactivity and viral evolution. A multi-coronavirus protein microarray used to map the binding of COVID-19 patient antibodies to SARS-CoV-2 proteins and protein fragments as well as to the proteins of four other coronaviruses that infect humans has shown specific regions of SARS-CoV-2 proteins that are highly reactive with patient antibodies and revealed cross-reactivity of these antibodies with other human coronaviruses. These data and the multi-coronavirus protein microarray tool will help

guide further studies of the antibody response to COVID-19 and to vaccination against this worldwide pandemic.

KEYWORDS COVID-19, HCoV, SARS-CoV-2, antibody binding sites

A novel human coronavirus which causes severe acute respiratory syndrome, now known as SARS-CoV-2, emerged in December 2019. Infection with SARS-CoV-2 spread rapidly worldwide, and on 11 March 2020, it was declared a pandemic by the World Health Organization (WHO) (1). As of 7 May 2021, there are over 156 million confirmed cases of coronavirus disease (COVID-19) caused by this new virus, resulting in more than 3.2 million deaths, corresponding to a case mortality rate of ~2.1% (https://coronavirus.jhu.edu/). Best estimates indicate that SARS-CoV-2 has a basic reproductive number, R_0, of 2 to 2.5 and an incubation time of approximately 4.6 days (2, 3), which allows rapid spread of the virus. Diagnosis, treatment, and vaccination against COVID-19 will all benefit from a clear understanding of the immune response to SARS-CoV-2 infection.

Previous studies have shown that COVID-19 patients rapidly seroconvert to SARS-CoV-2 and produce IgM, IgG, and IgA antibodies directed to several viral proteins (4–7). Reinfection challenge studies in rhesus macaques showed that the humoral and cellular immune response to SARS-CoV-2 infection was effective in blocking reinfection (8, 9). Nevertheless, it is not clear whether all antibody responses are beneficial or whether some antibody responses to SARS-CoV-2 lead to a less favorable course of disease (10, 11). Moreover, enhancement of infection by antibodies has been reported for severe acute respiratory syndrome coronavirus (SARS-CoV), which is closely related to SARS-CoV-2 (12–15).

We have created and used a multi-coronavirus protein microarray containing over 900 coronavirus proteins, protein fragments, and peptides to map IgG, IgA, and IgM antibody epitopes in sera from COVID-19 patients. Our approach localizes the antibody reactivity of COVID-19 patients within SARS-CoV-2 proteins and allows us to map the bound antigenic regions. Furthermore, we can similarly measure the antibody reactivity of COVID-19 patients and healthy controls with endemic human coronaviruses and with the two previous epidemic coronaviruses, SARS-CoV and Middle East respiratory syndrome coronavirus (MERS-CoV). Our findings and the multi-coronavirus protein microarray we created will be useful in discerning which *de novo* and cross-reactive antibody responses to SARS-CoV-2 are protective and which may be less useful in preventing disease or may even be detrimental. In addition, if high levels of antibody to specific epitopes are found to be especially protective, the array could be used to screen convalescent plasma for therapeutic potential and vaccine recipient sera as a preliminary measure of efficacy (16–19).

RESULTS

The multi-coronavirus protein microarray created and used in this study encompasses over 900 features. It includes the 4 structural proteins and 5 accessory proteins of SARS-CoV-2 as well as overlapping 100-, 50-, and 30-amino acid (aa) protein fragments to map immunodominant domains within each of these 9 SARS-CoV-2 proteins. It also contains the structural proteins of SARS-CoV, MERS-CoV, HCoV-NL63, and HCoV-OC43, plus overlapping 13- to 20-aa peptides of the SARS-CoV structural proteins and of the S proteins of MERS-CoV, HCoV-NL63, and HCoV-OC43 (Table 1).

The multicoronavirus array was incubated with sera from two sets of COVID-19 patient samples and associated negative controls collected in different regions of the United States. The first set of sera from 10 COVID-19 patients and 10 prepandemic healthy donors was obtained from the Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia. The second set included sera from 10 COVID-19 patients and 9 prepandemic samples obtained from the Mayo Clinic in Rochester, Minnesota. The age, sex, and SARS-CoV-2 enzyme-linked immunosorbent assay (ELISA) results of the

TABLE 1 Features of the 1st-generation ADI multi-coronavirus protein microarray[a]

Virus	S protein	E protein	M protein	N protein	Accessory or other proteins	No. of features (total n = 934)
SARS-CoV-2	RBD[b], S[c] from BEI, S1, S2 and 30-, 50-, and 100-aa fragments by IVTT at ADI	Protein and 30- and 50-aa fragments by IVTT at ADI	Protein and 30-, 50-, and 100-aa fragments by IVTT at ADI	Protein and 30-, 50-, and 100-aa fragments by IVTT at ADI	ORF 3a, 6, 7a, 8, and 10 protein and 30-, 50-, and 100-aa fragments by IVTT at ADI	12 proteins, 310 fragments = 322 features
SARS-CoV	Protein[d] and peptide set from BEI Resources	IVTT by ADI, peptides from BEI Resources	Protein, peptides from BEI Resources	Protein, peptides from BEI Resources	3CL protease from BEI Resources	5 proteins, 265 peptides = 270 features
MERS-CoV	Protein from BEI Resources	Protein produced by IVTT at ADI	Protein produced by IVTT at ADI	BEI Resources	ORF 3a, 4a, 4b, 5, and 8b proteins produced by IVTT at ADI	9 proteins
HCoV-NL63	Protein by IVTT at ADI, peptides from BEI	Protein produced by IVTT at ADI	Protein produced by IVTT at ADI	Protein produced by IVTT at ADI	ORF 3 protein produced by IVTT at ADI	5 proteins and 96 peptides = 101 features
HCoV-OC43	Protein by IVTT at ADI, peptides from BEI	Protein produced by IVTT at ADI	Protein produced by IVTT at ADI	Protein produced by IVTT at ADI	HE, N2 protein IVTT at ADI	6 proteins and 226 peptides = 232 features

[a]IVTT means coupled *in vitro* transcription and translation.
[b]RBD is the receptor binding domain, aa 319 to 541 of the SARS-CoV-2 S protein.
[c]SARS-CoV-2 S protein is a stabilized form with a trimerization sequence and transmembrane domain deletion.
[d]SARS-CoV-S protein is a transmembrane domain deleted form.

TABLE 2 Serum donors for samples obtained from the CDC

Sample ID[a]	SARS-CoV-2 S ELISA (S/T)[b]	Age (decade)	Sex
N1	0.25	60s	M
N2	0.09	20s	M
N3	0.18	40s	M
N4	0.34	50s	M
N5	019	60s	M
N6	0.94	40s	M
N7	0.26	40s	M
N8	0.15	50s	M
N9	0.40	40s	M
N10	0.19	50s	M
P1	2.06	60s	F
P2	3.98	60s	M
P3	5.47	50s	F
P4	2.35	20s	F
P4	5.16	70s	F
P6	5.21	70s	M
P7	3.82	30s	M
P8	5.30	80s	M
P9	3.81	40s	F
P10	5.20	40s	M

[a]N, healthy negative control; P, COVID-19-positive patient.
[b]S/T, signal/threshold for positivity ratio; >1.0 is positive.

COVID-19 patients and healthy control blood donors in both sample sets are shown in Tables 2 and 3. ELISAs were performed separately at the two different sites. Both assays clearly discriminated the COVID-19 patient samples from the control samples.

Specific antibody reactivity to SARS-CoV-2 and SARS-CoV purified recombinant proteins in COVID-19 patients. The specimens from COVID-19 patients had robust anti-SARS-CoV-2 IgG and IgA antibodies. IgM antibody responses were weaker. The magnitude and specificity of antibody responses were similar in the samples obtained from the CDC and the Mayo Clinic, so they are presented together here. COVID-19 patient serum IgG, IgA, and IgM reacted strongly to purified SARS-CoV-2 spike (S) protein as well as SARS-CoV nucleocapsid (N), S, and membrane (M) proteins compared to healthy control sera (Fig. 1). SARS-CoV-2 N protein was unavailable at the time of publication and was

TABLE 3 Serum donors for samples obtained from the Mayo Clinic

Sample ID[a]	SARS-CoV-2 S ELISA (S/C)[b]	Age (decade)	Sex
N101	<0.4	10s	M
N102	<0.4	10s	F
N103	NA	10s	M
N104	<0.4	10s	M
N105	NA	10s	M
N106	NA	10s	M
N107	<0.4	10s	F
N108	NA	10s	F
N109	<0.4	20s	F
P101	2.76	50s	F
P102	3.71	50s	M
P103	3.83	70s	M
P104	1.42	30s	M
P105	4.24	10s	F
P106	2.48	40s	F
P107	1.5	50s	F
P108	3.05	60s	F
P109	2.03	70s	M
P110	3.6	70s	M

