Check for updates

SARS-CoV-2 immune repertoire in MIS-C and pediatric COVID-19

Supriya Ravichandran[1,6], Juanjie Tang[1,6], Gabrielle Grubbs[1,6], Youri Lee[1,6], Sara Pourhashemi[1,6], Laila Hussaini[2,3], Stacey A. Lapp[2,3], Robert C. Jerris[4], Vidisha Singh[2], Ann Chahroudi[2,3], Evan J. Anderson[2,3,5], Christina A. Rostad[2,3] and Surender Khurana[1✉]

There is limited understanding of the viral antibody fingerprint following severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in children. Herein, SARS-CoV-2 proteome-wide immunoprofiling of children with mild/moderate or severe coronavirus disease 2019 (COVID-19) versus multisystem inflammatory syndrome in children versus hospitalized control patients revealed differential cytokine responses, IgM/IgG/IgA epitope diversity, antibody binding and avidity. Apart from spike and nucleocapsid, IgG/IgA recognized epitopes in nonstructural protein (NSP) 2, NSP3, NSP12–NSP14 and open reading frame (ORF) 3a–ORF9. Peptides representing epitopes in NSP12, ORF3a and ORF8 demonstrated SARS-CoV-2 serodiagnosis. Antibody-binding kinetics with 24 SARS-CoV-2 proteins revealed antibody parameters that distinguish children with mild/moderate versus severe COVID-19 or multisystem inflammatory syndrome in children. Antibody avidity to prefusion spike correlated with decreased illness severity and served as a clinical disease indicator. The fusion peptide and heptad repeat 2 region induced SARS-CoV-2-neutralizing antibodies in rabbits. Thus, we identified SARS-CoV-2 antibody signatures in children associated with disease severity and delineate promising serodiagnostic and virus neutralization targets. These findings might guide the design of serodiagnostic assays, prognostic algorithms, therapeutics and vaccines in this important but understudied population.

SARS-CoV-2 has resulted in the global pandemic of COVID-19 that causes severe disease most often in adults and elderly individuals with comorbidities. In children, SARS-CoV-2 infection is often asymptomatic or causes mild disease; however, children are susceptible to develop severe disease manifestations of COVID-19 and the associated postinfectious multisystem inflammatory syndrome in children (MIS-C). Although the adult COVID-19 immune response has been studied[1–4], limited knowledge exists regarding the SARS-CoV-2 immune response in children with COVID-19 or MIS-C, especially against NSPs and accessory proteins of SARS-CoV-2 and their association with disease severity. Previous studies in children evaluated immune responses primarily focused on spike or nucleocapsid (N) antibodies but have provided limited insight into the diversity and quality of polyclonal antibody responses across the complete SARS-CoV-2 proteome[3,5–8]. These studies did not identify epitope specificity against the complete viral proteome, nor did they identify antibody kinetic parameters that correlate with disease severity or potential serodiagnostic candidates in these children. Moreover, the contribution of different antigenic sites in SARS-CoV-2 spike protein as an immunogen or vaccine in neutralization against SARS-CoV-2 is unknown.

Therefore, an in-depth understanding of the nature of the immune response to SARS-CoV-2 in hospitalized children is critical to deciphering host–pathogen interaction as well as identification and development of serologic and effective therapeutic/neutralizing targets for reducing SARS-CoV-2-induced disease in this understudied population.

Results

Immunological signatures of children with COVID-19 or MIS-C.
A total of 74 hospitalized children (mean age, 9.4 + 6.1 years) were enrolled (Fig. 1a and Supplementary Table 1) at Children's Healthcare of Atlanta (CHOA), which included 25 children meeting criteria for MIS-C, 35 with acute COVID-19, according to the Centers for Disease Control and Prevention (CDC) case definition, and 14 hospitalized controls (Fig. 1b and Supplementary Table 1). Viral RNA levels were determined by quantitative PCR with reverse transcription (RT–qPCR) using nasopharyngeal swabs (Supplementary Table 2). Demographic, clinical, laboratory, imaging, treatment and outcome data are summarized in Supplementary Table 1 and detailed clinical data for each individual is provided in the Supplementary Table 3. Patients were stratified into severe (S) or mild/moderate (M) disease groups based on admission to an intensive care unit (ICU) or not, respectively. The disease-severity classification of the patients correlated with the maximum pediatric sequential organ failure assessment (pSOFA) scores (Fig. 1a and Supplementary Table 3).

Sera were collected upon study enrollment, with all days being relative to days post onset of symptoms (DPoS) that prompted hospitalization. Cytokine and chemokine analyses as well as in-depth antibody profiling were performed on the sera collected at study entry from these 60 children with COVID-19 or MIS-C. The controls were 14 hospitalized children (5 patients with severe illness in ICU and 9 non-ICU patients with mild illness) who were SARS-CoV-2 negative by RT–qPCR and by SARS-CoV-2 nucleocapsid
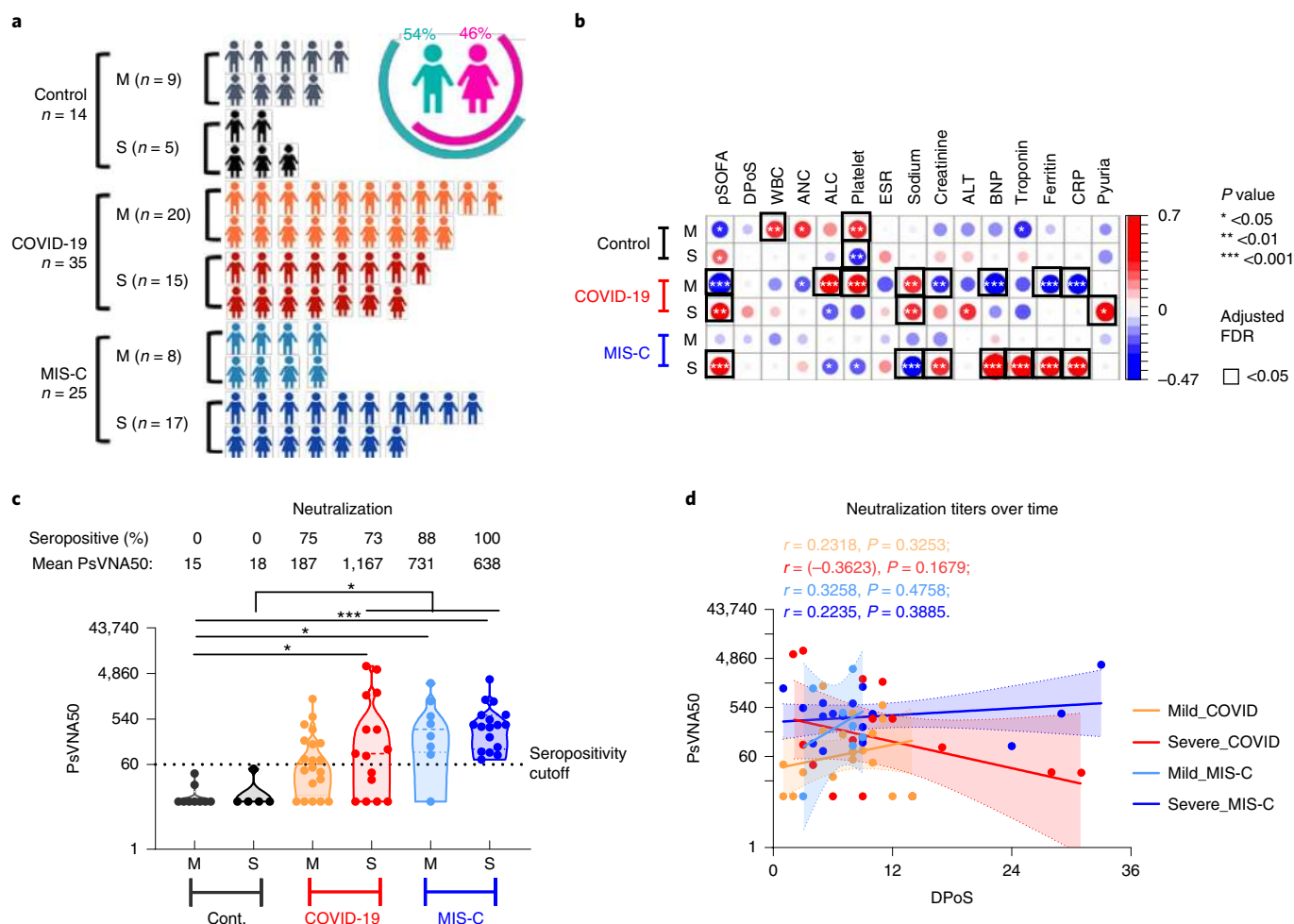
Fig. 1 | Demographics and neutralization titers of children with COVID-19 or MIS-C. a, Overview of pediatric cohort, including hospitalized mild/moderate or severe control (SARS-CoV-2 negative) and children with mild/moderate (M) or severe (S) COVID-19 or MIS-C. **b**, Spearman correlations of clinical parameters among the groups assigned according to clinical disease severity was performed using R on log-transformed data. Color intensity and size of the circles are proportional to the correlation coefficients. Heat scale in the correlogram ranges from −0.7 (blue; negative correlation) to +0.7 (red; positive correlation). Statistical analyses among groups were performed using R and differences were considered statistically significant with a 95% confidence interval when the P value was less than 0.05 (*$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$). Original P values that show significance are marked on the circles. A black box around the circle indicates the adjusted P value of false discovery rate (FDR) < 0.05. **c**, Violin plots of SARS-CoV-2 serum PsVNA50 across clinical groups, as determined by spike PsVNA. Each dot represents a distinct patient color coded by groups: gray, mild/moderate control ($n = 9$); black, severe control ($n = 5$); orange, mild/moderate COVID-19 ($n = 20$); red, severe COVID-19 ($n = 15$); cyan, mild/moderate MIS-C ($n = 8$); blue, severe MIS-C ($n = 17$). Statistical analyses among groups were performed using R. The tests were two-sided using Tukey multiple-group comparisons to reduce type 1 error. The differences were considered statistically significant with a 95% confidence interval when the P value was less than 0.05 (*$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$). Detailed statistical data are provided in Supplementary Table 9. **d**, Correlation and linear regression of serum neutralization titers with DPoS in children with MIS-C and COVID-19. The associated Pearson's correlation coefficients (r) and linear regression significance (P) are colored accordingly. The trend line fits are shown as center solid lines with the error bands representing the 95% confidence intervals shown as shaded colored areas. ALC, absolute lymphocyte count; ALT, alanine aminotransferase test; ANC, absolute neutrophil count; BNP, brain natriuretic peptide; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; WBC, white blood cell.

antibody test and were found to have alternative diagnoses. Thus, these children with phenotypically similar disease states served as disease-matched controls to identify differences in immune responses unique to SARS-CoV-2 infection. This is a more relevant clinical comparator group than healthy children, as diagnostic and immunophenotyping analyses are needed to differentiate children with COVID-19 and MIS-C from hospitalized children with similar clinical phenotypes, not from healthy children in the community. None of the patients with MIS-C reported having a preceding COVID-19-like illness before hospitalization; thus, DPoS refers to the days after onset of the symptoms that prompted hospitalization. Differential immune profiles at the time of hospitalization could aid

in establishing diagnosis and understanding immunopathogenesis of these distinct clinical phenotypes. Six of the 14 control children were suffering from Kawasaki disease, which serve as comparator control for children with MIS-C, as described in the study by Consiglio et al.[7].

COVID-19 symptoms typically occur on average 1 week (range 2–12 d) after SARS-CoV-2 infection, whereas MIS-C is a hyper-inflammatory syndrome that occurs approximately 4 weeks after infection. Because it is difficult to accurately determine the day of SARS-CoV-2 infection in these children, we defined these patients from the DPoS. Therefore, to control for differences in age, gender and DPoS among patients with COVID-19 and MIS-C, all

statistical calculations across disease groups, were performed on log2-transformed data in R package using age, gender and DPoS as covariates.

Pseudovirus neutralization assay (PsVNA) showed higher 50% pseudovirion neutralizing antibody (PsVNA50) titers in the severe COVID-19 or MIS-C groups compared with the mild disease group, but this difference did not reach statistical significance in comparison with mild/moderate disease (Fig. 1c) from these patients. Hospitalized controls did not show measurable PsVNA titers (PsVNA50 < 60). No gender differences were observed for neutralization titers (Extended Data Fig. 1). Over time, a trend in decline of neutralization titers was observed in the severe COVID-19 group (Fig. 1d).

Cytokine and chemokine profiling identified hyperinflammatory response and revealed elevation of multiple cytokines (interleukin (IL)-2, IL-6, IL-10, IL-13 and IL-17) in children with severe MIS-C or COVID-19, in comparison to children with mild/moderate disease (Fig. 2, Extended Data Fig. 2 and Supplementary Table 4). Tumor necrosis factor (TNF) was significantly higher in children with mild MIS-C, in comparison to children with mild/moderate or severe COVID-19 (Fig. 2k). Temporal analyses of cytokines/chemokines in serum, either by linear regression or grouped intervals, support a distinct disease course in SARS-CoV-2-infected children. Although IL-10 levels were sustained in both patients with severe MIS-C and those with COVID-19; IL-6, IL-8 and macrophage inflammatory protein (MIP)-1β levels declined in severe COVID-19/MIS-C over time (Fig. 2 and Extended Data Fig. 2). Because all eight children with mild MIS-C had samples collected within 9 d of symptom onset, the temporal analysis was performed to compare severe COVID-19 versus mild COVID-19 versus severe MIS-C only. Statistical analysis for the regression lines to test 'elevation' or the differences in the intercepts between the trend lines revealed statistically significant differences for IL-2, IL-6, IL-10, IL-17 and monocyte chemoattractant protein-1 (MCP-1) in mild COVID-19 versus severe COVID-19, or severe COVID-19 versus severe MIS-C (Supplementary Table 5).

To gain insights into key differences between COVID-19 and MIS-C with mild/moderate versus severe disease, we correlated the measurements of these soluble cytokine/chemokines using multivariable analysis (Fig. 2l). Positive associations of the cytokines linked to cytokine-release syndrome, such as IL-6 and IL-10, as well as IL-2, IL-7 and IL-13, were observed in children with severe disease.

SARS-CoV-2 antibody epitope repertoire in children. To elucidate the IgM, IgG and IgA antibody epitope repertoire across the entire SARS-CoV-2 proteome in children, epitope mapping was performed using the SARS-CoV-2 genome-fragment phage display library (GFPDL) containing sequences ranging from 50–1,500-bp long from the complete SARS-CoV-2 genome with

>10[7,9] unique phage clones (Extended Data Fig. 3). Previously, GFPDL analyses of highly pathogenic avian influenza virus, respiratory syncytial virus (RSV), Zika virus and Ebola virus were used to map the antibody repertoires of convalescent sera and in individuals after vaccinations. These studies revealed several diagnostic and protective targets[9–12].