[a]N, healthy negative control; P, COVID-19-positive patient.
[b]S/C, signal/calibrator ratio; >1.1 is positive. NA, not assayed, but collected prior to November 2019.
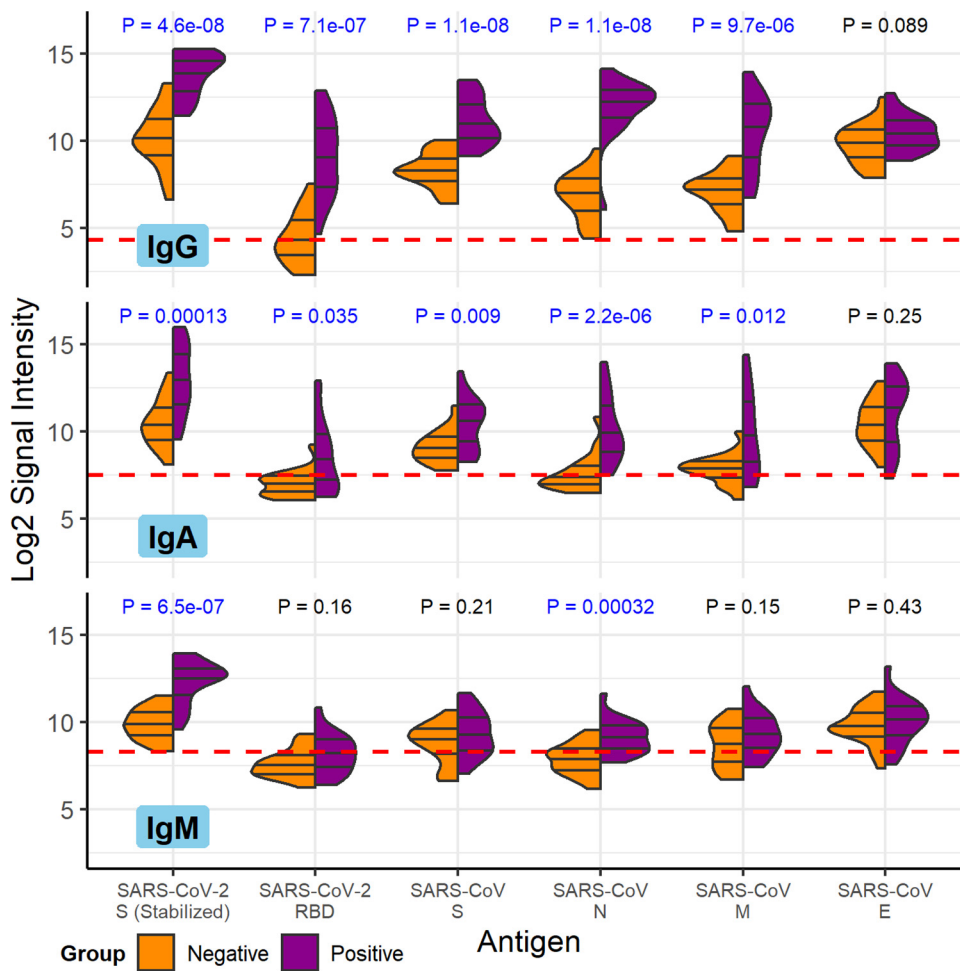
FIG 1 COVID-19 patient and healthy control antibody reactivity with purified SARS-CoV-2 and SARS-CoV proteins. The split violin plot shows the $\log_2$-transformed fluorescence signal intensity distribution of antibodies bound to each purified protein on the multi-coronavirus protein microarray. Within each half-violin are three lines representing the interquartile range and the median. Above each split violin is the Wilcoxon rank sum P value, colored blue for significant P values below 0.05. The three panels are split by isotype (IgG, top; IgA, middle; IgM, bottom). Horizontal red dashed lines are drawn at the median of all signal intensities against purified proteins ($n = 14$) and peptides ($n = 587$) plus 1.0, i.e., double the global median; this threshold serves as a point of reference but not necessarily a seropositivity cutoff for each protein.

not included. The receptor binding domain (RBD) of the SARS-CoV-2 S protein had overall weaker antibody binding signals but, nevertheless, was significantly more reactive with COVID-19 patient serum IgG and IgA than with control serum IgG and IgA. The signals shown in Fig. 1 are the base 2 logarithm of the raw intensities without normalization, since the background reactivity of each purified protein is different. The SARS-CoV-2 S and RBD as well as the SARS-CoV S, N, and M purified proteins had the largest mean differences between IgG binding of the negative and positive groups, and the differences are the most statistically significant (t test P values of $<10^{-5}$). The same five antigens had the largest significant mean differences between IgA binding of the negative and positive groups. Only SARS-CoV-2 S and SARS-CoV N, however, had significant differential IgM binding between the COVID-19 patients and the control group. These results are in agreement with the enzyme-linked immunosorbent assays (ELISA) shown in Tables 2 and 3.

SARS-CoV-2 protein fragments identify antigenic regions. Nine SARS-CoV-2 full-length proteins were produced by coupled *in vitro* transcription and translation (IVTT)—S, envelope (E), M, N, and open reading frames (ORFs) 3a, 6, 7a, 8, and 10. We used the same technique to produce overlapping 100-amino acid (aa), 50-aa, and

30-aa fragments of each of these 9 SARS-CoV-2 proteins and to produce the structural proteins and some accessory proteins of HCoV-NL63, HCoV-OC43, SARS-CoV, and MERS-CoV. Using amino acid start and end positions of each fragment within the protein, differential IgG reactivity between the COVID-19 and healthy donor groups was mapped in a circular heatmap for the structural proteins (Fig. 2). This analysis allowed us to identify antigenic regions in each SARS-CoV-2 structural protein.

The SARS-CoV-2 N protein showed the strongest reactivity in its carboxy-terminal 100-aa fragment, as well as in 50-aa fragments covering the same region. This region was recognized by IgG, IgA, and IgM with significant differential reactivity between COVID-19-positive patients and the healthy negative-control group (Fig. 2; see Fig. S1 in the supplemental material). The middle of the N protein also had a region recognized by IgG and IgA identified by two 100-aa fragments. This central region of the N protein is the location of two mutations found in circulating strains of SARS-CoV-2, S235F found in the UK strain B.1.1.7 and T205I found in the South African strain B.1.351. Together these two antibody-reactive regions encompass about two-thirds of the N protein that likely contains at least two epitopes.

The S1 protein also showed greatest IgG binding near its carboxy terminus, in the penultimate 100-aa fragment, aa 501 to 600 (Fig. 2). This antigenic region of S1 was defined further by IgG and IgA reactivity with the 50-aa fragment from aa 551 to 600. This region is also the site of a mutation found in strain B.1.1.7, A570D, and another mutation that was fixed early in the pandemic and is present in nearly all circulating strains of SARS-CoV-2, D614G. The region containing the RBD was not strongly reactive when produced by IVTT, but significant IgG, IgA, and IgM reactivity was detected to the purified RBD fragment (Fig. 2 and Fig. S1). Moreover, the RBD is the site of mutations in several circulating strains of SARS-Cov-2—K417N in B.1.351 and K417T in the Brazilian strain P.1, as well as L452R in the California strain CAL20.C. The S2 protein of SARS-CoV-2 was highly antigenic with three regions of strong IgG, IgA, and IgM binding and differential reactivity with full-length, 100-aa and 50-aa fragments. Only the region near the carboxy terminus, however, was also reactive as a 30-aa fragment. This reactive 30-aa fragment, from aa 451 to 480 of S2 (1,136 to 1,165 of S), therefore likely defines a linear IgG epitope in this highly antigenic protein. Notably, an epitope in the central S2 antigenic region was differentially reactive for IgG and IgA but showed equal levels of IgM reactivity in 100-aa and 50-aa fragments, perhaps indicating a region of cross-reactivity for IgM produced by memory B cells reactive with an endemic human coronavirus. This central antigenic region of the S2 protein includes the site of a mutation in the B.1.1.7 strain, S297A (S982A of S).

An additional short epitope was found in the amino terminal 30-aa fragment of the SARS-CoV-2 M protein. This short fragment was highly reactive with COVID-19 patient serum IgG compared to healthy donor serum IgG, while larger fragments containing it, and the full-length M protein, were not as highly discriminatory for COVID-19 patient sera. The SARS-CoV-2 E protein had only one 30-aa fragment that showed low-level reactivity with IgA and IgM (Fig. S1), in both the COVID-19-positive and -negative groups.

The antigenic regions of SARS-CoV-2 structural proteins we identified did not correlate with homology between SARS-CoV-2 and other human coronaviruses (percentage amino acid sequence identity shown in outer track of Fig. 2). There was a moderate to high level of correlation between antibody reactivity with S2, N, and M proteins produced *in vitro*, particularly for IgG (Pearson's correlation coefficient shown in the inner links of Fig. 2). Less reactivity was seen in nonstructural proteins, but significant reactivity of COVID-19 patient IgG and IgA compared to that of negative-control IgG and IgA was e identified in fragments of the 3a and 7a accessory proteins (Fig. S2).

Individual antibody response profiles to antigenic regions of SARS-CoV-2 and other human coronaviruses. Individual IgG responses to the antigenic regions of SARS-CoV-2 proteins identified by reactivity with protein fragments varied substantially between individuals, as they did for the structural proteins of other human coronaviruses (Fig. 3). Individual variation is also evident in IgG responses against all the SARS-CoV-2
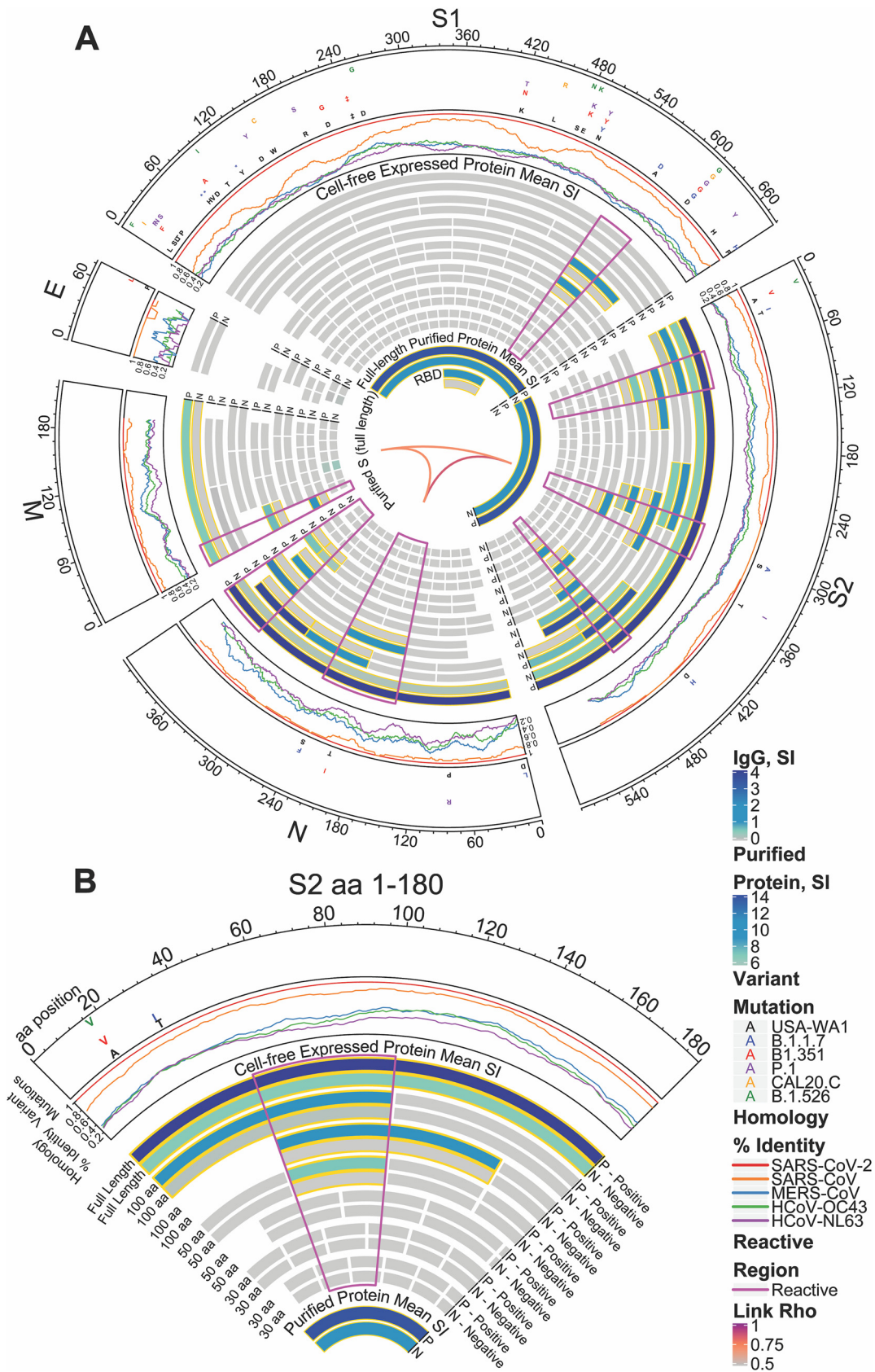
FIG 2 Reactivity of COVID-19 patient and healthy donor IgG to SARS-CoV-2 proteins and protein fragments. (A) The circular graphic maps the amino acid (aa) position of SARS-CoV-2 fragments, showing a heat map of IgG levels in each group for
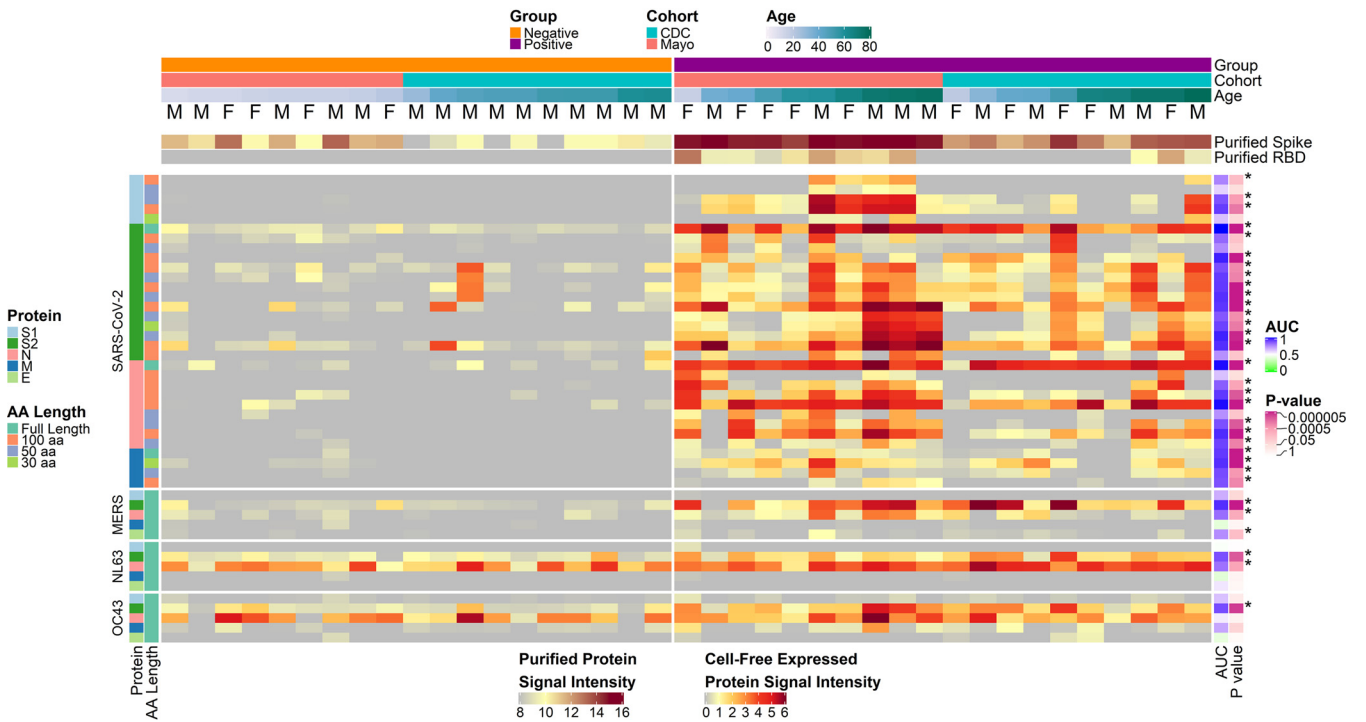
FIG 3 COVID-19-positive and -negative sample IgG reactivity to coronavirus proteins and protein fragments produced *in vitro*. The heatmaps present the signals of IgG binding to individual proteins and protein fragments within the antigenic regions of SARS-CoV-2, as well as the full-length structural proteins of MERS-CoV, HCoV-NL63, and HCoV-OC43 for individual samples. Columns represent serum samples ordered by increasing age within group and cohort, and rows represent proteins or protein fragments, including 32 SARS-CoV-2 proteins or fragments and 5 proteins each of MERS-CoV, HCoV-OC43, and HCoV-NL63. IgG signal intensity is shown on a color scale from gray to red. Sample information is overlaid above the heatmaps and includes sex (M/F), group (negative or positive), cohort (CDC or Mayo), and age (years). Protein/fragment information is annotated to the left of the heatmaps and includes the virus, the full-length protein name, and the amino acid length of the protein fragments ("AA Length," as full length, 100, 50, or 30 aa). The receiver operating characteristic area under the curve (AUC) and the unadjusted *t* test *P* value for each protein between negatives and positives are shown to the right of the heatmap. Asterisks next to the *P* values represent adjusted *P* values of <0.05.

S1, S2, N, and M protein fragments on the array (Fig. S3). Within the antigenic regions, some fragments, particularly 30-aa fragments, were nonreactive with COVID-19 patient sera, but others were reactive in a subset of individuals. Heterogeneity was higher and overall signal intensities were lower for IgA and IgM than for IgG (Fig. S4). There were no significant associations between age and sex with antibody levels in the positive group after adjustment for the false-discovery rate for any of the three isotypes (Table S1).

IgG from most negative controls and all COVID-19 patients was significantly reactive (normalized log$_2$ signal intensity, ≥1.0) with the HCoV-NL63 N protein; 17 of 19 negative sera and all 20 patient sera were reactive (proportion test *P* value of 0.4; Fig. 3). Similarly, most negative-control serum IgG and all patient serum IgG samples were significantly reactive with HCoV-OC43 N protein—15 of 19 negative sera and all 20 patient sera (*P* = 0.1). In contrast, IgG from only two control subjects reacted with the SARS-CoV-2 full-length N protein, while nearly all of the patients' serum IgG reacted—19 of 20 reactive (*P* = 6.8 × 10^{-7}).

FIG 2 Legend (Continued)

overlapping regions of different amino acid length. Proteins are indicated outside the circle plot, followed by a line graph showing the sequence homology of other CoVs with SARS-CoV-2 for each gene. Mutations, relative to the USA-WA1 strain, found in circulating mutant virus strains are shown in the next circular segment. Proteins and protein fragments produced *in vitro* are indicated by bars and show the length and position of each fragment in each protein. Each fragment is drawn twice and shows the group mean normalized log$_2$ signal intensity of IgG binding to each fragment for COVID-19-positive samples (P) and negative-control sera (N). The purified full-length S protein and the receptor binding domain (RBD) are shown for comparison. IgG signal intensity is shown by color gradient, from gray to blue. Bar pairs shown with a gold outline represent significantly differential IgG binding between COVID-19 patients and healthy controls, defined as a mean log$_2$ signal intensity of ≥0.1 in at least one group and a *t* test *P* value of ≤0.05. The regions of greatest reactivity for each protein are outlined in magenta. The Pearson's correlation coefficients ("Rho") between each full-length protein for IgG binding are shown as links between protein sectors in the center of the circle. (B) A slice of the circular graphic is amplified and labeled in more detail as a guide to interpreting the full figure. The first 180-aa sequence of S2 is shown.

Moreover, reactivity of the negative-control serum IgG with fragments of the SARS-CoV-2 N protein occurred rarely and exclusively in the C-terminal region of the protein (1 of 19 reactive), while COVID-19 patient serum IgG reacted frequently with fragments in the central region (12 of 20; $P = 2.1 \times 10^{-4}$) as well as the C-terminal region of the protein (19 of 20; $P = 1.3 \times 10^{-7}$).