The SARS-CoV-2 GFPDL displayed linear and conformational epitopes with random distribution of size and sequence of inserts that spanned the entire SARS-CoV-2 genome (Extended Data Fig. 3). In preliminary studies to characterize the SARS-CoV-2 spike GFPDL, epitope mapping of monoclonal antibodies targeting the SARS-CoV-2 spike protein or receptor-binding domain (RBD) identified the expected linear or conformation-dependent epitopes recognized by these monoclonal antibodies. Recently, we showed that SARS-CoV-2 spike GFPDL can recognize both linear, conformational and neutralizing epitopes in the post-vaccination sera of rabbits[13] and after SARS-CoV-2 infection in adults and the elderly[14,15]. Moreover, the SARS-CoV-2 full-genome GFPDL adsorbed >92% of SARS-CoV-2 virion-specific serum antibodies in patients with COVID-19 and in those with MIS-C (Extended Data Fig. 3), providing proof of concept for use of the SARS-CoV-2 GFPDL for epitope repertoire analyses of pediatric samples.

First-drawn samples collected from each child at the same time point to match DPoS from intravenous immunoglobulin (IVIG)-untreated, either mild/moderate or severe COVID-19 or MIS-C patients were pooled and evaluated to delineate the epitopes of IgM, IgG and IgA separately across the SARS-CoV-2 proteome compared with hospitalized controls (Fig. 3). Altogether, we analyzed targets of IgM, IgG, and IgA antibodies in five serum pools in duplicate, in total assessing >10[8] potential antibody epitope specificities. SARS-CoV-2 infection induced IgM/IgG/IgA antibodies with diverse overlapping epitope repertoire distribution, displaying small and large sequences spanning the entire SARS-CoV-2 proteome (Fig. 3 and Extended Data Figs. 3–5). Overall, IgM, IgG and IgA recognized similar protein antigenic sites across the SARS-CoV-2 proteome, suggesting that SARS-CoV-2 infection induced a shared epitope profile in these children, but IgM antibodies from severe COVID-19 and severe MIS-C serum recognized roughly a ten-fold higher number of phage clones compared with IgG or IgA phage titers in the same cohort (Extended Data Fig. 3). However, when fine epitope specificities of SARS-CoV-2-induced IgM, IgG and IgA were compared, we observed considerable divergence among some epitopes that were recognized especially by only IgG or only IgA in these children with COVID-19 and MIS-C (Extended Data Figs. 4 and 5).

The severe hospitalized control (SARS-CoV-2-negative) serum IgM, IgG or IgA bound ~1,000-fold fewer phages that primarily recognized epitopes in NSP3, NSP12–16, spike and N protein (Extended Data Fig. 3). The cross-reactivity observed in control serum potentially could be due to antibodies generated by endemic

---

Fig. 2 | Cytokine profile of children with COVID-19 or MIS-C. a, Heat map of 17 cytokines and chemokines in serum samples of children with mild/moderate or severe COVID-19 versus MIS-C and hospitalized SARS-CoV-2-negative controls. The color scale represents the concentration of each protein. **b–k**, Cytokine/chemokine levels reported over DPoS with linear regression analyses. IL-2, IL-6, IL-7, IL-8, IL-10, IL-13, IL-17, MCP-1, IFN-γ and TNF concentrations in serum are shown in log₁₀ of pg ml⁻¹. Statistical differences among groups were performed using R and considered statistically significant with a 95% confidence interval with a *P* value less than 0.05 (*P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001). The detailed statistical data are provided in Supplementary Table 9. Correlation of cytokine and chemokine measurements were reported over DPoS and continuously as linear regressions. The trend line fits were performed for mild/moderate COVID-19 (orange) versus severe COVID-19 (red) versus severe MIS-C (blue), using the linear regression model. Associated Pearson's correlation coefficients and linear regression significance are colored accordingly. The trend line fits are shown as central, colored solid lines with the error bands representing 95% confidence intervals shown as shaded colored areas. **l**, Spearman correlations and significance of cytokines/chemokines with disease-severity groups. Multivariate linear regression analysis on log₁₀-transformed data was performed using R. DOH, days of hospitalization; DOI, days of ICU admission. The Spearman's rank correlation coefficient (ρ) is indicated by circle size and heat scale in the correlogram. Heat scale is indicated as gradient color from ρ = −0.7 (blue; strong negative correlation) to ρ = 0 (white, no correlation) to ρ = 0.7 (red; strong positive correlation). The *P* values that show significance are marked on the circles; *P < 0.05; **P < 0.01; ***P < 0.001. A black box around the circle indicates the adjusted *P* value of FDR < 0.05.

human coronaviruses (hCoVs) that cross-react due to conservation in homologous protein sequences (for example, NSP3, NSP12, NSP13, NSP14 and spike; Supplementary Table 6). This may also

result in an anamnestic response boosting B cell memory against hCoVs following SARS-CoV-2 infection in children. Within NSPs, the COVID-19 IgM recognized several antigenic sites in NSP3,
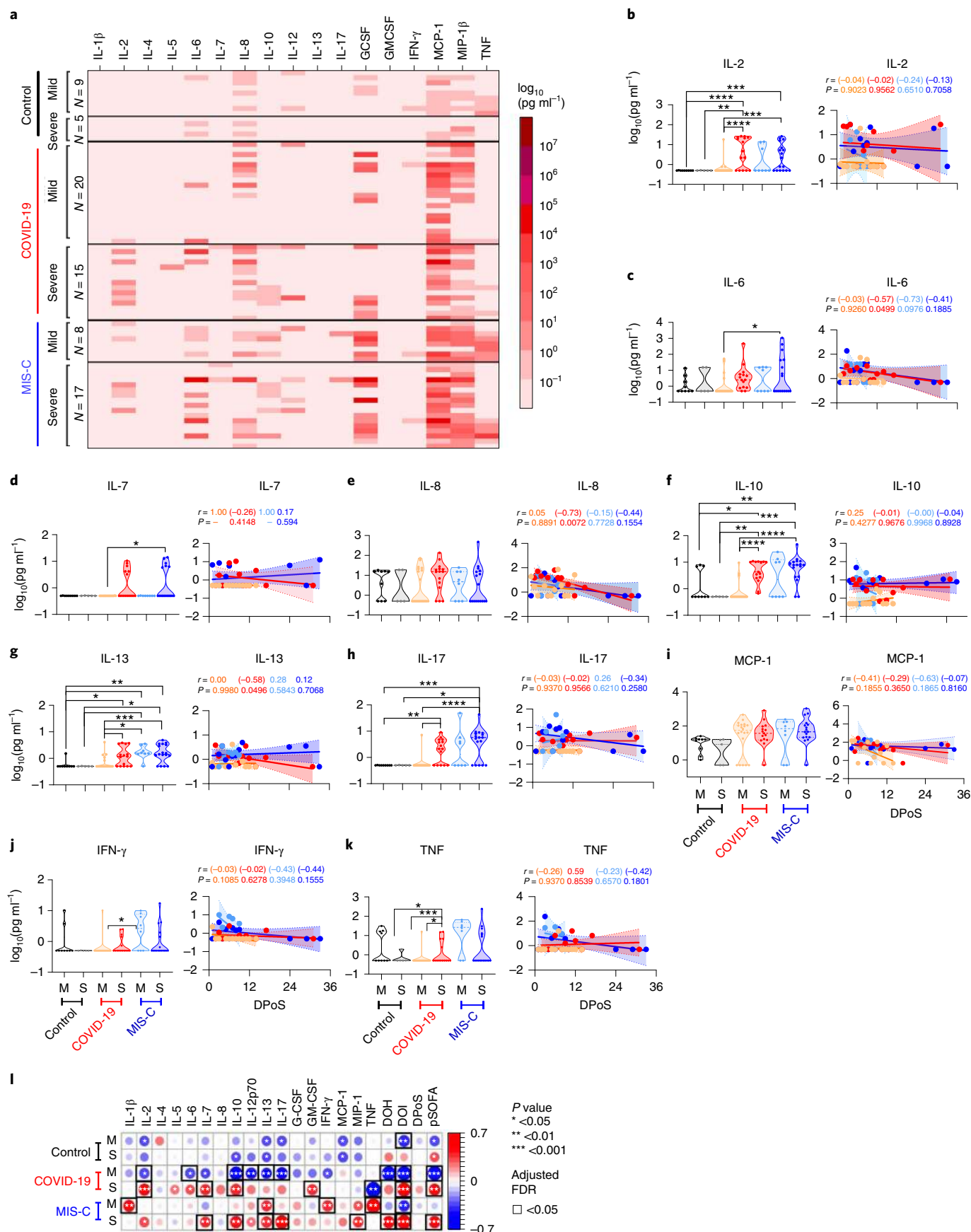
NSP12, NSP13 and NSP14, followed by few sites within NSP1, NSP2, NSP5, NSP15 and NSP16 (Fig. 3 and Extended Data Fig. 5). In many cases, we observed recognition of overlapping epitope sequences, indicating an epitope in the common region. IgM in children with mild/moderate COVID-19 recognized NSP2 (249–331), few sites in NSP3 (1,329–1,483, 1,678–1,815 and 2,475–2,660) and NSP12 (4,724–4,791) at higher frequency compared with IgM from severe COVID-19 (Extended Data Fig. 5). In ORFs, IgM from children with COVID-19 preferentially recognized N, followed by sites in ORF5 (M), ORF8 and ORF9b (9,786–9,828). In children with severe COVID-19, a strong response was observed to S1 immuno-dominant antigenic epitopes (7,729–7,761 and 7,738–7,799) at the C terminus of the RBD as well as a sequence containing a fusion peptide (7,972–8,013), compared to that in patients with mild/moderate COVID-19. The IgM repertoire in MIS-C was diverse across the entire SARS-CoV-2 proteome; however, higher antibody binding to epitopes in NSP3 (1,071–1,158 and 2,437–2,526), NSP5/6 and NSP7/8, ORF7a (9,110–9,173) and N (9,586–9,600) was observed compared with that in patients with COVID-19 (Extended Data Figs. 4 and 5). IgM from mild/moderate MIS-C recognized sites in NSP2 (249–331), NSP3 (1,265–1,465), NSP7/8 (3,905–3,994), ORF7a (9,110–9,173), ORF8 (9,301–9,314) and N (9,413–9,485) at a higher frequency than that from severe MIS-C, whereas severe MIS-C IgM showed a higher frequency of epitopes in NSP3 (1,071–1,158 and 2,437–2,526) compared with mild/moderate MIS-C IgM (Extended Data Fig. 5).

The IgG epitope repertoire mimicked the diverse IgM epitope profile, with primary recognition of sites in NSP3, NSP12, NSP14, spike, ORF3a, ORF4, ORF5, ORF8 and N (Fig. 3 and Extended Data Figs. 4 and 5). IgG from children with severe COVID-19 showed a higher frequency of sites in NSP3 (1,265–1,465 and 1,447–1,552), NSP12 (5,164–5,218), NSP14 (5,942–6,040), ORF3a (8,407–8,438), ORF8 (9,301–9,314), and few sites in N (9,655–9,663) and spike (7,323–7,430 and 8,066–8,257) compared with IgG from children with mild/moderate COVID-19 (Extended Data Fig. 5). The IgG repertoire in MIS-C was also diverse, with additional recognition of ORF4 (8,760–8,782) and ORF5 (8,844–8,912) by mild/moderate MIS-C in comparison with severe MIS-C, and stronger IgG binding to NSP3 (1,732–1,775) and ORF3a (8,407–8,438) in severe MIS-C compared with mild MIS-C (Fig. 3 and Extended Data Figs. 4 and 5). Compared with severe MIS-C, IgG from severe COVID-19 selected a significantly higher frequency of clones displaying NSP3 (1,447–1,552 and 1,573–1,658), spike (8,066–8,257) and N (9,647–9,663), whereas severe MIS-C recognized NSP3 (1,732–1,775) and ORF3a (8,407–8,438) significantly higher than severe COVID-19 (Extended Data Fig. 5).

The IgA repertoire was equally diverse, wherein all sera recognized immunodominant epitopes in NSP3 (1,678–1,815), NSP12 (5,164–5,218) and IgM/IgG/IgA epitopes in spike (7,738–7,799) and N (9,655–9,663; Fig. 3 and Extended Data Figs. 4 and 5). Interestingly, the severe MIS-C IgA uniquely recognized several epitopes located in NSP5/6 (3,506–3,598), NSP14 (6,296–6,503) and ORF4 (E; 8,760–8,782) that were absent or minimally observed by IgA/IgM/IgG from patients with mild/moderate or severe COVID-19 or patients with mild/moderate MIS-C (Extended Data Fig. 5).

To test the probability that host autoantibodies may contribute to the identified SARS-CoV-2 antibody epitope repertoire, we performed homology searches for the GFDPL-identified antigenic sites with a human protein database. This analysis identified some sequences in nucleocapsid, ORF9b and ORF10 proteins that demonstrated >70% sequence conservation hits with human proteins (Supplementary Table 7).

Antibody binding and avidity profile in MIS-C versus COVID-19. To further investigate the individual antibody profiles to different proteins encoded by SARS-CoV-2, quantitative and qualitative surface plasmon resonance (SPR) analyses were performed with 24 recombinant SARS-CoV-2 proteins, including 4 structural proteins (spike, S; nucleocapsid, N; envelope, E; and membrane, M), 8 accessory proteins and 13 NSPs using serial dilutions of polyclonal serum. These proteins were produced in mammalian cells, insect cells or bacterial cells, and were either obtained from commercial sources or produced in-house (Supplementary Table 8). Owing to lack of protein availability, NSP4, NSP6, NSP10, NSP15 and ORF6 were not included in this SPR analysis. Moreover, weak or no binding was observed to NSP4, NSP6, NSP10 or ORF6 in GFPDL selection (Extended Data Fig. 5), in agreement with earlier studies in convalescent adults[16]. In addition, antibody-binding kinetics were analyzed using three spike subdomains (S1, RBD and S2), prefusion spike proteins from six different hCoVs (SARS-1, Middle East respiratory syndrome (MERS), 229E, HKU, OC43 and NL63) as well as influenza H1N1 hemagglutinin (HA) and RSV prefusion F protein, as controls. Most humans are exposed to RSV by 2 years of age. RSV fusion protein (F) is widely used for antibody evaluation following RSV infection and for development of RSV vaccines[17].

Control pediatric serum ($n = 14$) did not display antibody binding against most SARS-CoV-2 proteins. However, weak binding was observed against a few NSPs (NSP3, NSP7/8, NSP12, NSP13 and NSP16) in serum from seven older controls (ages 8–14 years), but not in serum from the seven younger hospitalized controls (ages 0.3–4 years; Fig. 4a and Extended Data Fig. 6). Serum antibody-binding titers to most SARS-CoV-2 proteins were higher in children with severe MIS-C or COVID-19 than in children with mild/moderate disease; however, interestingly, NSP2 and NSP13 showed significantly higher reactivity in children with mild/moderate disease (Fig. 4a). Severe MIS-C serum showed higher antibody reactivity to ORF3a and ORF9c than COVID-19 serum, whereas severe COVID-19 serum showed higher antibody binding to ORF5 (M) than did MIS-C serum (Fig. 4a and Extended Data Fig. 6). Analysis over time revealed that differential antibody levels to ORF3a, ORF3b and ORF9c were sustained in patients with severe MIS-C, whereas children with severe COVID-19 showed sustained

---

Fig. 3 | SARS-CoV-2 IgM, IgG and IgA epitope repertoires in children. IgM, IgG and IgA antibody epitope repertoire recognized in the SARS-CoV-2-infected sera from children with mild/moderate (M) or severe (S) COVID-19 or MIS-C across the whole proteome of SARS-CoV-2 encoding four structural proteins (S, E, M and N) as well as 16 NSPs and 9 accessory proteins. The schematic on top shows various proteins in the full proteome of SARS-CoV-2. The ORF1ab polypeptide encoding various NSPs (numbered sequentially based on occurrence) is shown. Other proteins are labeled as spike (S), ORF3a (3a), ORF4 (E), ORF5 (M), ORF7a (7a), ORF8 (8), ORF9a (N), ORF9b (9b), ORF9c (9c) and ORF10 (10). Epitope coverage map shows alignment position (amino acid position) on the x axis and coverage (phage clone frequency) of each position on the y axis defined by epitopes recognized by IgG, IgM and IgA in mild/moderate and severe COVID-19 (orange and red, respectively) and in mild/moderate and severe MIS-C (cyan and blue, respectively). The y axis shows phage clone frequency in the range of 0–200 for IgM, IgG or IgA selections with these samples, 0–1,000 for IgM clones with severe COVID-19 and 0–500 for severe MIS-C. Owing to higher frequency of selected phage clones, the original clone frequencies (y axis) of spike and ORF3a antigenic sites were scaled (a scaling ratio of 5.25 was applied for spike sites 7,738–7,799 and 3.05 for ORF3a sites 8,407–8,438) to aid in data visualization. Samples were fresh and not frozen and thawed. All samples were run in one batch at the same time for each assay. The GFPDL affinity selection was performed in duplicate (two independent experiments were performed by a researcher in the laboratory, who was blinded to sample identity). Similar numbers of bound phage clones and epitope repertoire were observed in the GFPDL panning.

titers against ORF5 (M; Fig. 4b). Because all eight children with mild MIS-C had samples collected at early time point of hospitalization (within 9 d of symptom onset), the temporal analysis in Fig. 4b was performed to compare mild COVID-19 versus severe-COVID-19 versus severe MIS-C only.

Serum antibody binding to the SARS-CoV-2 prefusion form of spike and S-RBD was significantly higher in patients with severe disease, compared with that in controls (Fig. 4a). However,

binding antibodies against the S2 domain were observed in six of the seven older uninfected controls (ages 8–14 years) but were absent in the seven younger children (ages 0.3–4 years; Fig. 4a). SARS-CoV-2-induced S2-binding antibodies were not significantly different between children with MIS-C and those with COVID-19. Regression analysis did not identify differences among groups for spike-, RBD- or S2-specific antibody responses (Fig. 4b).
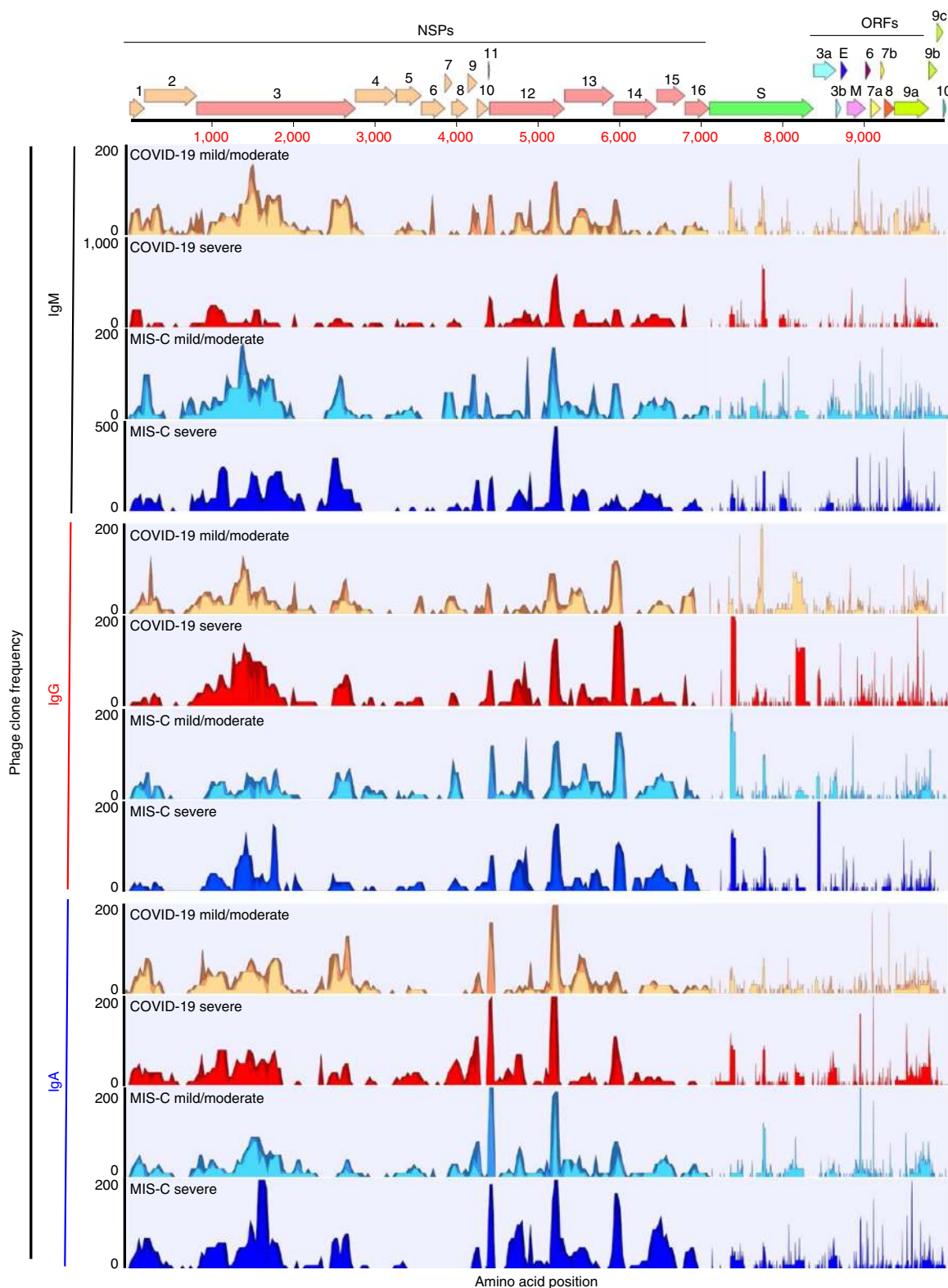
To determine the antibody avidity maturation against different SARS-CoV-2 proteins following infection, the dissociation kinetics (off-rate constants) of antigen–antibody complexes that are independent of antibody concentration were used as a surrogate for overall average avidity of polyclonal antibody against SARS-CoV-2 proteins using SPR[10,13,18]. Previously, we showed that antibody kinetics measured under optimized SPR conditions primarily represent the monovalent interactions between the antibody–antigen complex, as antigen–antibody-binding off-rates of the IgG and Fab interaction with protein antigens were similar[12–15]. Because sera from pediatric controls have minimal or no binding to the SARS-CoV-2 proteins, antibody avidity was determined for children with COVID-19 or MIS-C. SARS-CoV-2 infection of children induced weak antibody avidity against most NSPs (apart from NSP2 and NSP12) with off-rates ranging from 0.8 to ~0.1 per second (Extended Data Fig. 7). In addition to antibodies against spike and N proteins, SARS-CoV-2-induced antibodies showed higher avidity (off-rates of ~0.1–0.01 per second) to several accessory proteins including ORF3a, ORF3b, ORF4 (E), ORF5 (M), ORF7a, ORF7b, ORF8 and ORF9b (Fig. 4c and Extended Data Fig. 7). Children with milder disease showed significantly higher antibody avidity (slower off-rates) against NSP2 and NSP13, whereas children with severe disease demonstrated higher antibody avidity to NSP11 and NSP12 (Fig. 4c and Extended Data Fig. 7). Severe MIS-C serum contained higher avidity antibodies against NSP7/8 complex and ORF9c compared with either mild or severe COVID-19 serum (Fig. 4c and Extended Data Fig. 7). Although children with mild MIS-C demonstrated significantly higher antibody avidity against NSP9 and ORF3a compared to patients with either mild/severe COVID-19 or severe MIS-C (Extended Data Fig. 7).

Significantly higher avidity against SARS-CoV-2 prefusion spike protein was observed in children with mild/moderate COVID-19 and MIS-C than in children with severe disease (Fig. 4c). But this difference in antibody avidity did not translate to spike subdomains S1, RBD or S2, suggesting that antibodies recognize epitopes present on the native trimeric conformation of prefusion spike, but not those retained in individual spike subdomains.

SPR analysis with prefusion spike protein from different hCoVs including pathogenic SARS-CoV-1 and MERS-CoV, as well as seasonal hCoVs (229E, HKU1, OC43 and NL63), revealed SARS-CoV-2 infection in children induced antibodies that cross-reacted with the prefusion spike protein of SARS-CoV-1, MERS-CoV and seasonal hCoVs (Fig. 4a and Extended Data Fig. 6). Patients with severe disease demonstrated higher antibody binding to the prefusion spike protein of other hCoVs compared with that of children with mild/moderate disease. High levels of binding antibodies against RSV-F protein were observed in children with severe MIS-C (Extended Data Fig. 6); however, there was no significant correlation between SARS-CoV-2 spike protein-binding antibodies with RSV-F or influenza HA-binding antibodies. RSV-F protein lacks sequence conservation with SARS-CoV-2 spike protein, and most children get infected with RSV by the age of 2 years[17]. The 14 age-matched hospitalized SARS-CoV-2-negative controls showed no or minimal antibody binding to SARS-CoV-2 proteins in SPR (Fig. 4a) and no SARS-CoV-2 neutralization (Fig. 1a), demonstrating no nonspecific cross-reactivity of anti-RSV antibodies with SARS-CoV-2. This observation was further confirmed by SPR analysis of an additional set of sera from hospitalized seronegative controls and SARS-CoV-2-seropositive COVID-19 and MIS-C patients obtained from later time points from some of the children in our study (Extended Data Fig. 6). Therefore, the increase in anti-RSV-F-binding antibodies following SARS-CoV-2 infection in children with severe COVID-19 and MIS-C seems to be a true biological observation, which agrees with inflamed antibody profiles specific for multiple pathogens, including RSV, generated following SARS-CoV-2 infection in children[19].
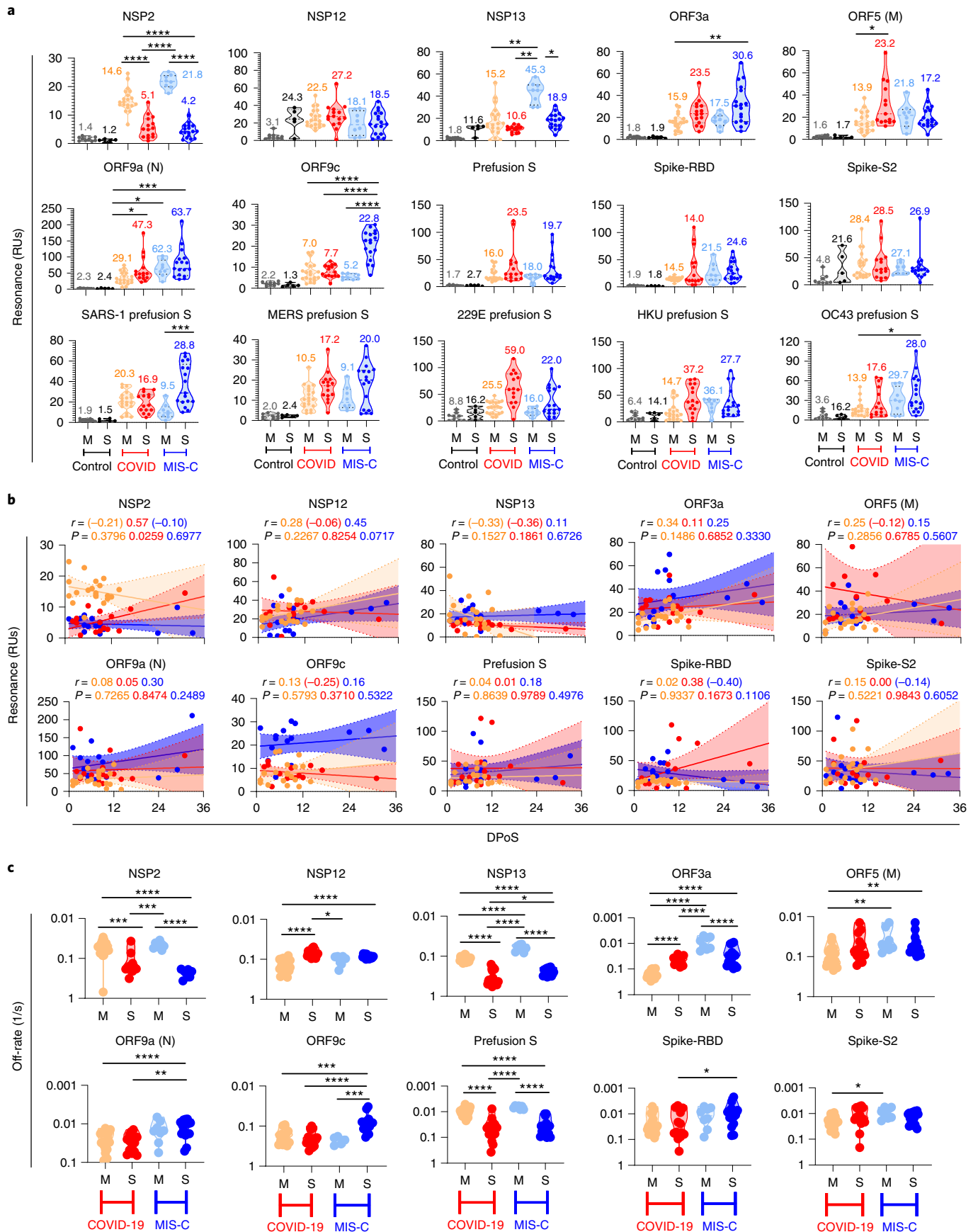
Immune correlates of disease severity in COVID-19 and MIS-C. To identify the immune markers that associate with clinical disease severity in pediatric MIS-C and COVID-19, linear regression analysis was performed for maximum pSOFA score for each patient and corresponding serum cytokine/chemokine, total antibody-binding titers and antibody avidity against each of the SARS-CoV-2 proteins (Fig. 5 and Extended Data Fig. 8–10). Higher neutralization titers did not associate with the decline of pSOFA scores in these children (Fig. 5a). A significant positive correlation was observed between pSOFA scores and IL-2, IL-7 and IL-10, which were upregulated in most patients with severe disease (Fig. 5b and Extended Data Fig. 8).