HCoV S2 proteins were reactive with COVID-19 patient IgG at a much higher frequency than in the controls for both HCoV-NL63 (16 of 19 and 5 of 20 positives, respectively; $P = 2.4 \times 10^{-3}$) and HCoV-OC43 (18 of 19 and 4 of 20, respectively; $P = 5.9 \times 10^{-5}$). The higher frequencies in the COVID-19-positive subjects provide strong evidence of increased responses due to their exposure to SARS-CoV-2. Some negative-control subject's IgG reacted with the C-terminal (4 of 19) or central regions (4 of 19) of the SARS-CoV-2 S2 protein, but none reacted with the N-terminal region; this includes one individual who had unique reactivity to the 100-aa SARS-CoV-2 S2 fragment, 401 to 500 (S 986 to 1085; Fig. S3). By ELISA, this serum had a signal to threshold (S/T) ratio of 0.94, which was just below the positivity threshold of 1.0 and much higher than that of other healthy donor sera. This reactivity was unique among negative-control donors but did not directly translate to reactivity with OC43 or NL63 full-length S proteins. Overall, COVID-19 patient serum IgG reacted with the SARS-CoV-2 S2 protein C-terminal (19 of 20; $P = 1.3 \times 10^{-5}$), central (17 of 20; $P = 2.3 \times 10^{-4}$), and/or N-terminal (12 of 20; $P = 2.1 \times 10^{-4}$) 100-aa fragments much more frequently than prepandemic negative-control sera.

The reactivity of COVID-19 patient serum IgA compared to IgA of healthy donor sera was similar to results obtained for IgG. The IgA results had lower statistical significance than the IgG results, however, likely due to the lower concentration of IgA in serum compared to IgG. Nevertheless, many of the same proteins were the most differentially reactive with COVID-19 patient serum IgA compared to healthy donor serum IgA, including the N and S proteins and RBD of SARS-CoV-2 as well as the N, S, and M proteins of SARS-CoV with t test P values ranging from 2.1×10^{-6} to 1.1×10^{-3} (Fig. S4). The COVID-19 patient sera used in this study had less coronavirus reactive IgM than IgG or IgA, perhaps because the samples were obtained during the convalescent phase of disease. Nevertheless, significantly greater IgM reactivity was seen in patient sera compared to control donor sera for four proteins and two protein fragments produced $in\ vitro$ (Fig. S4). These were the N, S2, and M proteins of SARS-CoV-2, the MERS-CoV N protein, the carboxy-terminal 100-aa fragment of the SARS-CoV-2 N protein, and the amino terminal 30-aa fragment of the SARS-CoV-2 M protein.

A library of 587 peptides, 15 to 20 aa in length, from the epidemic SARS-CoV (covering S, N, M, and E proteins) and 2 endemic human coronaviruses (covering S protein) was printed on the multicoronavirus microarray at the same concentration as full-length purified recombinant proteins. The peptides, however, showed lower antibody reactivity than full-length proteins or protein fragments of 30, 50, or 100 aa (data not shown). Exceptionally, a single 17-aa peptide from HCoV-OC43 S protein with sequence CSKASSRSAIEDLLFDK spanning residues 905 to 921 had approximately 3.5-fold higher mean reactivity with COVID-19 patient sera ($P = 0.001$, not significant after adjustment for the false-discovery rate). This peptide mapped to the SARS-CoV-2 sequence PSKPSKRSFIEDLLFNK at residues 124 to 140 of S2 (809 to 825 of S protein) with identical residues in 12 of 17 positions.

To visualize the relative importance of antibody isotype binding in differentiating COVID-19-positive sera from negative sera, the samples were projected in two dimensions for each isotype using t-distributed stochastic neighbor embedding (tSNE; Fig. 4A), a nonlinear machine learning dimensionality reduction method which clusters together similar sets of multidimensional data. The 30 most reactive proteins for all isotypes were selected for this analysis to reduce the effect of differing isotype background levels that would be notable in low-reactivity proteins (Fig. 4B). Each of the isotypes clusters separately, but only IgG gave a clear delineation of positives and negatives (at $\sim$2.6 in tSNE dimension 2).
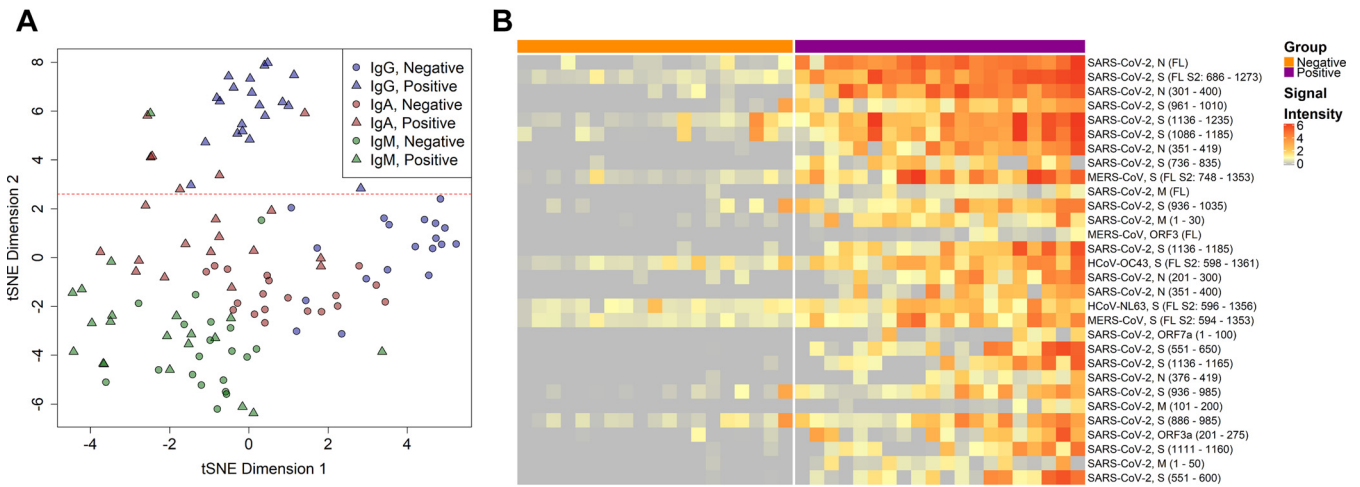
FIG 4 IgG responses give the best delineation of COVID-19 patient sera from healthy donor sera. (A) The IgG, IgA, and IgM responses against the 30 most reactive IVTT proteins by means of all samples and isotypes were projected for each sample across two dimensions using t-distributed stochastic neighbor embedding (tSNE). Points represent individual samples and are colored according to the isotype measurement. The shape represents the group in which the sample belonged, either the negative healthy donor group or the positive COVID-19 patient group. The horizontal red dashed line ($y = 2.6$) separates the IgG responses in negative and positive individuals. (B) The heatmap shows the 30 most differentially reactive IgG responses to IVTT proteins between the negative and positive groups. Columns represent serum samples, separated by group with colored headers. Rows represent full-length or fragmented proteins produced by cell-free expression *in vitro*. The protein annotations to the right of the heatmap denote the virus, protein, and in parentheses, the amino acid range of the fragment or full-length "FL" protein. Normalized $\log_2$ signal intensity is displayed on a gray to red color scale.

The full-length SARS-CoV-2 N and S2 proteins as well as several fragments of both proteins had the top nine largest mean differences in IgG reactivity between COVID-19 patients and healthy controls (Fig. 4B). These results were also statistically significant, with *t* test *P* values ranging from 2.1×10^{-6} to 4.3×10^{-2} (Table S1). Antibody responses to HCoV-NL63, HCoV-OC43, and MERS-CoV proteins were also among the 30 most discriminatory antigens for differentiating COVID-19 patients from control donors due to high reactivity with COVID-19-positive sera, while also demonstrating a considerable reactivity with negative sera. Nearly all the same epitopes and regions of reactivity found for IgG were recapitulated by IgA reactivity when reactivity to the overlapping 100-aa, 50-aa, and 30-aa protein fragments was analyzed (Fig. S4). This includes the epitopes mapped in the SARS-CoV-2 N, S1, S2, and M proteins (Fig. 2).