Antibody-binding titers (maximum RUs) to ORF3a, ORF5 (M) and ORF9c demonstrated a significant positive correlation with pSOFA scores, whereas antibody binding to NSP2 correlated with decrease in pSOFA scores (Fig. 5c and Extended Data Fig. 9). A significant positive association of antibody avidity against ORF3a, NSP11, NSP12 and NSP16 with disease severity was observed (Fig. 5d and Extended Data Fig. 10). However, the increase in antibody avidity (decreasing off-rates) against NSP2, NSP13 and SARS-CoV-2 prefusion spike protein (but not S1, RBD or S2) significantly correlated with the decrease in clinical pSOFA score in these MIS-C and COVID-19 pediatric patients (Fig. 5d and Extended Data Fig. 10). These data suggest that both the levels of cytokines as well as antibody specificity and avidity may be associated with SARS-CoV-2-mediated disease severity or its resolution in children.

To characterize the SARS-CoV-2 immune profile difference more finely between patients with MIS-C and pediatric patients with COVID-19, we conducted principal-component analysis (PCA) for

Fig. 4 | Serum antibody kinetics of COVID-19 or MIS-C with SARS-CoV-2 proteins. Serum from children with mild/moderate (M) or severe (S) COVID-19 or MIS-C or hospitalized controls were analyzed for antibody binding to purified proteins by SPR. **a,b**, Total antibody binding of tenfold-diluted serum is represented in SPR resonance units (RUs; **a**) and as linear regressions with DPoS (**b**). Each dot represents a distinct patient and DPoS, color coded by groups: gray, mild/moderate control (*n* = 9); black, severe control (*n* = 5); orange, mild/moderate COVID-19 (*n* = 20); red, severe COVID-19 (*n* = 15); cyan, mild/moderate MIS-C (*n* = 8); blue, severe MIS-C (*n* = 17). **a**, The mean values for each group are shown. **b**, Correlations of binding antibodies reported over DPoS and continuously as linear regressions were performed for mild/moderate COVID-19 (orange) versus severe COVID-19 (red) versus severe MIS-C (blue), using linear regression analysis in GraphPad Prism. Associated Pearson's correlation coefficients and linear regression significance values are colored accordingly. The trend line fits are shown as the center solid lines, with error bands representing 95% confidence intervals (shaded colored areas). **c**, Antibody–antigen off-rate constants as a surrogate of antibody avidity of post-SARS-CoV-2 infection serum to SARS-CoV-2 proteins were measured using SPR and plotted in the respective color coding for each patient group. Statistical differences among groups were calculated using R. The differences were considered statistically significant with a 95% confidence interval when the *P* value was less than 0.05 (*P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001). Antibody binding of all SARS-CoV-2-infected children (COVID-19 or MIS-C) were significantly higher than control samples for different SARS-CoV-2 proteins. Therefore, for clarity of representation, only statistically significant differences between SARS-CoV-2-infected patient groups are shown (COVID-19 versus MIS-C with mild versus severe disease). Detailed statistical data are provided in Supplementary Table 9. All SPR experiments were performed twice, and the researchers performing the assay were blinded to sample identity. In these optimized SPR conditions, the variation for each sample in duplicate SPR runs was <5%.
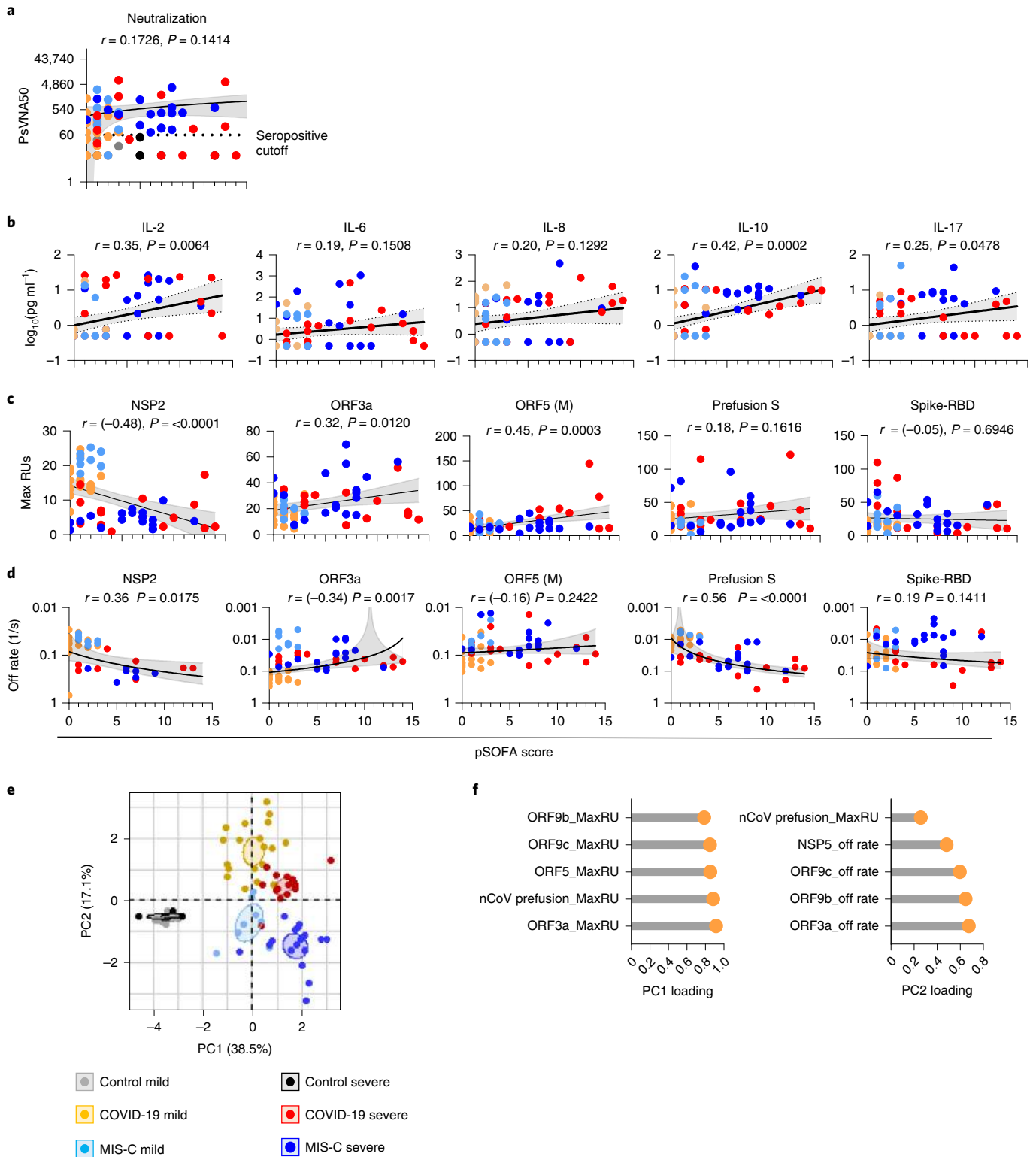
Fig. 5 | Immune markers of disease severity in children with COVID-19 or MIS-C. a–d, Correlation and linear regression of serum neutralization titers (**a**), cytokines (**b**), total antibody binding (**c**) and antibody avidity (**d**) for the first-drawn time-point sample for each patient, with corresponding maximum pSOFA clinical disease-severity scores across all pediatric MIS-C and COVID-19 groups, regardless of disease-severity classification across all SARS-CoV-2-infected children. The associated Pearson's correlation coefficients and linear regression significance values are shown. Each dot represents a unique patient antibody binding and associated pSOFA score: orange, mild/moderate COVID-19 ($n = 20$); red, severe COVID-19 ($n = 15$); cyan, mild/moderate MIS-C ($n = 8$); blue, severe MIS-C ($n = 17$). The black center lines indicate the trend lines for all patients with the error bands representing 95% confidence intervals, shown as shaded areas. **e,** PCA for antibody binding and avidity to selected SARS-CoV-2 proteins (prefusion spike, NSP5, ORF3a, ORF5 (M) and ORF9c) of the first-drawn serum samples collected from 74 children. The percentage of variables explained by each principal component is indicated in the parenthesis. Points are colored by sample groups: gray, mild/moderate control; black, severe control; orange, mild/moderate COVID-19; red, severe COVID-19; cyan, mild/moderate MIS-C; blue, severe MIS-C. Ellipses show the 95% confidence interval around the colored centroid for each group. **f,** The top five antibody variables describing component loadings (PC1 and PC2) of PCA analysis are shown in the loading plots.
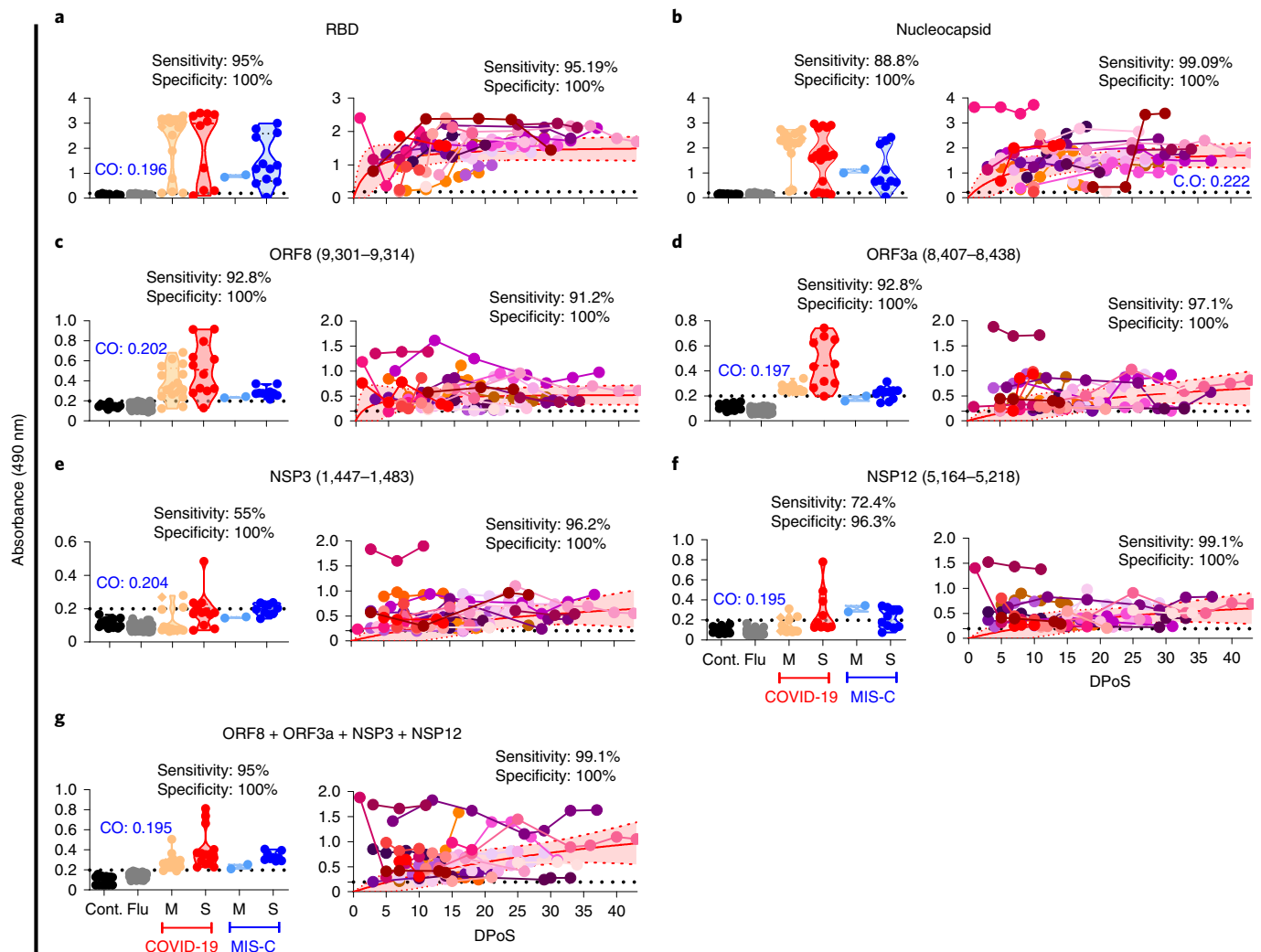
Fig. 6 | Peptide-based SARS-CoV-2 serodiagnostic ELISA for SARS-CoV-2. IgG detection in post-SARS-CoV-2-infection sera from children (black, control ($n = 20$; 10 hospitalized controls and 10 additional nonhospitalized healthy pediatric samples (ages 12–35 months, collected in 2009)); orange, mild/moderate COVID-19 ($n = 16$); red, severe COVID-19 ($n = 10$); cyan, mild/moderate MIS-C ($n = 2$); blue, severe MIS-C ($n = 10$)) in comparison with 72 pre-COVID influenza-infected adult controls (flu; collected in 2010; gray) are shown on the left, against each protein (**a** and **b**), or individual peptides (**c–f**) or a mixture of four peptides (**g**), by ELISA using HRP-conjugated anti-human IgG-Fc-specific antibody. Longitudinal analysis of IgG responses to the SARS-CoV-2 proteins and peptides in serum ($n = 100$) collected at different DPoS from 27 COVID-19 adult patients (shades of red) are shown on the right of each protein/peptide. Dots represent individual serum samples collected at the indicated DPoS, and the samples from the same COVID-19 patients are connected by the lines. Absorbance (y axis) is for 100-fold dilution serum. Cutoff values for each antigen in ELISA were determined by area under the receiver operating characteristic (ROC) curve analysis of the negative controls (using ten control hospitalized children serum samples at 100-fold serum dilution) in GraphPad Prism. The trend line fits were performed for longitudinal adult COVID-19 patients using a nonlinear regression model with polynomial distribution through the origin in GraphPad Prism. The trend line is depicted as the central, colored line with the error bands representing 95% confidence intervals (shaded colored area). Area under the ROC curve analysis using GraphPad Prism was performed to calculate sensitivity (%) and specificity (%) values with 95% confidence intervals. Data are the average values of two experimental ELISA runs. The variation for each sample in duplicate runs was <7%.

cytokines/chemokines, antibody titers and avidity, or both antibody parameters and cytokines/chemokines. PCA identified unique antibody parameters that separated hospitalized controls from patients with SARS-CoV-2 and appear to distinguish between children with mild/moderate versus severe COVID-19 or MIS-C (Fig. 5e,f). This antibody signature includes antibody-binding titers and antibody avidity to SARS-CoV-2 prefusion spike protein, ORF9b and ORF3a, which could potentially serve as markers for disease-severity classification of children with SARS-CoV-2 infection (Fig. 5e,f).