Correlation of SARS-CoV-2 and endemic human coronavirus responses. By comparing the correlation between antibody responses to the S2 and N proteins of SARS-CoV-2 with responses to the S2 and N proteins of endemic human coronaviruses, in both COVID-19-positive and -negative sera, we can estimate to what extent antibody responses to SARS-CoV-2 are the result of *de novo* immune responses or of boosting preexisting immunity. There were significantly stronger correlations between SARS-CoV-2 S2 protein IgG and HCoV-OC43 S2 proteins in the positive group (Pearson's correlation coefficient, $\rho = 0.6$) than the negative group ($\rho = 0.24$; Fig. 5A, top left). In the negative group, SARS-CoV-2 N protein IgG had no correlation with HCoV-OC43 N protein ($\rho = 0.02$) or HCoV-NL63 N protein ($\rho = 0.09$), whereas the correlations in the positive group were higher; HCoV-OC43 and HCoV-NL63 had a ρ value of 0.44 with SARS-CoV-2 N protein. These results suggest that *de novo* responses to SARS-CoV-2 S2 and N proteins are predominant. HCoV-OC43 and HCoV-NL63 N protein reactivity exhibited strong correlations in both positive and negative groups; $\rho = 0.54$ and $\rho = 0.62$, respectively. S2 protein reactivity correlations between these endemic human coronaviruses, however, were lower in the negative group than the positive group; $\rho = 0.29$ and $\rho = 0.49$, respectively. Further inspection of the IgG correlation scatterplot matrix (Fig. S5) showed an outlier sample in the CDC positive group for SARS-CoV-2 N protein, which had a normalized signal intensity of 0.65, 3 normalized signal intensity units lower than the next lowest sample. This had an outsized effect on SARS-CoV-2 N protein correlations; for example, removal of the sample increased correlation between
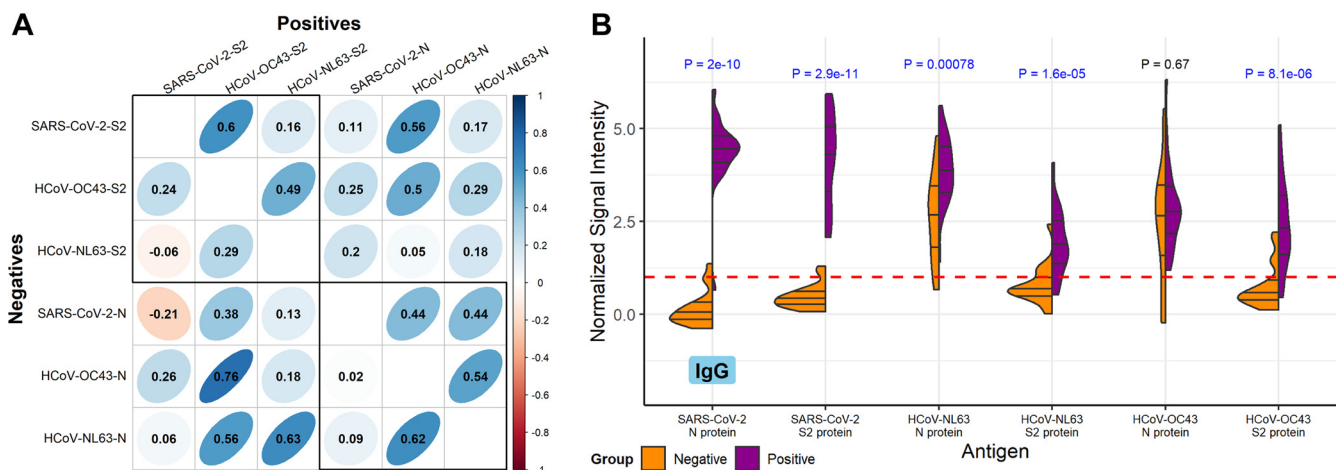
FIG 5 Correlation and concordance between IgG responses to SARS-CoV-2 and endemic human coronavirus N and S2 proteins. (A) The correlogram shows the Pearson's correlation coefficient (ρ) between IgG normalized signal intensity to SARS-CoV-2, HCoV-OC43, and HCoV-NL63 N and S2 full-length proteins produced *in vitro*. The lower half of the diagonal shows the correlation between reactivity of sera in the negative group, and the upper half of the diagonal shows the positive group serum correlations. The color scale indicates positive correlation in darker shades of blue and negative correlation in darker shades of red, and ρ is overlaid on each comparison. Additionally, the narrowness and slope of the ellipses represent increasing positive or negative correlation. Boxes are drawn around the intra-S2 and intra-N protein comparisons. (B) The split violin plot shows the normalized $\log_2$ IgG signal intensity distribution for each N and S2 protein produced *in vitro*. Within each half-violin are three lines representing the interquartile range and the median. Above each split violin is the Wilcoxon rank sum P value, colored blue for significant P values below 0.05. The red dashed line represents the 1.0 seropositivity cutoff.

SARS-CoV-2 N protein and HCoV-OC43 N protein among COVID-19-positive IgG, from 0.44 to 0.75 (data not shown). Differential IgG reactivity between the COVID-19-positive and -negative groups was also observed with the S2 and N proteins of SARS-CoV-2, HCoV-OC43, and HCoV-NL63. Positive COVID-19 patient sera had significantly higher IgG levels to S2 and N than the negative healthy donor sera for all three coronaviruses (Fig. 5B), with the exception of HCoV-OC43 N protein; this protein also showed higher IgA reactivity in the negatives (Fig. S6).

Correlation of multi-coronavirus protein microarray responses with ELISA and virus neutralization assays. S protein-based ELISA results from the CDC cohort, taken on all COVID-19 and healthy negative donor samples, were compared with IgG reactivity in the protein microarray data by Pearson's correlation coefficient for the highly reactive IVTT S2 protein ($\rho = 0.85$), IVTT N protein ($\rho = 0.9$), purified recombinant full-length S protein ($\rho = 0.88$), and purified recombinant RBD ($\rho = 0.85$), shown in Fig. 6A to D. The data clustered separately for negative responders and positive responders for all proteins. Virus neutralization titers were only available for the CDC COVID-19 patients and one healthy donor sample that tested near the 1.0 cutoff for ELISA reactivity ($n = 11$). In all cases, neutralization activity was low, with positive neutralization titers at dilution factors of 20 or 40. Despite the low values and few samples, a trend was observed using linear regression for IVTT S2 ($\beta = 6.5$, $P = 0.076$), IVTT N ($\beta = 6$, $P = 0.036$), and stabilized purified S ($\beta = 6.3$, $P = 0.077$) (Fig. 6E to H). There was no association, however, of neutralization activity with IgG reactivity to purified RBD ($\beta = 2.8$, $P = 0.27$). The linear regression models were specified with values of 0 for titers of <20. However, since the true titer is between 0 and 20, neutralization was also modeled as an ordinal variable using ordinal logistic regression. Similar results were obtained for IVTT S2 and stabilized purified S, whereas association with IVTT N protein was no longer significant. The complete correlation results for all proteins are shown in Table S1.

DISCUSSION

In this study of 20 COVID-19 patients, the strongest antibody responses to the SARS-CoV-2 proteins used on this array, for all antibody isotypes, were directed to the N and S2 proteins as has been previously seen in other studies (4, 7, 20, 21). We also detected antibody responses to S1, M, and accessory proteins 3a and 7a. Moreover, we
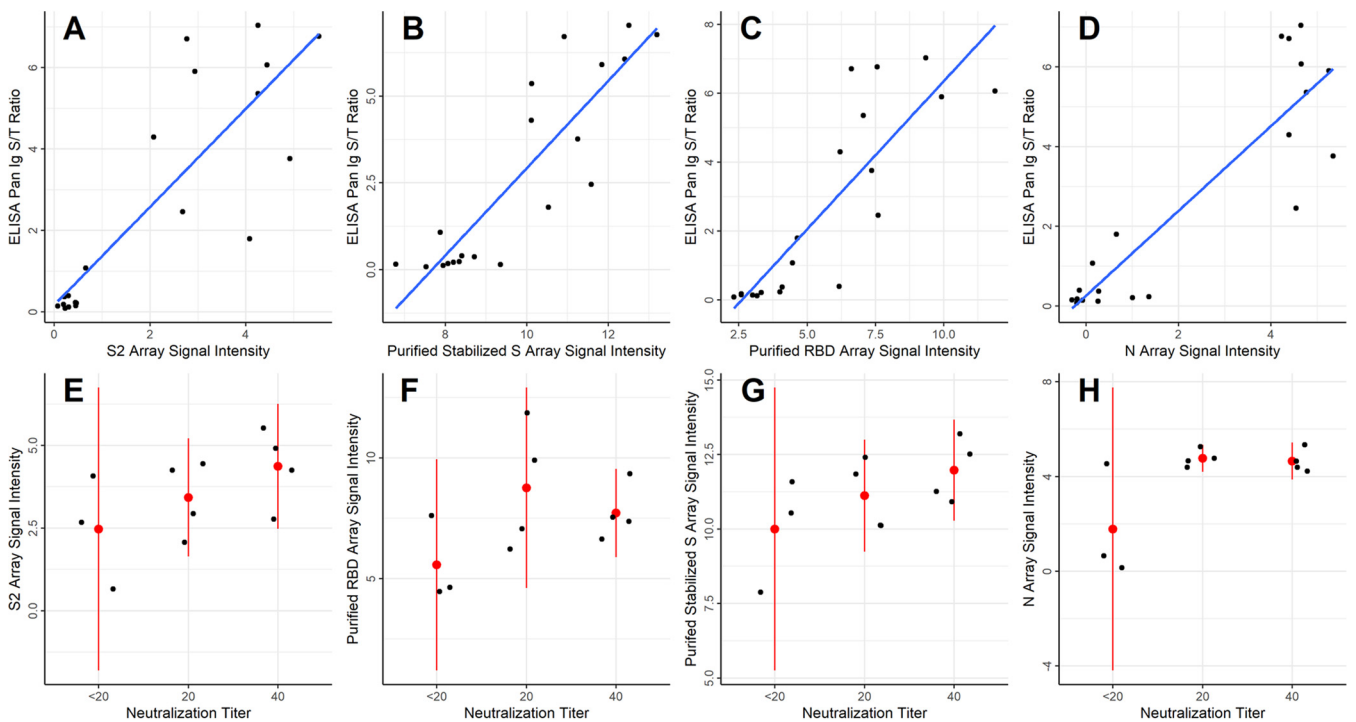
FIG 6 Correlation of IVTT and purified protein microarray results with ELISA and virus neutralization assays. (A to D) The scatterplots show the SARS-CoV-2 S protein-based ELISA Pan Ig signal/threshold ratio (y axis) plotted against the protein microarray $\log_2$ normalized IgG signal intensity for S2 and N proteins produced *in vitro*, as well as for the stabilized, purified full-length S protein and the purified RBD fragment of S1 protein (x axis), respectively. The blue lines were fit to the data using linear regression. (E to H) The dot plots show individual values of each patient for the protein microarray $\log_2$ normalized IgG binding intensity (y axis) of the four proteins shown in panels A to D at each neutralization titer (x axis). Red dots are plotted at the means of each stratum, and the red lines represent the 95% confidence intervals.

localized regions of each of these SARS-CoV-2 proteins to which antibodies bound, by antibody reactivity with overlapping protein fragments of three different lengths—100, 50, and 30 aa. Our results were internally consistent in that reactive proteins had more reactive fragments than nonreactive proteins and 100-aa reactive fragments contained reactive 50-aa fragments and sometimes they also contained reactive 30-aa fragments. We found little reactivity of COVID-19 patient sera with 13- to 20-aa peptides from SARS-CoV S, M, E, or N, HCoV-OC43 S, or HCoV-NL63 S with the exception of one S2 peptide from HCoV-OC43.