Serodiagnosis and neutralization targets of SARS-CoV-2. To evaluate the relevance of immunodominant epitopes identified

in SARS-CoV-2-infected children, we sought to determine if synthetic peptides, reflective of these antigenic sites, would be targets for SARS-CoV-2 serodiagnosis or, upon immunization, generate SARS-CoV-2-neutralizing responses. Currently, in addition to spike-based vaccines, several whole-virus-inactivated vaccines (for example, Sinovac and Covaxin) containing spike and nucleocapsid proteins are being used for mass vaccination in several countries including China, India and others[20–23]. These inactivated vaccines will induce antibodies that will react in SARS-CoV-2 S-based or N-based serodiagnostic tests leading to the phenomenon of vaccine-induced seropositivity (VISP), which would confound the interpretation of current S-based and N-based SARS-CoV-2
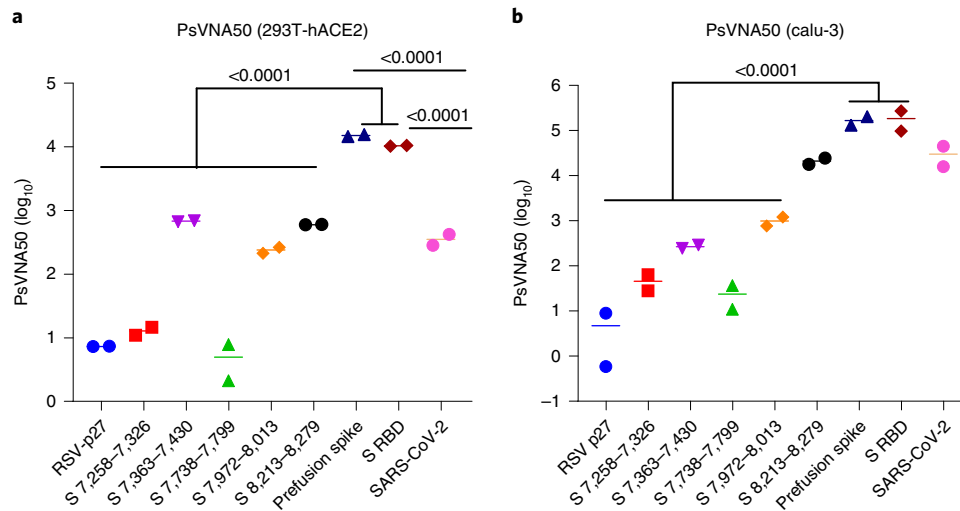
Fig. 7 | SARS-CoV-2 neutralization following GFPDL peptide immunization. a,b, SARS-CoV-2 neutralization titers of serum from rabbits immunized with GFPDL-identified SARS-CoV-2 spike peptides or control RSV p27 peptide by PsVN assay using either 293T-hACE2 (**a**) or Calu-3 (**b**) cells. Sera from rabbits immunized with SARS-CoV-2 prefusion spike protein, RBD protein or heat-killed SARS-CoV-2 virions were used as comparators. The data (50% neutralization titers; PsVNA50) represent the median value of two rabbits immunized with each immunogen from assays run in duplicate. Pre-vaccination rabbit sera did not neutralize SARS-CoV-2 tested in PsVN assay. Samples were fresh and not thawed. All samples were run in one batch at the same time for each assay. Differences among groups were analyzed using multiple-group comparisons by non-parametric (Kruskal–Wallis) statistical testing using Dunn's post hoc analysis in GraphPad Prism. The differences were considered statistically significant with a 95% confidence interval when the P value was less than 0.05 (*$P \le 0.05$, **$P \le 0.01$, ***$P \le 0.001$, ****$P \le 0.0001$).

serodiagnostic assays. Therefore, to address the phenomenon of VISP, we focused on differential antigenic sites in non-spike and non-nucleocapsid proteins using SARS-CoV-2 GFPDL analysis to determine potential serodiagnosis candidates that identifies true SARS-CoV-2 infection but does not react with vaccine-induced antibodies (Fig. 3 and Extended Data Fig. 5). Four peptides representing immunodominant antigenic sites from ORF8, ORF3a, NSP3 and NPS12 were chemically synthesized and evaluated by IgG ELISA (Fig. 6). In addition, SARS-CoV-2 RBD and nucleocapsid proteins were used as controls. Although both peptides and proteins showed high specificity, only ORF3a (8,407–8,438) and ORF8 (9,301–9,314) peptides matched >90% sensitivity observed for the RBD and N protein. The absolute total binding antibodies against ORF8 and ORF3a antigenic site peptides varied; patients with severe disease showed higher antibody binding compared to those with mild/moderate disease versus controls. NSP3 (1,447–1,483) and NSP12 (5,164–5,218) showed high specificity with low reactivity. The combination of four peptides demonstrated 95% sensitivity for the SARS-CoV-2 sera (Fig. 6g). The RBD and N proteins demonstrated high reactivity to serum from children with COVID-19/MIS-C representing an aggregate sum of different antibodies recognizing multiple epitopes in these two large proteins compared with fewer overlapping epitopes in the shorter peptide sequence.

To substantiate the findings with children's sera, we tested serum from 72 pre-COVID influenza-infected adult controls (collected in 2010) and 100 sequential serum samples collected from 27 adults with COVID-19 (ref. [14]) for longitudinal antibody analysis against RBD or N, or individual peptides, or a mixture of four peptides in the serodiagnostic ELISA (Fig. 6). The antibody binding of post-COVID-19 adult samples to most peptides evolved with similar trends. In addition to RBD and N proteins, all individual peptides and peptide mixtures demonstrated >90% sensitivity and 100% specificity with SARS-CoV-2-infected adults.

Because SARS-CoV-2 spike is the major target for vaccine and therapeutic development, we evaluated the functional importance of GFPDL-identified sites within the spike protein as an immunogen to generate SARS-CoV-2-neutralizing antibodies. Two rabbits

each were immunized intramuscularly with either of five individual keyhole limpet hemocyanin (KLH)-conjugated synthetic peptides representing GFPDL-identified immunodominant antigenic sites (Extended Data Fig. 5), a control RSV-p27 peptide, SARS-CoV-2 prefusion spike protein, RBD or heat-killed SARS-CoV-2 particles. Sera from three of the five SARS-CoV-2 peptide immunized rabbits (but not p27 peptide) following the second immunization generated SARS-CoV-2-neutralizing antibodies, representing S7,363–7,430 in the RBD, S7,972–8,013 encompassing the fusion peptide, and S8,213–8,279 containing the heptad repeat 2 (HR2) sequence, in both the 293-ACE2 cell-based and Calu-3-based neutralization assays (Fig. 7). The fusion peptides of S2 (S7,972–8,013) and HR2 domain (S8,213–8,279) generated stronger neutralizing titers in Calu-3 cells, but were lower than those generated by the prefusion spike protein, RBD protein and virus particles (Fig. 7b). These rabbit studies confirmed that GFPDL-identified antigenic spike peptides are immunogenic, and that three of these peptides, including the RBD, conserved antigenic sites in fusion peptides and HR2 in the S2 domain, eliciting SARS-CoV-2-neutralizing antibodies.

Discussion

Deciphering the humoral response to SARS-CoV-2 infection in children is critical for understanding host–pathogen interaction and for developing diagnostics, therapeutics and vaccines in this important but understudied population. In this study, we provide a comprehensive SARS-CoV-2 proteome-wide in-depth antibody profiling in children with mild COVID-19 versus severe COVID-19 versus mild MIS-C versus severe MIS-C that revealed antibody markers associated with clinical disease severity. We identified serodiagnosis peptides and neutralizing targets that could guide rational design of diagnostics, therapeutics and vaccines in this critically underrepresented and understudied SARS-CoV-2 infection in children.

To summarize our key findings, SARS-CoV-2 proteome-wide antibody profiling revealed that higher antibody avidity to NSP2 and the SARS-CoV-2 prefusion spike protein was associated with decreased clinical disease severity and could potentially be used to develop a prognostic test of clinical disease outcomes in children.
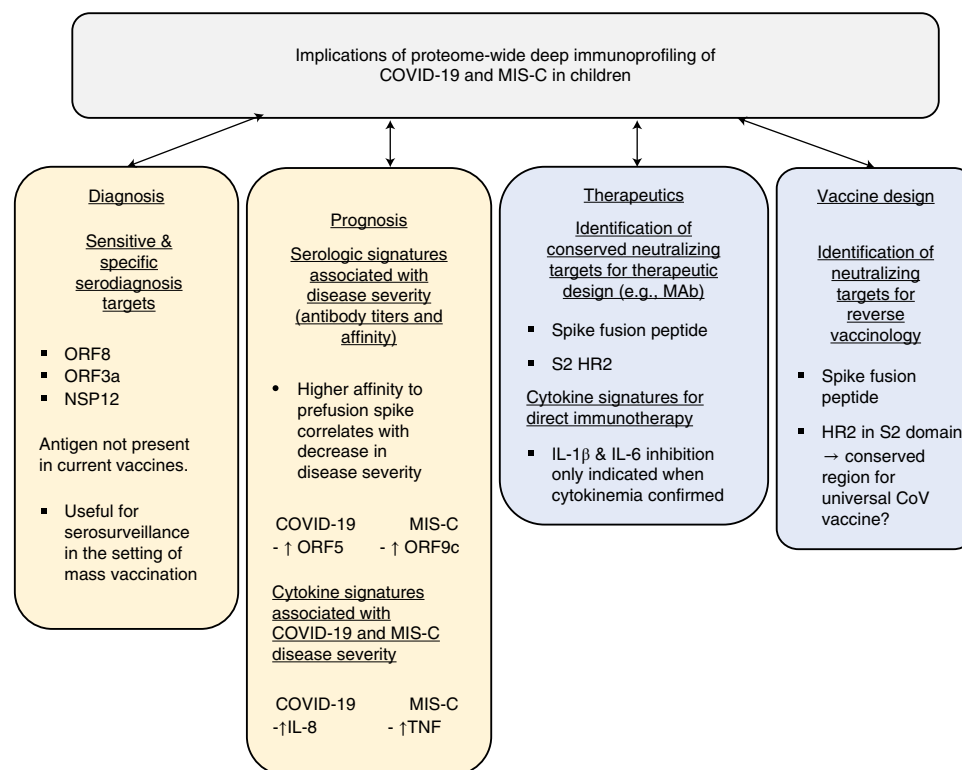
Fig. 8 | Implications of SARS-CoV-2 deep immunoprofiling of COVID-19 and MIS-C. Schematic highlighting the importance of the key messages for SARS-CoV-2-induced human antibody signature against the viral proteome in children with COVID-19 or MIS-C with different levels of disease severity for various diagnostic, prognostic, therapeutics and vaccine countermeasures against SARS-CoV-2.

Children with mild COVID-19 and mild MIS-C showed higher antibody binding and antibody avidity to NSP2 and NSP13, compared to children with severe COVID-19 or severe MIS-C. Moreover, we identified ORF8 and ORF3a peptides for differential serodiagnosis of SARS-CoV-2 infection, as well as the spike fusion peptide and HR2 in the S2 domain as potential candidates for development of better therapeutics and vaccines against SARS-CoV-2 (Fig. 8).

Although MIS-C is temporally associated with SARS-CoV-2 infection in children, its immunopathogenesis is still poorly understood. Consiglio et al. found that children with MIS-C lacked the IL-17-driven cytokine storm characteristic of Kawasaki disease[7]. In our study, we found that the distinguishing pro-inflammatory cytokines associated with disease severity in children with COVID-19 included IL-2, IL-6, IL-10, IL-13 and IL-17, complementing and extending the findings of Pierce et al.[24]. Children with MIS-C, however, had higher TNF, which may be indicative of a resolved viral infection or waning antiviral cytokinemia. The non-COVID-19 IVIG used for treatment early during the pandemic was unlikely to have any neutralizing antibodies to SARS-CoV-2 (ref. [25]). Sub-analyses of antibody fingerprinting between children treated without or with non-COVID IVIG did not reveal any differences in antibody response. Robust serologic and hyperinflammatory immune responses were observed in children with MIS-C despite administration of any immunomodulatory therapies. In our study, children with MIS-C did not have elevated levels of IL-1β, which brings into question the utility of these targeted biologic treatment modalities.

We observed a differentially evolving diverse antibody epitope repertoire and avidity maturation to various SARS-CoV-2 proteins in children with mild/moderate versus severe COVID-19 or MIS-C. Higher antibody avidity against NSP8 and prefusion spike protein (but not neutralization titers) was associated with lower clinical disease severity in these children. Antibodies against both

NSPs and spike protein could help in disease resolution possibly by Fc-mediated functions. Thus, vaccines or antibody-based therapeutics that can elicit high-affinity antibodies may have a substantial advantage for in vivo clinical outcomes in SARS-CoV-2 and contribute to amelioration of disease in infected individuals. Some of the current SARS-CoV-2 vaccines are based on a 'stabilized' prefusion spike protein[20]. The potential role of antibodies against the prefusion spike protein in protection from disease, provides support for the prefusion spike protein as a vaccine or therapeutic target[23,26,27].

Current vaccines have been shown to reduce disease severity, but they do not provide sterilizing immunity; therefore, vaccinated individuals get infected with SARS-CoV-2 and remain asymptomatic[28,29]. These breakthrough infections will be missed or under-reported if they are not tested in the acute phase of infection by PCR, because the SARS-CoV-2 PCR-based tests are typically negative after 2 weeks of SARS-CoV-2 infection. This means that the true number of breakthrough infections will not be captured and may result in transmission to susceptible unvaccinated individuals like children <12 years of age, who are not yet eligible for vaccination. To address VISP, we identified differential immunodominant epitopes covering non-spike and non-N antigenic sites that showed sensitive and specific antibody detection using ORF3a (8,407–8,438), ORF8 (9,301–9,314) and NSP12 peptide or the peptide combination. ORF3a and ORF8 were recently identified as serologic markers for SARS-CoV-2 infection in adults[4]. These peptide antigens are not part of current advanced vaccine candidates, are not present in virion particles, and are only expressed following SARS-CoV-2 infection. As SARS-CoV-2-specific antibodies are expected to persist for longer compared with SARS-CoV-2 RNA following viral infection, this serodiagnostic test may allow greater sensitivity to detect breakthrough infections long after the acute phase (1–4 weeks) of SARS-CoV-2 infection.

Although GFPDL-based assessments are likely to detect both conformational and linear epitopes on SARS-CoV-2 proteins, they are unlikely to detect paratopic interactions that require post-translational modifications and rare quaternary epitopes that cross protomers. However, ~90% of anti-SARS-CoV-2 antibodies either from post-vaccination or postinfection plasma were removed by adsorption with the SARS-CoV-2 GFPDL, supporting the use of the SARS-CoV-2 GFPDL for analyses of human sera[13–15]. Another limitation is our sample size for COVID-19 and MIS-C needs to be expanded to a larger cohort. Even with these limitations, we identified unique pediatric serum antibody-binding and avidity signatures against few SARS-CoV-2 proteins (prefusion spike, ORF3a and ORF9b) that could be used for development of diagnostic and/or prognostic tests for clinical outcomes and disease-severity classification of children with SARS-CoV-2 infection. In a recent study of elderly patients with COVID-19 where samples were collected every 1–3 d, an increase in prefusion spike-specific antibody affinity maturation was immediately followed by decrease in symptom severity[15]. Identification of fusion peptide and HR2 domain neutralization targets that are conserved among various SARS-CoV-2 variants may facilitate design of targeted therapeutics and more effective SARS-CoV-2 vaccines using reverse vaccinology approaches.