Many previous publications have predicted B cell epitopes in SARS-CoV-2 proteins using a variety of immunoinformatic approaches (22–25). Crooke et al. (22) predicted 26 potential linear B cell epitopes in the S protein, 14 potential epitopes in the N protein, and 3 potential epitopes in the M protein. We noted antibody reactivity with regions containing some, but not all, of these predicted epitopes. In particular, of the top six predicted B cell epitopes in the S protein we found significantly stronger reactivity with COVID-19 patient sera than with healthy donor sera for regions containing three epitopes—DIADTT (residues 568 to 573 near the carboxy terminus of S1), PPIKD (residues 792 to 796 near the amino terminus of S2), and VYDPLQPELDSF (residues 1137 to 1148 near the carboxy terminus of S2). The other three top predicted B cell epitopes of the S protein, residues 405 to 428, 440 to 450, and 496 to 507, were not in highly reactive regions of the S protein in our experiments, perhaps due to the overall low reactivity of the S1 protein except for its carboxy-terminal region or a need for native structure not found in protein fragments produced *in vitro*. Similarly, we found COVID-19-specific reactivity for regions including 9 of the 14 B cell epitopes in the N protein and 1 of 3 B cell epitopes in the M protein predicted by Crooke et al.

A few other groups have used protein or peptide arrays to map antibody reactivity to SARS-CoV-2 protein (21, 26–29). Two studies included full-length purified structural proteins from SARS-CoV-2, other human coronaviruses, and diverse human retroviruses (21,

26). Their results are consistent with ours but do not include accessory proteins or the ability to map reactive regions in each protein. Several groups used peptides to map epitopes in the SARS-CoV-2 S protein (27–29); Li et al. (27) found 4 epitopes defined by 12 amino acid peptides, 3 of which are in regions of antibody reactivity that we found. Poh et al. (28) found 2 epitopes defined by 18 amino acid peptides. Both are in regions of antibody reactivity that we described here. Finally, Zhang et al. (29) used 15 amino acid peptides overlapping by 5 aa covering the whole SARS-CoV-2 proteome, plus full-length N and E as well as 5 truncated forms of S to map IgM and IgG responses of acute COVID-19 patients (median 4 days post-onset of symptoms). They found more robust IgM responses than we did, since their specimens were collected earlier during infection. Zhang et al. identified five peptides as the most specific for COVID-19 patient IgG binding compared to controls, two in the S protein, two in N, and 1 in ORF-1ab. Both S protein peptides are in regions where we found IgG reactivity; one is in the N-terminal region of reactivity we found in S2, and the other is in the central reactivity region of S2. The N peptides of this group were not in a reactive region in our work, and we did not assay antibody reactivity of the ORF-1ab polyproteins.

Two groups published epitope maps of SARS-CoV-2 using phage display (30, 31). One group analyzed 56 aa and 20 aa fragments of the SARS-CoV-2 proteome, while the other group analyzed 38 aa fragments of the proteome. Both studies also included other human coronaviruses and used COVID-19 patient sera and control sera to identify specifically reactive epitopes in the SARS-CoV-2 proteome. Their data are largely in agreement with data presented here. Both studies found the greatest reactivity of COVID-19 patient sera in the S2 and N proteins. Moreover, the epitopes they mapped overlapped with the ones we found here by different methods.

SARS-CoV-2-infected subjects have much higher antibody levels to SARS-CoV-2 proteins, primarily N and S, than uninfected individuals, but it is also clear that even in the small sample sets evaluated here, some SARS-CoV-2-naive individuals have substantial preexisting antibody to some epitopes of these two proteins. These preexisting antibody levels have been shown to vary according to many different factors, including age, and may therefore have relevance to the clinical course of disease (32–34). Antibody reactivity of both positive and negative sera with endemic HCoVs and MERS-CoV was strongest for the S2 and N proteins as expected due to their abundance and conservation. Correlations between antibody responses to SARS-CoV-2 S2 and N proteins and to HCoV S2 and N were much stronger in COVID-19 convalescent-phase sera compared to negative sera, indicating that new cross-reactive responses to SARS-CoV-2 predominated over preexisting antibodies to HCoVs that cross-reacted with SARS-CoV-2 proteins.

The multi-coronavirus protein array is a tool that can help us improve our understanding of the immune response to SARS-CoV-2 and other coronaviruses. With these first two sets of convalescent-phase sera provided by the Mayo Clinic and the CDC, we have shown that SARS-CoV-2-naive subjects have clearly measurable cross-reactive antibody to the whole N and S2 proteins and that this reactivity is limited to specific epitopes. Importantly, there are epitopes that are more specific to SARS-CoV-2, that might serve as useful biomarkers of infection. Conversely, we have shown that infection with SARS-CoV-2 elicits or boosts the level of antibodies that bind to the N and S2 proteins of other coronaviruses, including SARS-CoV, MERS-CoV, HCoV-NL63, and HCoV-OC43.

The limitations of our study are the small sample size and the inclusion of only convalescent-phase samples. Despite these limitations, we identified clear differences in the antibody response from COVID-19 patients and healthy, nonexposed controls. The ideal data set to further investigate associations between preexisting antibody to specific epitopes and protection from severe disease would be longitudinal, with at least a preexposure sample, an acute-phase sample, and a convalescent-phase sample from each subject. Inclusion of samples from COVID-19 patients with a range of clinical symptoms will also provide an important comparison. In upcoming projects, we are

seeking to analyze these types of samples paired with detailed clinical data on disease outcomes ranging from asymptomatic to fatal to further improve our understanding of the complex role antibodies play in SARS-CoV-2 infection. It may also prove interesting to test convalescent plasma samples, especially given the variable efficacy results that have been reported in the literature (16–19). An assay providing more granular detail on the humoral response in these samples, such as the protein microarray described here, may provide valuable insights into factors that determine the effects of convalescent plasma treatment.

MATERIALS AND METHODS

Protein microarray analysis of serum samples. The first-generation multi-coronavirus protein microarray, produced by Antigen Discovery, Inc. (ADI, Irvine, CA, USA), included 935 full-length coronavirus proteins, overlapping 100-, 50-, and 30-aa protein fragments, and overlapping 13- to 20-aa peptides from SARS-CoV-2 (WA-1), SARS-CoV, MERS-CoV, HCoV-NL63, and HCoV-OC43. Purified proteins and peptides were obtained from BEI Resources. All these coronavirus proteins were produced in *Escherichia coli* except the SARS-CoV-2 and SARS-CoV S proteins, which were made in Sf9 insect cells, and the SARS-CoV-2 RBD, made in HEK-293 cells. Other proteins and protein fragments were expressed using an *E. coli in vitro* transcription and translation (IVTT) system (rapid translation system; Biotechrabbit, Berlin, Germany) and printed onto nitrocellulose-coated glass AVID slides (Grace Bio-Labs, Inc., Bend, OR, USA) using an Omni Grid Accent robotic microarray printer (Digilabs, Inc., Marlborough, MA, USA). Microarrays were probed with sera and antibody binding detected by incubation with fluorochrome-conjugated goat anti-human IgG, IgA, or IgM (Jackson ImmunoResearch, West Grove, PA, USA, or Bethyl Laboratories, Inc., Montgomery, TX, USA). Slides were scanned on a GenePix 4300A high-resolution microarray scanner (Molecular Devices, Sunnyvale, CA, USA), and raw spot and local background fluorescence intensities, spot annotations, and sample phenotypes were imported and merged in R (35), in which all subsequent procedures were performed. Foreground spot intensities were adjusted by subtraction of local background, and negative values were converted to 1. All foreground values were transformed using the base 2 logarithm. The data set was normalized to remove systematic effects by subtracting the median signal intensity of the IVTT controls for each sample. With the normalized data, a value of 0.0 means that the intensity is no different than the background, and a value of 1.0 indicates a doubling with respect to background. For full-length purified recombinant proteins and peptide libraries, the raw signal intensity data were transformed using the base 2 logarithm for analysis.

Control sera and COVID-19 patient samples. COVID-19-positive and pre-COVID-19 negative-control sera provided by the CDC were acquired from commercial laboratories or through partnership with Emory University. Samples were provided with only clinical and demographic information retained. The majority of samples (7/10) were from patients that were not hospitalized, with blood collected between 26 and 60 days post-symptom onset. Negative-control sera were collected pre-COVID-19, in the fall of 2019. This activity was reviewed by the CDC and was conducted consistently with applicable federal law and CDC policy (45 C.F.R. part 46, 21 C.F.R. part 56). The COVID-19-positive samples provided by the Mayo Clinic were deidentified residual sera from clinical testing with only age and sex information available. The COVID-19-negative samples were collected pre-COVID-19 pandemic, between 2005 and 2012. These samples were from participants in prior Mayo Clinic vaccine studies who had provided informed consent for future use of their biospecimens. The original blood collection was collected through Mayo Clinic IRB-approved protocols. Samples were tested for SARS-CoV-2-specific antibodies and the presence of neutralizing antibodies as described below.

Enzyme linked immunosorbent assay (ELISA). The CDC provided samples were tested using an enzyme-linked immunosorbent assay (ELISA) against the prefusion stabilized ectodomain of SARS-CoV-2 spike protein (36). This validated assay has been shown to have sensitivity and specificity of 96% and 99%, respectively (37). Briefly, plates were coated with purified spike protein and incubated overnight at 4°C followed by 37°C incubation steps and subsequent phosphate-buffered saline + 0.05% Tween 20 (PBST) washings with 2.5× Stabilcoat blocker (Surmodics), 1:25 to 1:1,600 diluted serum in 1× PBST + 5% skim milk for 1 h, 1:2,000 goat anti-human Ab conjugated to horseradish peroxidase (KPL) for 1 h, and ABTS peroxidase substrate for 30 min. Reactions were then quenched with stop solution. Plates were read at 405 nm and 490 nm, with resulting optical densities (ODs) calculated as 490 nm to 405 nm absorbance for each sample and minus PBS-only-coated wells. Results are reported as a ratio of the calculated sample OD/cutoff threshold OD (signal/threshold, or S/T); values of >1.0 are defined as positive. The Mayo Clinic COVID-19-positive samples were tested using an IgG SARS-CoV-2 spike protein-specific ELISA (EuroImmune, Inc.) performed according to the manufacturer's recommendations. This validated assay has been shown to have sensitivity and specificity of 90% and 100%, respectively (38). Results are reported as a ratio of the sample OD/calibrator OD (signal/calibrator, or S/C); values of >1.1 are defined as positive.