In conclusion, our comprehensive immunoprofiling across the SARS-CoV-2 proteome demonstrated differential epitope repertoire diversity, antibody specificity and antibody avidity maturation in children with COVID-19 or MIS-C with different levels of disease severity. Antibody avidity measurement against the prefusion spike protein can be used to develop a prognostic test for clinical disease outcomes in children. We identified differential serodiagnosis peptides and demonstrated an important role for the fusion peptide and HR2 in the S2 domain for the generation of SARS-CoV-2-neutralizing antibodies. These observations provide a more in-depth understanding of quantitative and qualitative aspects of immune responses that could aid in the development and evaluation of targeted, more effective SARS-CoV-2 serodiagnostics, therapeutics and next-generation vaccines for these understudied pediatric populations.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data and Supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-021-01051-8.

References

1. Lucas, C. et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463–469 (2020).
2. Mathew, D. et al. Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* **369**, eabc8511 (2020).
3. Ng, K. W. et al. Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. *Science* **370**, 1339–1343 (2020).
4. Hachim, A. et al. ORF8 and ORF3b antibodies are accurate serological markers of early and late SARS-CoV-2 infection. *Nat. Immunol.* **21**, 1293–1301 (2020).
5. Yonker, L. M. et al. Pediatric severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): clinical presentation, infectivity, and immune responses. *J. Pediatr.* **227**, 45–52 (2020).
6. Rostad, C. A. et al. Quantitative SARS-CoV-2 serology in children with multisystem inflammatory syndrome (MIS-C). *Pediatrics* **146**, e2020018242 (2020).
7. Consiglio, C. R. et al. The immunology of multisystem inflammatory syndrome in children with COVID-19. *Cell* **183**, 968–981 (2020).
8. Gruber, C. N. et al. Mapping systemic inflammation and antibody responses in multisystem inflammatory syndrome in children (MIS-C). *Cell* **183**, 982–995 (2020).
9. Khurana, S. et al. Antigenic fingerprinting of H5N1 avian influenza using convalescent sera and monoclonal antibodies reveals potential vaccine and diagnostic targets. *PLoS Med.* **6**, e1000049 (2009).
10. Khurana, S. et al. MF59 adjuvant enhances diversity and affinity of antibody-mediated immune response to pandemic influenza vaccines. *Sci. Transl. Med.* **3**, 85ra48 (2011).
11. Ravichandran, S. et al. Differential human antibody repertoires following Zika infection and the implications for serodiagnostics and disease outcome. *Nat. Commun.* **10**, 1943 (2019).
12. Khurana, S. et al. Longitudinal human antibody repertoire against complete viral proteome from Ebola virus survivor reveals protective sites for vaccine design. *Cell Host Microbe* **27**, 262–276 e264 (2020).
13. Ravichandran, S. et al. Antibody signature induced by SARS-CoV-2 spike protein immunogens in rabbits. *Sci. Transl. Med.* **12**, eabc3539 (2020).
14. Tang, J. et al. Antibody affinity maturation and plasma IgA associate with clinical outcome in hospitalized COVID-19 patients. *Nat. Commun.* **12**, 1221 (2021).
15. Ravichandran, S. et al. Longitudinal antibody repertoire in 'mild' versus 'severe' COVID-19 patients reveals immune markers associated with disease severity and resolution. *Sci. Adv.* **7**, eabf2467 (2021).
16. Jiang, H. W. et al. SARS-CoV-2 proteome microarray for global profiling of COVID-19-specific IgG and IgM responses. *Nat. Commun.* **11**, 3581 (2020).
17. Fuentes, S., Coyle, E. M., Beeler, J., Golding, H. & Khurana, S. Antigenic fingerprinting following primary rsv infection in young children identifies novel antigenic sites and reveals unlinked evolution of human antibody repertoires to fusion and attachment glycoproteins. *PLoS Pathog.* **12**, e1005554 (2016).
18. Khurana, S. et al. Human antibody repertoire after VSV-Ebola vaccination identifies novel targets and virus-neutralizing IgM antibodies. *Nat. Med.* **22**, 1439–1447 (2016).
19. Bartsch, Y. C. et al. Humoral signatures of protective and pathological SARS-CoV-2 infection in children. *Nat. Med.* **27**, 454–462 (2021).
20. Callaway, E. The race for coronavirus vaccines: a graphical guide. *Nature* **580**, 576–577 (2020).
21. Voysey, M. et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa and the UK. *Lancet* **10269**, P99–P111 (2020).
22. Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
23. Jackson, L. A. et al. An mRNA vaccine against SARS-CoV-2—preliminary report. *N. Engl. J. Med.* **383**, 1920–1931 (2020).
24. Pierce, C. A. et al. Immune responses to SARS-CoV-2 infection in hospitalized pediatric and adult patients. *Sci. Transl. Med.* **12**, eabd5487 (2020).
25. Tang, J. et al. Epitope diversity of SARS-CoV-2 hyperimmune intravenous human immunoglobulins and neutralization of variants of concern. *iScience* **24**, 103006 (2021).
26. Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
27. Graham, B. S. Rapid COVID-19 vaccine development. *Science* **368**, 945–946 (2020).
28. Hacisuleyman, E. et al. Vaccine breakthrough infections with SARS-CoV-2 variants. *N. Engl. J. Med.* **384**, 2212–2218 (2021).
29. CDC COVID-19 Vaccine Breakthrough Case Investigations Team. COVID-19 vaccine breakthrough infections reported to CDC—United States, 1 January–30 April 2021. *MMWR Morb. Mortal. Wkly. Rep.* **70**, 792–793 (2021).

Methods

Patient enrollment. Children and adolescents, from 0 to 21 years of age, hospitalized at CHOA with confirmed or suspected COVID-19 or MIS-C were enrolled into an institutional review board (IRB)-approved specimen collection protocol (Emory University IRB protocol nos. STUDY00022371 and STUDY00000723) following informed consent and assent as appropriate for age[6]. Children were classified as having MIS-C if they met the CDC case definition (https://emergency.cdc.gov/han/2020/han00432.asp; https://www.cdc.gov/mis-c/hcp/). They were classified as having acute COVID-19, per the CDC case definition (https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html), if they were hospitalized with symptomatic disease and had SARS-CoV-2 detected by real-time PCR (Supplementary Table 2). Clinical disease severity was classified as either severe or mild/moderate. Patients with severe disease were those who required ICU admission. Patients with mild/moderate disease consisted of children who were hospitalized with symptomatic COVID-19 or MIS-C, but who were never admitted to ICU (Supplementary Table 4). These criteria for disease classification and severity scales were previously described for SARS-CoV-2-infected children[6]. All children with COVID-19 or MIS-C were confirmed as having SARS-CoV-2 infection by the US Food and Drug Administration (FDA) Emergency Use Authorization-approved SARS-CoV-2 RT–PCR and/or SARS-CoV-2 antibody ELISA. The molecular tests performed at CHOA for inpatients during this time period included the DiaSorin Molecular Simplexa COVID-19 Direct RT–PCR assay[30], the Cepheid GeneXpert Xpress SARS-CoV-2 test[31] and the Hologic Panther Fusion SARS-CoV-2 test[32]. The qualitative SARS-CoV-2 nucleocapsid antibody test (Abbott Laboratories) was used for serologic testing. Prospective blood samples were collected from patients at the time of enrollment, and residual leftover samples from clinical laboratories were also obtained from the clinical laboratory. SARS-CoV-2-seropositive samples were collected 1 month after initial sample collection on a subset of enrolled patients. Specimens from children who were evaluated for COVID-19 or MIS-C, but tested negative, were utilized as hospitalized contemporanea controls. The controls were alike-disease hospitalized controls (not infected with SARS-CoV-2), which served as true disease-matched controls. Of these 14 control children, 6 had Kawasaki disease, 5 were admitted to ICU and were categorized as 'severe', and 9 were never admitted to ICU and served as the 'mild' disease comparator group. Because all the samples used in the study were collected from children during their stay in hospitalization, we have not considered any samples as convalescent. Convalescence refers to the later stage of an infectious disease or illness when the patient recovers and returns to previous health. Demographic, clinical, laboratory, imaging, treatment and outcome data were abstracted from the medical record and summarized in Supplementary Table 1, and individual data are provided in Supplementary Table 3. As our study samples were collected early during the pandemic, the effects of non-COVID IVIG administration on SARS-CoV-2 serology were likely negligible. The maximum pSOFA score within a 24-h period during hospitalization was calculated for each patient[33]. Healthy pediatric control samples from 1- to 3-year-old children were collected in 2009 (ref. [34]). The longitudinal plasma samples from patients hospitalized with COVID-19 were described previously[14].

Samples were tested in different antibody assays with approval from the FDA's Research Involving Human Subjects Committee (FDA-RIHSC) under exemption protocol '252-Determination-CBER-2020-08-19'. Sample were frozen on the day of collection and kept frozen below −70 °C. The samples were thawed only a single time on the day when different assays were run simultaneously. All samples were run in one batch at the same time for each assay. All the data are presented in an unbiased way, and there was no normalization unless explicitly mentioned in the text or figure legends.

Ethics statement for human studies. The study at the Center for Biologics Evaluation and Research (CBER), FDA was conducted with de-identified samples (Emory University IRB protocol nos. STUDY00022371 and STUDY0000723) under RIHSC protocol no. '252-Determination-CBER-2020-04-02'. Samples were collected under IRB-approved specimen collection protocol (Emory University IRB protocol nos. STUDY00022371 and STUDY00000723) after informed consent was obtained from at least one parent or legal guardian. All assays performed fell within the permissible usages in the original consent.

Proteins. Purified recombinant SARS-CoV-2 proteins that were produced in mammalian, insect or *Escherichia coli* cells or chemically synthesized were developed in-house or obtained from commercial sources (Supplementary Table 8).

The plasmids expressing genetically stabilized SARS-CoV-2 prefusion 2019-nCoV_S-2P spike ectodomain, and prefusion spike protein of different hCoVs, SARS-CoV-1, MERS, HKU, OC43, NL63 and 229E were a kind gift from B. Graham (VRC, National Institutes of Health (NIH))[26]. This expression vector was used to transiently transfect FreeStyle293F cells (Thermo Fisher, R79007) and protein was purified from filtered cell supernatants using StrepTactin resin (Cytiva, 29401326) and by size-exclusion chromatography and tested for binding to human ACE2 protein[13].

Measurement of cytokine levels in serum. All samples were diluted fourfold in Bio-Plex Sample Diluent HB buffer and analyzed via a Bio-Plex Pro Human Cytokine Panel 17-Plex assay (Bio-Rad, M5000031YV) per the manufacturer's instructions.

SARS-CoV-2 pseudovirus production and neutralization assay. Human codon-optimized cDNA encoding SARS-CoV-2 S glycoprotein (NC_045512) was cloned into eukaryotic cell expression vector pcDNA 3.1 between the *BamH*I and *Xho*I sites. Pseudovirions were produced by co-transfection of Lenti-X 293T cells (Takara, 632180) with psPAX2(gag/pol), pTrip-luc lentiviral vector and pcDNA 3.1 SARS-CoV-2-spike-deltaC19, using Lipofectamine 3000 (Thermo Fisher, L3000015). The supernatants were collected at 48 h after transfection and filtered through 0.45-μm membranes and titrated using 293T-ACE2 cells (HEK293T cells expressing ACE2 protein). The following reagent was obtained through BEI Resources, National Institute of Allergy and Infectious Diseases (NIAID), NIH: Human Embryonic Kidney Cells (HEK-293T) Expressing Human Angiotensin-Converting Enzyme 2, HEK-293T-hACE2 cell line, NR-52511.

For the neutralization assay, 50 μl of SARS-CoV-2 S pseudovirions were pre-incubated with an equal volume of varying dilutions of serum at RT for 1 h, then virus–antibody mixtures were added to 293T-ACE2 cells in a 96-well plate. After 3 h of incubation, the inoculum was replaced with fresh medium. After 24 h, cells were lysed and luciferase activity was measured. Controls included cell-only control, virus without any antibody control and positive control sera.

SARS-CoV-2 gene fragment phage display library. DNA plasmid encoding the full-length SARS-CoV-2 isolate Wuhan-Hu-1 strain (GenBank: MN908947.3) was a kind gift from V. Thiel and used for cloning[35]. A gIII display-based in-house phage vector, fSK-9-3, was used where the desired polypeptide can be displayed on the surface of the phage as a gIII-fusion protein. Purified DNA containing the entire SARS-CoV-2 genome was digested with DNase I (NEB, M0303L) to obtain gene fragments of 50–1,500 bp in size (18 to 500 amino acids) and used for GFPDL construction[12,13,18].

Adsorption of SARS-CoV-2 serum on SARS-CoV-2 GFPDL phages. Before panning of GFPDL, 500 μl of tenfold-diluted serum antibodies from SARS-CoV-2-infected samples was adsorbed by incubation with SARS-CoV-2 GFPDL phage-coated Petri dishes. To ascertain the residual antibodies specificity, an ELISA was performed with wells coated with 10^4 per 100 μl of heat-inactivated SARS-CoV-2 virion particles. After blocking with PBST (PBS containing 0.1 % Tween-20) containing 2% BSA, serial dilutions of human serum (with or without adsorption) in blocking solution were added to each well, incubated for 1 h at room temperature (RT), followed by addition of 5,000-fold-diluted HRP-conjugated goat anti-human IgA + IgG + IgM-specific antibody (Jackson ImmunoResearch, 109-035-064) and developed by 100 μl of OPD substrate (Thermo Fisher, 34006) solution. Absorbance was measured at 490 nm.

Affinity selection of SARS-CoV-2 GFPDL with human serum. Before panning of GFPDL with polyclonal serum antibodies, serum components, which could nonspecifically interact with phage proteins, were removed by incubation with ultraviolet-killed M13K07 (NEB, N0315S) phage-coated Petri dishes[13,18].

Equal volumes of first-drawn time-point samples from the same time point (DPoS) to match the days after onset of symptoms from each patient in the corresponding cohort were pooled from the control, COVID-19 or MIS-C samples from mild/moderate (M) or severe (S) disease and were used for GFPDL panning. GFPDL in-solution affinity selection was carried out with anti-IgM, or protein A/G (IgG), or anti-IgA-specific affinity resin[12,13,36]. Briefly, the pooled serum was incubated with the full-length SARS-CoV-2 GFPDL and the specific resin, and the unbound phages were removed by PBST wash followed by PBS. Bound phages were eluted by addition of 0.1 N Gly-HCl (pH 2.2). After panning, antibody-bound phage clones were amplified, the inserts were sequenced, and the sequences were aligned to the SARS-CoV-2 genome.

The GFPDL affinity selection was performed in duplicate (two independent experiments by a researcher in the laboratory, who was blinded to sample identity). Similar numbers of bound phage clones and epitope repertoire were observed by GFPDL panning.