Neutralization assay. All SARS-CoV-2 microneutralization assays (MNT) were performed following biosafety level-3 precautions, using a SARS-CoV-2 clinical isolate. The WA1 strain of SARS-CoV-2 was employed using a modified version of a previously established protocol. A total of 27 Vero cell suspensions (ATCC CCL-81) were prepared at 2.2×10^5 to 2.5×10^5 cells/ml in Dulbecco modified Eagle medium (DMEM; Thermo Fisher; catalog no. 11965118) plus 10% fetal bovine serum (FBS, defined; HyClone; catalog no. SH30070.03) (heat-inactivated at 56°C for 30 min) plus 2× antibiotic-antimycotic (Thermo Fisher; catalog no. 15240062) plus 2× penicillin-streptomycin (Thermo Fisher; catalog no. 15140122)

immediately before use. Sera were 2-fold serial diluted in serum-free DMEM in a 96-well flat-bottom plate, from 1:10 to 1:320, in triplicate, to a final volume of 50 μl/well. Then 50 μl SARS-CoV-2 was added to each well, such that final serum dilution titers ranged from 1:20 to 1:640. After a 30-min incubation at 37°C and 5% CO_2, 100 μl of Vero cells in suspension were added to each well, for a final concentration of 2.2 to 2.5 $\times$ 104 cells/well. After 5 days cells were stained and fixed with crystal violet fixative (0.15% crystal violet, 2.5% ethanol, 11% formaldehyde, 50% PBS, 0.01 M pH 7.4). The endpoint concentration at which antibodies were determined to be neutralizing for SARS-CoV-2 infection was the lowest concentration of antibody at which 3 replicate wells were protected against virus infection.

Statistical analysis. Student's t tests were used for comparison of the individual antibody response means between the positive and negative groups. Comparison of the medians was done using Wilcoxon's rank sum test. The area under the receiver operating characteristics curve (AUC) was calculated to estimate delineation of groups for each antigen. The tSNE analyses were calculated after 25,000 iterations with a perplexity parameter of 30 using the R package Rtsne (39). Comparisons of the proportions of responders to each protein between groups was done using two-proportion z tests implemented by the prop.test function in R. Correlation between antibody features and between protein microarray and ELISA measurements used Pearson's correlation coefficient (ρ), and association between antibody measurements and sample information such as sex, age, and cohort were modeled using linear regression. The association of specific antibody responses with virus neutralization titers was estimated using linear regression, with the values below detection levels ($<$20) coded as zero, or by converting neutralization titers to ordinal values and estimating the proportional odds ratio by ordinal logistic regression, whereby P values were estimated by comparing the t value against the standard normal distribution. Adjustment for the false-discovery rate was performed using the p.adjust function in R (40). Data visualization was performed using the circlize (41), ComplexHeatmap (42), ggplot2, heatmap2, and corrplot (38) packages in R. Unadjusted P values were shown in graphics.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
SUPPLEMENTAL FILE 1, PDF file, 2.1 MB.

REFERENCES

1. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020.
2. Kissler S, Tedijanto C, Goldstein E, Grad Y, Lipsitch M. 2020. Projecting transmission dynamics of SARS-CoV-2 through the postpandemic period. Science 368:860–868. https://doi.org/10.1126/science.abb5793.
3. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, Cuomo-Dannenburg G, Thompson H, Walker PGT, Fu H, Dighe A, Griffin JT, Baguelin M, Bhatia S, Boonyasiri A, Cori A, Cucunubá Z, FitzJohn R, Gaythorpe K, Green W, Hamlet A, Hinsley W, Laydon D, Nedjati-Gilani G, Riley S, van Elsland S, Volz E, Wang H, Wang Y, Xi X, Donnelly CA, Ghani AC, Ferguson NM. 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet

Infect Dis 20:669–677. https://doi.org/10.1016/S1473-3099(20)30243-7.

4. Long Q-X, Liu B-Z, Deng H-J, Wu G-C, Deng K, Chen Y-K, Liao P, Qiu J-F, Lin Y, Cai X-F, Wang D-Q, Hu Y, Ren J-H, Tang N, Xu Y-Y, Yu L-H, Mo Z, Gong F, Zhang X-L, Tian W-G, Hu L, Zhang X-X, Xiang J-L, Du H-X, Liu H-W, Lang C-H, Luo X-H, Wu S-B, Cui X-P, Zhou Z, Zhu M-M, Wang J, Xue C-J, Li X-F, Wang L, Li Z-J, Wang K, Niu C-C, Yang Q-J, Tang X-J, Zhang Y, Liu X-M, Li J-J, Zhang D-C, Zhang F, Liu P, Yuan J, Li Q, Hu J-L, Chen J, et al. 2020. Antibody responses to SARS-CoV-2 in patients with COVID-19. Nat Med 26:845–848. https://doi.org/10.1038/s41591-020-0897-1.

5. Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, Wang X, Yuan J, Li T, Li J, Qian S, Hong C, Wang F, Liu Y, Wang Z, He Q, Li Z, He B, Zhang T, Fu Y, Ge S, Liu L, Zhang J, Xia N, Zhang Z. 2020. Antibody responses to SARS-CoV-2 in patients with novel coronavirus disease 2019. Clin Infect Dis 71:2027–2034. https://doi.org/10.1093/cid/ciaa344.

6. Yu H-Q, Sun B-Q, Fang Z-F, Zhao J-C, Liu X-Y, Li Y-M, Sun X-Z, Liang H-F, Zhong B, Huang Z-F, Zheng P-Y, Tian L-F, Qu H-Q, Liu D-C, Wang E-Y, Xiao X-J, Li S-Y, Ye F, Guan L, Hu D-S, Hakonarson H, Liu Z-G, Zhong N-S. 2020. Distinct features of SARS-CoV-2-specific IgA response in COVID-19 patients. Eur Respir J 56:2001526. https://doi.org/10.1183/13993003.01526-2020.

7. Guo L, Ren L, Yang S, Xiao M, Chang D, Yang F, Dela Cruz CS, Wang Y, Wu C, Xiao Y, Zhang L, Han L, Dang S, Xu Y, Yang Q-W, Xu S-Y, Zhu H-D, Xu Y-C, Jin Q, Sharma L, Wang L, Wang J. 2020. Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). Clin Infect Dis 71:778–785. https://doi.org/10.1093/cid/ciaa310.

8. Chandrashekar A, Liu J, Martinot AJ, McMahan K, Mercado NB, Peter L, Tostanoski LH, Yu J, Maliga Z, Nekorchuk M, Busman-Sahay K, Terry M, Wrijil LM, Ducat S, Martinez DR, Atyeo C, Fischinger S, Burke JS, Slein MD, Pessaint L, Van Ry A, Greenhouse J, Taylor T, Blade K, Cook A, Finneyfrock B, Brown R, Teow E, Velasco J, Zahn R, Wegmann F, Abbink P, Bondzie EA, Dagotto G, Gebre MS, He X, Jacob-Dolan C, Kordana N, Li Z, Lifton MA, Mahrokhian SH, Maxfield LF, Nityanandam R, Nkolola JP, Schmidt AG, Miller AD, Baric RS, Alter G, Sorger PK, Estes JD, et al. 2020. SARS-CoV-2 infection protects against rechallenge in rhesus macaques. Science 369:812–817. https://doi.org/10.1126/science.abc4776.

9. Deng W, Bao L, Liu J, Xiao C, Liu J, Xue J, Lv Q, Qi F, Gao H, Yu P, Xu Y, Qu Y, Li F, Xiang Z, Yu H, Gong S, Liu M, Wang G, Wang S, Song Z, Liu Y, Zhao W, Han Y, Zhao L, Liu X, Wei Q, Qin C. 2020. Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. Science 369:818–823. https://doi.org/10.1126/science.abc5343.

10. Atyeo C, Fischinger S, Zohar T, Slein MD, Burke J, Loos C, McCulloch DJ, Newman KL, Wolf C, Yu J, Shuey K, Feldman J, Hauser BM, Caradonna T, Schmidt AG, Suscovich TJ, Linde C, Cai Y, Barouch D, Ryan ET, Charles RC, Lauffenburger D, Chu H, Alter G. 2020. Distinct early serological signatures track with SARS-CoV-2 survival. Immunity 53:524–532.e4. https://doi.org/10.1016/j.immuni.2020.07.020.

11. McAndrews KM, Dowlatshahi DP, Dai J, Becker LM, Hensel J, Snowden LM, Leveille JM, Brunner MR, Holden KW, Hopkins NS, Harris AM, Kumpati J, Whitt MA, Lee JJ, Ostrosky-Zeichner LL, Papanna R, LeBleu VS, Allison JP, Kalluri R. 2020. Heterogeneous antibodies against SARS-CoV-2 spike receptor binding domain and nucleocapsid with implications for COVID-19 immunity. JCI Insight 5:e142386. https://doi.org/10.1172/jci.insight.142386.

12. Yang ZY, Werner HC, Kong WP, Leung K, Traggiai E, Lanzavecchia A, Nabel GJ. 2005. Evasion of antibody neutralization in emerging severe acute respiratory syndrome coronaviruses. Proc Natl Acad Sci U S A 102:797–801. https://doi.org/10.1073/pnas.0409065102.

13. Jaume M, Yip MS, Cheung CY, Leung HL, Li PH, Kien F, Dutry I, Callendret B, Escriou N, Altmeyer R, Nal B, Daëron M, Bruzzone R, Peiris JSM. 2011. Anti-severe acute respiratory syndrome coronavirus spike antibodies trigger infection of human immune cells via a pH- and cysteine protease-independent FcgammaR pathway. J Virol 85:10582–10597. https://doi.org/10.1128/JVI.00671-11.

14. Yip MS, Leung NHL, Cheung CY, Li PH, Lee HHY, Daëron M, Peiris JSM, Bruzzone R, Jaume M. 2014. Antibody-dependent infection of human macrophages by severe acute respiratory syndrome coronavirus. Virol J 11:82. https://doi.org/10.1186/1743-422X-11-82.