Surface plasmon resonance-based antibody-binding kinetics of human serum. Steady-state equilibrium binding of pediatric serum was monitored at 25 °C using a ProteOn SPR (Bio-Rad). The purified recombinant SARS-CoV-2 or other proteins were captured to a Ni-NTA sensor chip (Bio-Rad, 176-5031) with 200 RUs in the test flow channels. The protein density on the chip was optimized so as to measure monovalent interactions independent of the antibody isotype[12–15,37]. Serial dilutions (10-, 30- and 90-fold) of freshly prepared sample in BSA-PBST buffer (PBS pH 7.4 buffer with Tween-20 and BSA) were injected at a flow rate of 50 μl min⁻¹ (120-s contact duration) for association, and disassociation was performed over a 600-s interval. Responses from the protein surface were corrected for the response from a mock surface and for responses from a buffer-only injection. SPR was performed with serially diluted serum of each individual time point in this study. Total antibody binding was calculated

with Bio-Rad ProteOn manager software (version 3.1). All SPR experiments were performed twice, and the researchers performing the assay were blinded to sample identity. In these optimized SPR conditions, the variation for each sample in duplicate SPR runs was <5%. The maximum RU data shown in the figures are the RU signal for the tenfold-diluted serum sample.

Antibody off-rate constants, which describe the stability of the antigen–antibody complex, that is, the fraction of complexes that decays per second in the dissociation phase, were determined directly from the human polyclonal sample interaction with recombinant purified SARS-CoV-2 proteins using SPR. To that end, serially diluted samples at 10-, 30- and 90-fold dilutions were analyzed to determine antibody off-rate constants in the dissociation phase only for the sensorgrams with maximum values in the range of 5–150 RUs and calculated using the Bio-Rad ProteOn manager software for the heterogeneous sample model as described before[10,12–15,37]. Off-rate constants were determined from two independent SPR runs. All SPR experiments were performed twice, and the researchers performing the assay were blinded to sample identity. The variation for each sample in duplicate SPR runs was <5%. The data shown are average values of two experimental runs.

Seroreactivity to SARS-CoV-2 antigenic peptides by ELISA. Ninety-six-well Immulon plates were coated with 100 ng/100 µl of recombinant spike, N protein or synthetic peptides in PBS overnight at 4 °C. Starting at a 1:100 dilution, serum samples were serially diluted at a 1:5 ratio and applied to the coated well for 1 h at ambient temperature. Serum samples were assayed in duplicate. Naïve serum samples were assayed along with the experimental samples. After three washes with PBS/0.05% Tween 20, bound antibodies were detected with a 1:5,000 dilution of horseradish peroxidase (HRP)-conjugated anti-human IgG Fc-specific antibody (Jackson ImmunoResearch, 109-035-098). After 1 h, plates were washed as before and o-phenylenediamine dihydrochloride (oPD) was added for 10 min. Absorbance was measured at 492 nm. The absorbance shown is reported for the 100-fold-diluted serum/plasma sample. Cutoff values for each antigen in ELISA were determined by the area under the ROC curve analysis of the negative controls (using samples from ten hospitalized control children at 100-fold serum dilution) in GraphPad Prism.

Rabbit immunization studies. GFPDL-selected immunodominant SARS-CoV-2 spike peptide conjugation to Maleimide Activated KLH was performed as described in the product manual of Imject Maleimide Activated KLH (Thermo Scientific, 77605). Female New Zealand white rabbits (4–6 weeks old) were immunized twice intra-muscularly at 21-d intervals with 25 µg of KLH-conjugated peptides or 50 µg of purified recombinant proteins or heat-inactivated 10^5 plaque-forming units of SARS-CoV-2 virion particles mixed with Emulsigen Adjuvant. Two rabbits were used per immunogen. Sera collected before vaccination and 8 d after the second vaccination were analyzed in a pseudovirion neutralization assay.

All animal experiments were approved by the FDA Institutional Animal Care and Use Committee under protocol no. 2008-10. This protocol meets NIH guidelines.

Statistical analysis. All experimental data to compare differences between groups were analyzed using lme4 and emmeans packages in R (RStudio version 1.1.463). Data from SPR (antibody-binding and antibody off-rates), cytokine and neutralization titers (used as independent variables) were analyzed for statistical significance among various disease groups to control for age, gender and DPoS, which served as covariates (predictor variables), using a linear regression model. To ensure robustness of the results, absolute measurements were $\log_2$ transformed before performing the analysis. Within each disease group, the respective measurement was regressed on age, gender and DPoS. For comparisons between various disease categories (factor variable; mild and severe controls, mild and severe COVID-19, mild and severe MIS-C), pairwise comparisons were extracted using emmeans, and Tukey-adjusted P values were used for denoting significance to reduce type 1 error due to multiple testing. The tests were two sided. The statistical analysis tested for significant differences in cytokine, antibody binding, antibody off-rates and neutralization titer measurements among various disease categories. For the temporal analysis and correlations, the DPoS values for COVID-19 and MIS-C were controlled statistically by multivariable analysis in R. Detailed statistical data for the figures and extended data figures are provided in Supplementary Table 9.

PCA analysis of 74 antibody-binding and avidity variables and 17 cytokines/chemokines from 74 samples was conducted using R packages factoMineR and factoextra. Antibody-binding variables (maximum RUs), avidity (off rate) and cytokines/chemokines (concentration) data were $\log_{10}$ transformed before analysis. Correlation and regression analyses were performed by computing Pearson's coefficient in GraphPad Prism. Multivariate linear regression models for cytokine and clinical disease (Fig. 1b,p) were performed in R and used to calculate Spearman's rank correlation coefficient and significance.

For rabbit data, differences among groups were analyzed using multiple-group comparisons by non-parametric (Kruskal–Wallis) statistical testing using Dunn's post hoc analysis in GraphPad Prism.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
Antigenic sites were depicted on the SARS-CoV-2 structures (https://www.rcsb.org/). Sequences were downloaded from https://www.ncbi.nlm.nih.gov/genbank/. SARS-CoV-2 proteome (GenBank no. NC_045512.2); SARS-CoV-1 BJ01 strain GenBank ID or protein ID: ORF1ab (AAP30028.1), spike (AAP30030.1), ORF3a (AAP30031.1), E (AAP30033.1), M (AAP30034.1), N (AAP30037.1) and ORF10 (civet, UnitProt no. Q3ZTD2). MERS CoV GenBank or protein ID: ORF1ab (AKS48060.1), spike (AKN11075.1), ORF3a (YP_009047205.1), E (YP_009047209.1), M (YP_009047210.1) and N (AKS48069.1); human CoV NL63 GenBank ID or protein ID: ORF1ab (YP_003766.2), spike (YP_003767.1), ORF3a (UnitProt no. Q06XP3), E (YP_003769.1) and M (YP_003770.1); human CoV 229E GenBank ID or protein ID: ORF1ab (ARU07599.1), spike (ARU07601.1), ORF3a (ORF4; ARU07602.1), E (ARU07603.1), M (ARU07604.1), N (ARU07605.1); human CoV HKU-1 GenBank or protein ID: ORF1ab (YP_173236.1), spike (YP_173238.1), ORF3a (YP_173239.1), E (YP_173240.1), M (AYN64564.1) and N (UnitProt no. Q0ZME3); human CoV OC43 GenBank or Protein ID: ORF1ab (AAT84359.1), ORF3a (NS5A; AOL02454.1), M (YP_009555244.1), E (YP_009555243.1) and N (UnitProt no. P33469). Homology searches for selected antigenic sites with the human protein database were performed with blastp using the NCBI human sequence database (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). Source data are provided with this paper. All other data generated in this article are provided in the figures and Supplementary Information.

References
30. Bordi, L. et al. Rapid and sensitive detection of SARS-CoV-2 RNA using the Simplexa COVID-19 direct assay. *J. Clin. Virol.* **128**, 104416 (2020).
31. Goldenberger, D. et al. Brief validation of the novel GeneXpert Xpress SARS-CoV-2 PCR assay. *J. Virol. Methods* **284**, 113925 (2020).
32. Hogan, C. A. et al. Five-minute point-of-care testing for SARS-CoV-2: not there yet. *J. Clin. Virol.* **128**, 104410 (2020).
33. Matics, T. J. & Sanchez-Pinto, L. N. Adaptation and validation of a pediatric sequential organ failure assessment score and evaluation of the sepsis-3 definitions in critically ill children. *JAMA Pediatr.* **171**, e172352 (2017).
34. Verma, N. et al. Influenza virus H1N1pdm09 infections in the young and old: evidence of greater antibody diversity and affinity for the hemagglutinin globular head domain (HA1 domain) in the elderly than in young adults and children. *J. Virol.* **86**, 5515–5522 (2012).
35. Thi Nhu Thao, T. et al. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *Nature* **582**, 561–565 (2020).
36. Fuentes, S. et al. Antigenic fingerprinting of respiratory syncytial virus (RSV)-A-infected hematopoietic cell transplant recipients reveals importance of mucosal anti-RSV G antibodies in control of RSV infection in humans. *J. Infect. Dis.* **221**, 636–646 (2020).
37. Khurana, S. et al. Repeat vaccination reduces antibody affinity maturation across different influenza vaccine platforms in humans. *Nat. Commun.* **10**, 3338 (2019).

Author contributions
Designed research: S.K. Performed research: S.R., J.T., G.G., Y.L., S.P. and S.K. Conducted clinical study, contributed clinical samples and provided clinical data: L.H., S.A.L., R.C.J.,

V.S., A.C., E.J.A. and C.A.R. Conducted rabbit study and SARS-CoV-2 neutralization assays: G.G., J.T. and S.K. Contributed to writing: S.K.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41590-021-01051-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-021-01051-8.

Correspondence and requests for materials should be addressed to Surender Khurana.