15. Wang Q, Zhang L, Kuwahara K, Li L, Liu Z, Li T, Zhu H, Liu J, Xu Y, Xie J, Morioka H, Sakaguchi N, Qin C, Liu G. 2016. Immunodominant SARS Coronavirus epitopes in humans elicited both enhancing and neutralizing effects on infection in non-human primates. ACS Infect Dis 2:361–376. https://doi.org/10.1021/acsinfecdis.6b00006.

16. Liu STH, Lin H-M, Baine I, Wajnberg A, Gumprecht JP, Rahman F, Rodriguez D, Tandon P, Bassily-Marcus A, Bander J, Sanky C, Dupper A, Zheng A, Nguyen FT, Amanat F, Stadlbauer D, Altman DR, Chen BK, Krammer F, Mendu DR, Firpo-Betancourt A, Levin MA, Bagiella E, Casadevall A, Cordon-Cardo C, Jhang JS, Arinsburg SA, Reich DL, Aberg JA, Bouvier NM. 2020. Convalescent plasma treatment of severe COVID-19: a propensity score–matched control study. Nat Med 26:1708–1713. https://doi.org/10.1038/s41591-020-1088-9.

17. Wang M, Wu T, Zuo Z, You Y, Yang X, Pan L, Hu Y, Luo X, Jiang L, Xia Z, Deng M. 2021. Evaluation of current medical approaches for COVID-19: a systematic review and meta-analysis. BMJ Support Palliat Care 11:45–52. https://doi.org/10.1136/bmjspcare-2020-002554.

18. Bakhtawar N, Usman M, Khan MMU. 2020. Convalescent plasma therapy and its effects On COVID-19 patient outcomes: a systematic review of current literature. Cureus 12:e9535. https://doi.org/10.7759/cureus.9535.

19. Wooding DJ, Bach H. 2020. Treatment of COVID-19 with convalescent plasma: lessons from pastcoronavirus outbreaks. Clin Microbiol Infect 26:1436–1446. https://doi.org/10.1016/j.cmi.2020.08.005.

20. Brouwer PJM, Caniels TG, van der Straten K, Snitselaar JL, Aldon Y, Bangaru S, Torres JL, Okba NMA, Claireaux M, Kerster G, Bentlage AEH, van Haaren MM, Guerra D, Burger JA, Schermer EE, Verheul KD, van der Velde N, van der Kooi A, van Schooten J, van Breemen MJ, Bijl TPL, Sliepen K, Aartse A, Derking R, Bontjer I, Kootstra NA, Wiersinga WJ, Vidarsson G, Haagmans BL, Ward AB, de Bree GJ, Sanders RW, van Gils MJ. 2020. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. Science 369:643–650. https://doi.org/10.1126/science.abc5902.

21. de Assis RR, Jain A, Nakajima R, Jasinskas A, Felgner J, Obiero JM, Adenaiye O, Tai S, Hong F, Norris PJ, Stone M, Simmons G, Bagri A, Schreiber M, Buser A, Holbro A, Battegay M, Hosimer P, Noesen C, Milton DK, Davies DH, Contestable P, Corash LM, Busch MP, Felgner PL, Khan S, Prometheus Study Group. 2020. Analysis of SARS-CoV-2 antibodies in COVID-19 convalescent blood using a coronavirus antigen microarray. bioRxiv. 2020.04.15.043364.

22. Crooke SN, Ovsyannikova IG, Kennedy RB, Poland GA. 2020. Immunoinformatic identification of B cell and T cell epitopes in the SARS-CoV-2 proteome. Sci Rep 10:14179. https://doi.org/10.1038/s41598-020-70864-8.

23. Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B, Sette A. 2020. A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. Cell Host Microbe 27:671–680. https://doi.org/10.1016/j.chom.2020.03.002.

24. Feng Y, Qiu M, Liu L, Zou S, Li Y, Luo K, Guo Q, Han N, Sun Y, Wang K, Zhuang X, Zhang S, Chen S, Mo F. 2020. Multi-epitope vaccine design using an immunoinformatics approach for 2019 novel coronavirus (SARS-CoV-2). Biorxiv. https://doi.org/10.1101/2020.03.03.962332.

25. Fast E, Altman R, Chen B. 2020. Potential T-cell and B-cell epitopes of 2019-nCoV. Biorxiv. https://doi.org/10.1101/2020.02.19.955484.

26. Khan S, Nakajima R, Jain A, de Assis RR, Jasinskas A, Obiero JM, Adenaiye O, Tai S, Hong F, Milton DK, Davies H, Felgner PL, Prometheus Study Group. 2020. Analysis of serologic cross-reactivity between common human coronaviruses and SARS-CoV-2 using coronavirus antigen microarray. Biorxiv. https://doi.org/10.1101/2020.03.24.006544.

27. Li Y, Lai D-Y, Zhang H-N, Jiang H-W, Tian X, Ma M-L, Qi H, Meng Q-F, Guo S-J, Wu Y, Wang W, Yang X, Shi D-W, Dai J-B, Ying T, Zhou J, Tao S-C. 2020. Linear epitopes of SARS-CoV-2 spike protein elicit neutralizing antibodies in COVID-19 patients. Cell Mol Immunol 17:1095–1097. https://doi.org/10.1038/s41423-020-00523-5.

28. Poh CM, Carissimo G, Wang B, Amrun SN, Lee CY-P, Chee RS-L, Fong S-W, Yeo NK-W, Lee W-H, Torres-Ruesta A, Leo Y-S, Chen MI-C, Tan S-Y, Chai LYA, Kalimuddin S, Kheng SSG, Thien S-Y, Young BE, Lye DC, Hanson BJ, Wang C-I, Renia L, Ng LFP. 2020. Two linear epitopes on the SARS-CoV-2 spike protein that elicit neutralising antibodies in COVID-19 patients. Nat Commun 11:2806. https://doi.org/10.1038/s41467-020-16638-2.

29. Zhang X, Wu X, Wang D, Lu M, Hou X, Wang H, Liang T, Dai J, Duan H, Xu Y, Li Y, Yu X. 2020. Proteome-wide analysis of differentially-expressed SARS-CoV-2 antibodies in early COVID-19 infection. Biorxiv. https://doi.org/10.1101/2020.04.14.20064535.

30. Shrock E, Fujimura E, Kula T, Timms RT, Lee I-H, Leng Y, Robinson ML, Sie BM, Li MZ, Chen Y, Logue J, Zuiani A, McCulloch D, Lelis FJN, Henson S, Monaco DR, Travers M, Habibi S, Clarke WA, Caturegli P, Laeyendecker O, Piechocka-Trocha A, Li JZ, Khatri A, Chu HY, Villani A-C, Kays K, Goldberg MB, Hacohen N, Filbin MR, Yu XG, Walker BD, Wesemann DR, Larman HB, Lederer JA, Elledge SJ, MGH COVID-19 Collection & Processing Team. 2020. Viral epitope profiling of COVID-19 patients reveals cross-reactivity and correlates of severity. Science 370:eabd4250. https://doi.org/10.1126/science.abd4250.

31. Zamecnik CR, Rajan JV, Yamauchi KA, Mann SA, Loudermilk RP, Sowa GM, Zorn KC, Alvarenga BD, Gaebler C, Caskey M, Stone M, Norris PJ, Gu W, Chiu CY, Ng D, Byrnes JR, Zhou XX, Wells JA, Robbiani DF, Nussenzweig MC, DeRisi JL, Wilson MR. 2020. ReScan, a multiplex diagnostic pipeline, pans human sera for SARS-CoV-2 antigens. Cell Rep Med 1:100123. https://doi.org/10.1016/j.xcrm.2020.100123.

32. Aydillo T, Rombauts A, Stadlbauer D, Aslam S, Abelenda-Alonso G, Escalera A, Amanat F, Jiang K, Krammer F, Carratala J, García-Sastre A. 2021. Immunological imprinting of the antibody response in COVID-19 patients. Nat Commun 12:3781. https://doi.org/10.1038/s41467-021-23977-1.

33. Yang F, Nielsen SCA, Hoh RA, Röltgen K, Wirz OF, Haraguchi E, Jean GH, Lee JY, Pham TD, Jackson KJL, Roskin KM, Liu Y, Nguyen K, Ohgami RS, Osborne EM, Nadeau KC, Niemann CU, Parsonnet J, Boyd SD. 2021. Shared B cell memory to coronaviruses and other pathogens varies in human age groups and tissues. Science 372:738–741. https://doi.org/10.1126/science.abf6648.

34. Chaudhury S, Hutter J, Bolton JS, Hakre S, Mose E, Wooten A, O'Connell W, Hudak J, Krebs SJ, Darden JM, Regules JA, Murray CK, Modjarrad K, Scott P, Peel S, Bergmann-Leitner ES. 2021. Serological profiles of pan-coronavirus-specific responses in COVID-19 patients using a multiplexed electrochemi-luminescence-based testing platform. PLoS One 16:e0252628. https://doi.org/10.1371/journal.pone.0252628.

35. R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

36. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O, Graham BS, McLellan JS. 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367:1260–1263. https://doi.org/10.1126/science.abb2507.

37. Freeman B, Lester S, Mills L, Rasheed M, Moye S, Abiona O, Hutchinson GB, Morales-Betoulle M, Krapinunaya I, Gibbons A, Chiang C-F, Cannon D, Klena J, Johnson JA, Owen SM, Graham BS, Corbett KS, Thornburg NJ. 2020. Validation of a SARS-CoV-2 spike protein ELISA for use in contact investigations and serosurveillance. bioRxiv. https://doi.org/10.1101/2020.04.24.057323.

38. FDA. 2021. EUA authorized serology test performance. https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/eua-authorized-serology-test-performance.

39. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B 57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

40. Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. circlize implements and enhances circular visualization in R. Bioinformatics 30:2811–2812. https://doi.org/10.1093/bioinformatics/btu393.

41. Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32:2847–2849. https://doi.org/10.1093/bioinformatics/btw313.

42. Friendly M. 2002. Corrgrams: exploratory displays for correlation matrices. Am Stat 56:316–324. https://doi.org/10.1198/000313002533.