Peer review information *Nature Immunology* thanks Carrie Lucas, Paul Moss and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. N. Bernard was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.
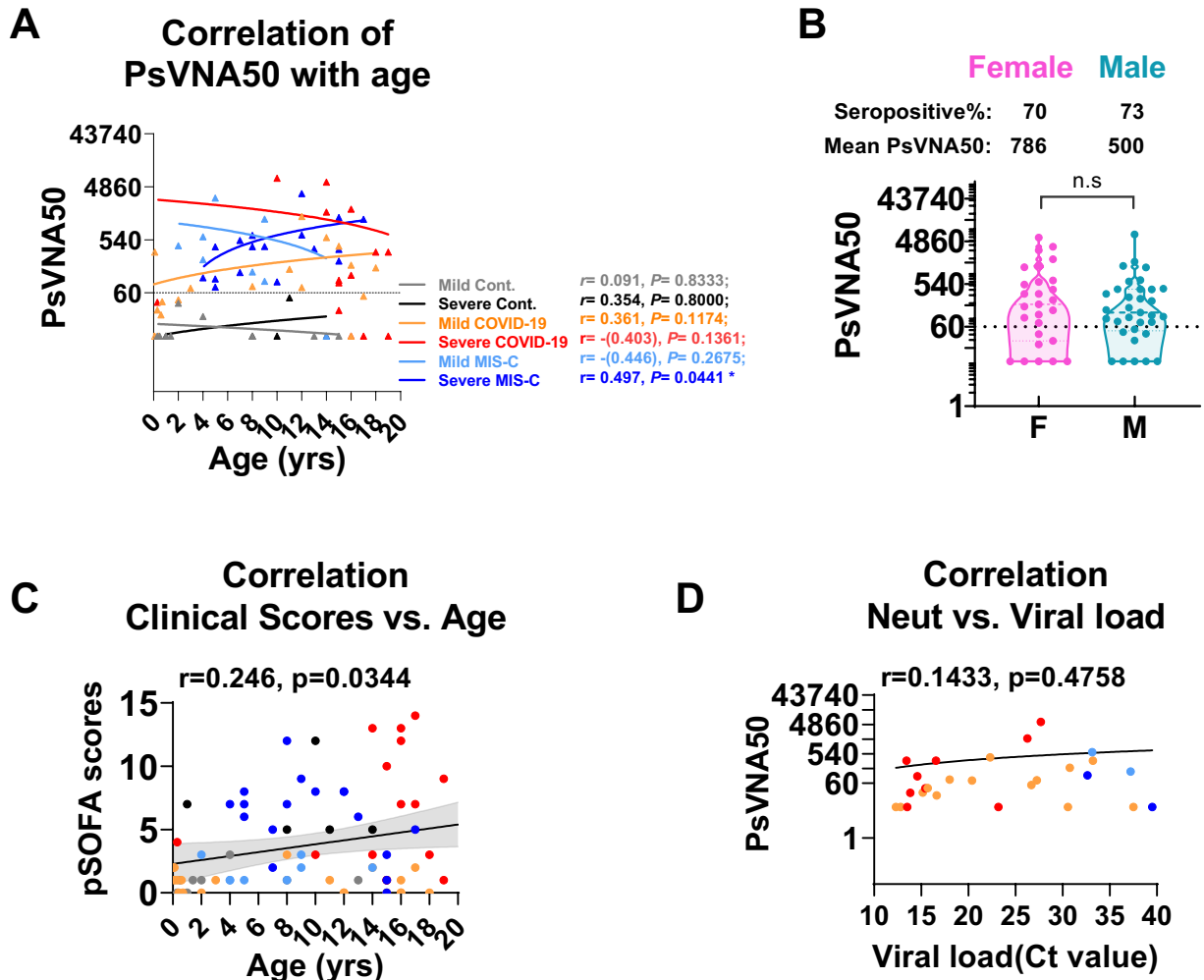
Extended Data Fig. 1 | Neutralization antibody levels in children with COVID-19 or MIS-C. (a) Correlation analysis between neutralization antibody levels of the patients with age in different disease groups. Dotted line indicates the neutralization titer cutoff (PsVNA50 of 60) for seropositive determination. The trend line is depicted as solid colored line. (b) Neutralization antibody levels between males (M) vs females (F). **(c)** Correlation analysis between age of the patients and their pSOFA clinical scores. The trendline is depicted as solid center line with the error bands representing 95% confidence interval shown as shaded area. **(d)** Correlation analysis between neutralizing titers with viral load (measured by RT-qPCR). Correlation analysis was performed using linear regression model and associated Pearson's correlation coefficients (r) and linear regression significance (p) are shown. The trend line is depicted as solid line with the error bands representing 95% confidence interval shown as shaded colored area. Differences among groups were performed using R. The tests were two-sided using Tukey multiple group comparisons to reduce Type 1 error. The differences were considered statistically significant with a 95% confidence interval when the p value was less than 0.05. (* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001).
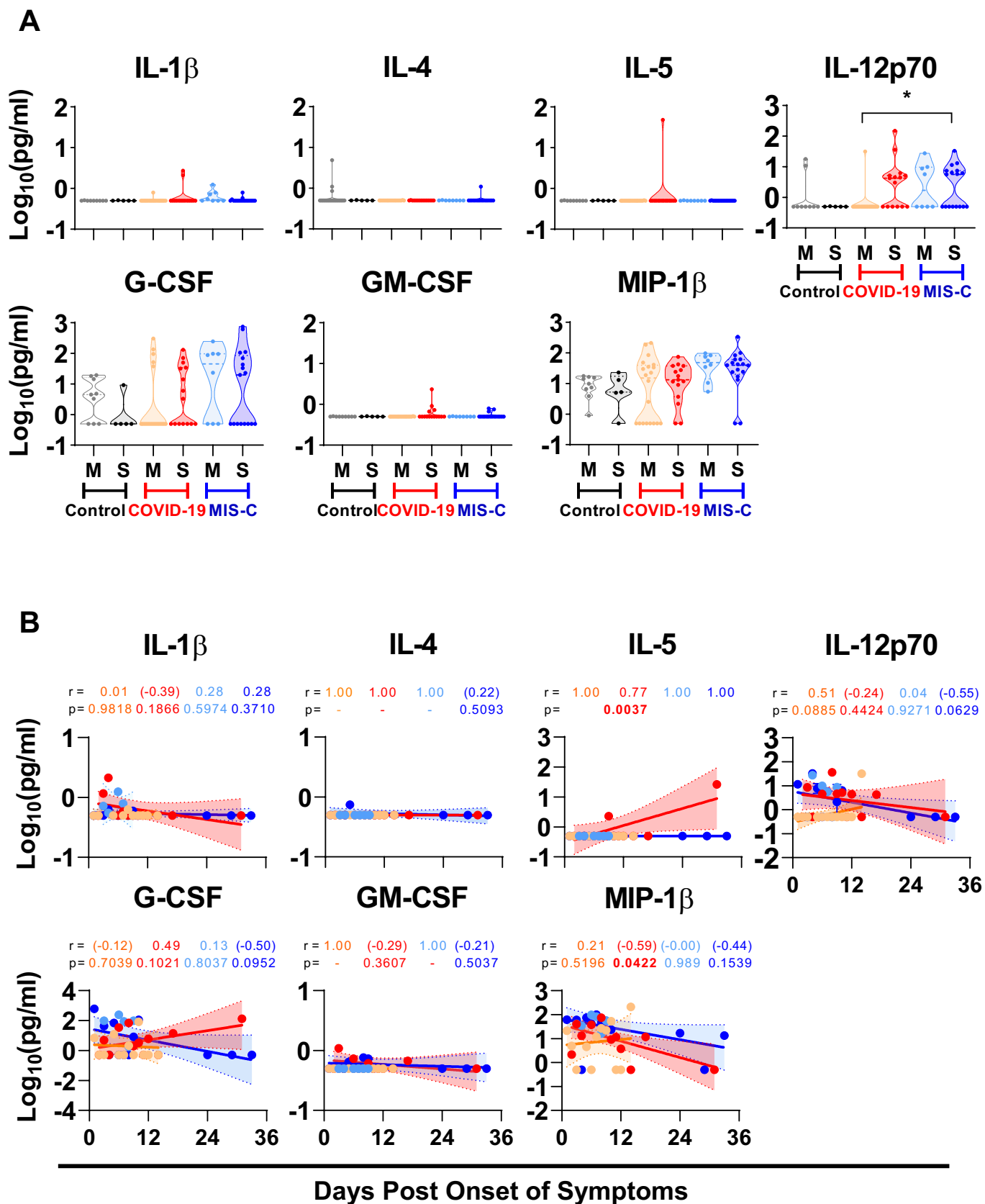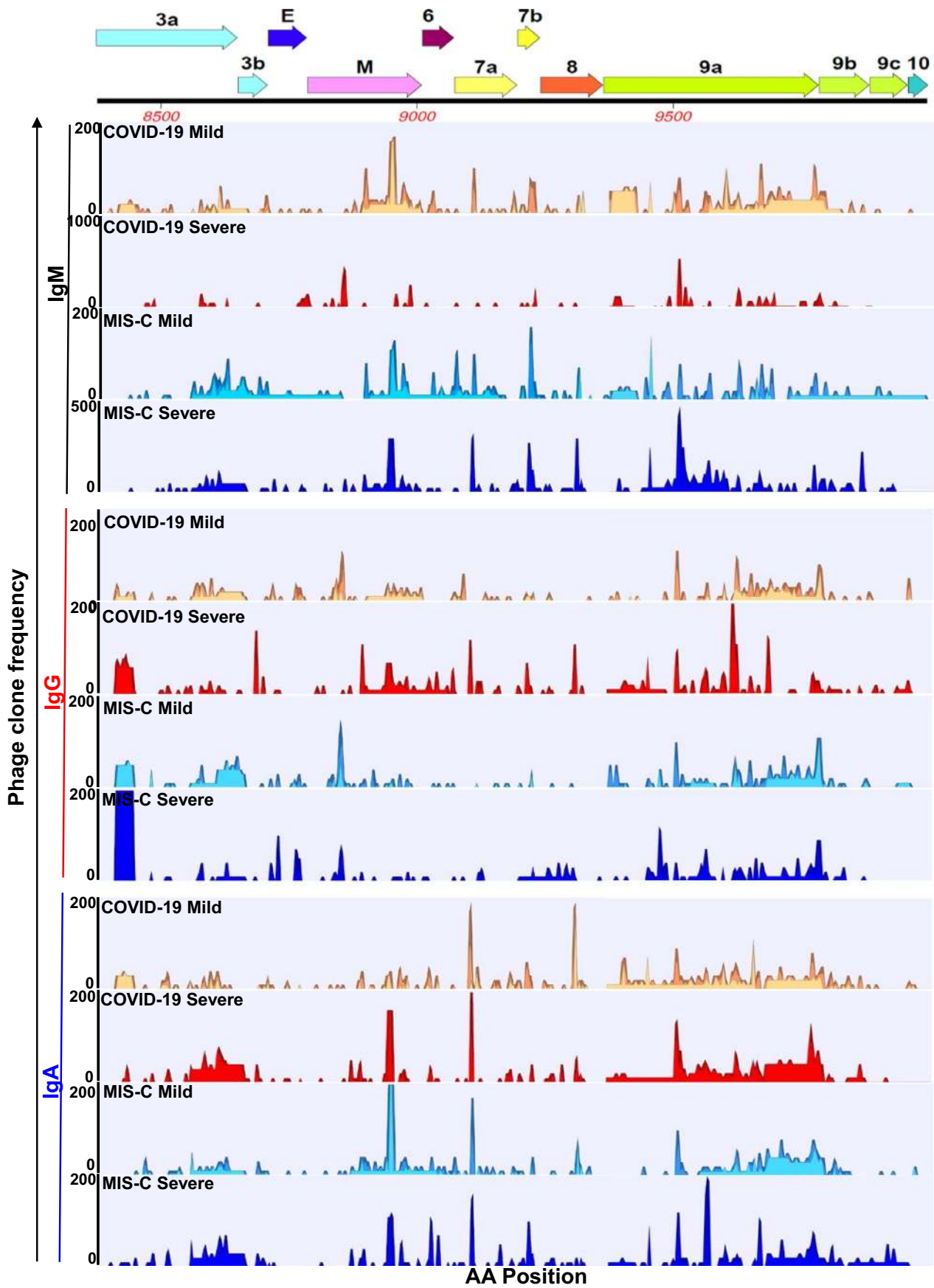
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Cytokine/chemokine levels in serum samples from each patient. (a) Cytokine/chemokine levels of 4-fold diluted serum samples were analyzed by a Bio-Plex Pro Human Cytokine Panel 17-Plex assay. Rest of cytokines are shown in Fig. 2. The samples are color coded by groups.: grey, Mild/moderate Control (n = 9); black, Severe control (n = 5); orange, Mild/moderate COVID-19 (n = 20); red, Severe COVID-19 (n = 15); cyan, Mild/moderate MIS-C (n = 8); blue, Severe MIS-C (n = 17). Differences among groups were performed using R. The tests were two-sided using Tukey multiple group comparisons to reduce Type 1 error. The differences were considered statistically significant with a 95% confidence interval when the p value was less than 0.05. (* $P \leq 0.05$). **(b)** Correlation of cytokine and chemokines with days post-onset of symptoms in first time-point drawn serum of children with mild COVID-19 vs severe COVID-19 vs severe MIS-C. Cytokine and chemokine measurements reported over days post-onset of symptoms and continuously as linear regressions. Each dot represents a distinct patient and days post-onset of symptoms: orange, Mild/moderate COVID-19 (n = 20); red, Severe COVID-19 (n = 15); cyan, Mild/moderate MIS-C (n = 8); blue, Severe MIS-C (n = 17), using linear regression model in GraphPad Prism. Associated Pearson's correlation coefficients (r) and linear regression significance (p) are colored accordingly. The trendline is depicted as solid colored center line with the error bands representing 95% confidence interval shown as shaded colored area.

A

ORFs

NSPs

B

Anti-SARS-CoV-2 virion ELISA
following GFPDL adsorption

- ● COVID-19
- ● COVID-19 - GFPDL Ads
- ● MIS-C
- ● MIS-C - GFPDL Ads
- ● Control Sera

C

Sample	Phage titers/mL		
	IgM	IgG	IgA
Mild COVID-19	1.25E+05	8.60E+04	1.42E+05
Severe COVID-19	1.02E+06	1.01E+05	1.45E+05
Mild MIS-C	2.14E+05	4.80E+04	1.72E+05
Severe MIS-C	3.12E+05	4.00E+04	3.16E+04
Uninfected Severe Control	6.78E+02	1.80E+02	1.17E+02

D

ORFs

NSPs

Amino Acid Position
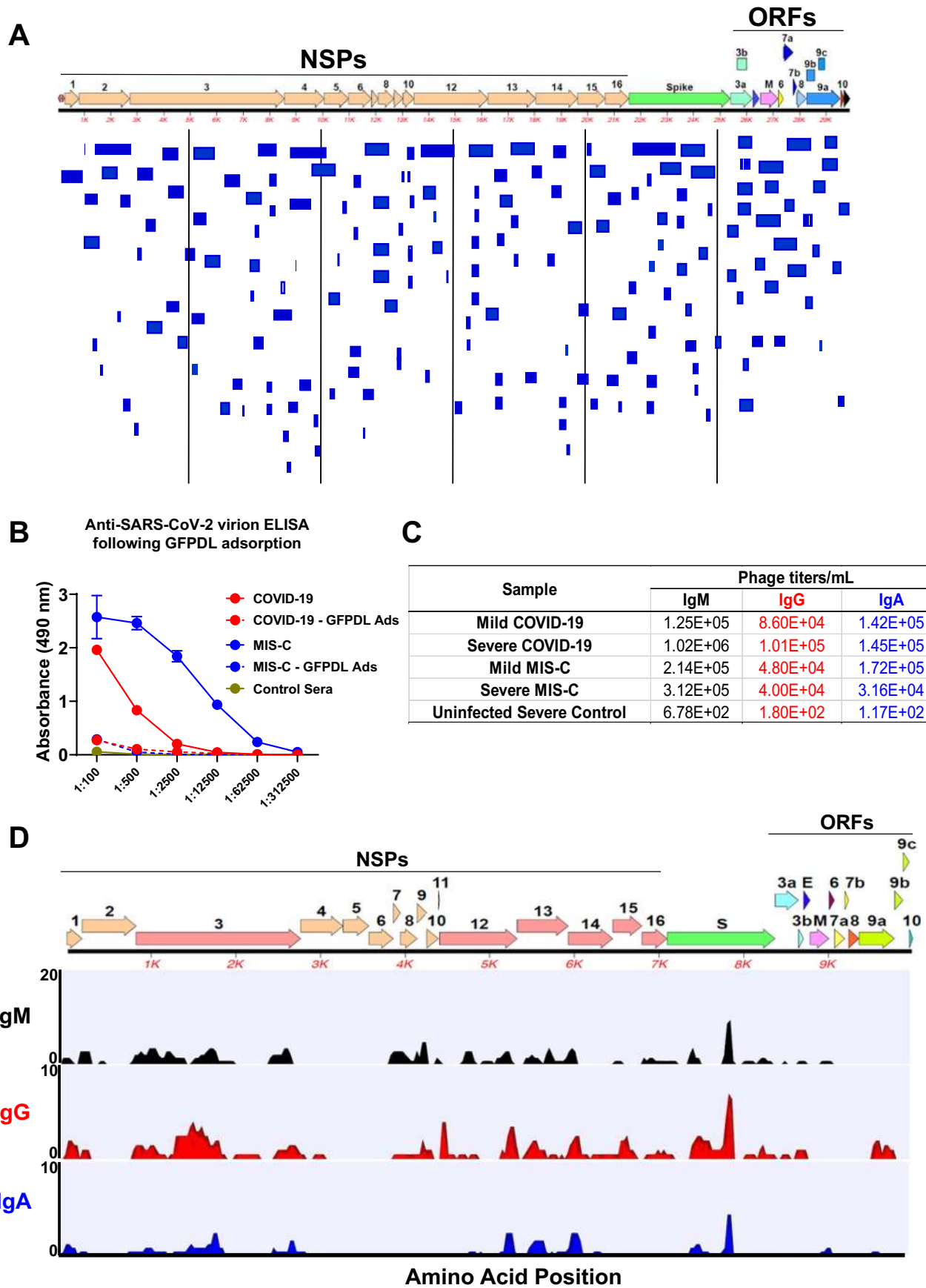
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Characterization of SARS-CoV-2 GFPDL. (a) Random distribution of size and sequence of the SARS-CoV-2 GFPDL. Sequencing of SARS-CoV-2 Wuhan-Hu-1 strain sequences expressed by the phages of SARS-CoV-2 GFPDL were aligned to SARS-CoV-2 Wuhan-Hu-1 whole proteome translated sequence. **(b)** Anti-SARS-CoV-2 reactivity of serum from COVID-19 or MIS-C children in ELISA before and after SARS-CoV-2 GFPDL adsorption. Sera from either severe COVID-19 or severe MIS-C children was adsorbed on SARS-CoV-2 GFPDL coated petri dishes. Binding to SARS-CoV-2 virus particles coated wells is shown for serum antibodies before (solid lines) and after (dashed lines) GFPDL-adsorption in ELISA using 1:5000 dilution of HRP-conjugated goat anti-human IgA + IgG + IgM specific antibody. The data is presented as mean values +/- SEM for two biologically independent experiments. **(c)** Number of IgM, IgG and IgA binding phage clones selected using SARS-CoV-2 GFPDL with COVID-19 and MIS-C serum of different disease severity (mild vs severe) and hospitalized severe controls. **(d)** IgM, IgG and IgA epitope repertoire recognized in the control (SARS-CoV-2 negative) hospitalized children and their alignment to the whole proteome of SARS-CoV-2 Wuhan-Hu-1 strain (GenBank: MN908947.3) showing four structural proteins (S, E, M and N) as well as 25 putative non-structural proteins. Graphical distribution of representative clones with a frequency of >2, obtained after affinity selection, are shown. Epitope coverage map showing alignment position (amino acid positions) on X-axis and coverage (clone frequency units) at each position on Y-axis. Data and alignment analysis were done using MacVector with Assembler (v 12.7.5) and CLC Genomics Workbench 20 (v 20.0.4). The GFPDL affinity selection data was performed in duplicate (two independent experiments by researcher in the lab, who was blinded to sample identity), and similar number of phage clones and epitope repertoire was observed in both phage display analysis.
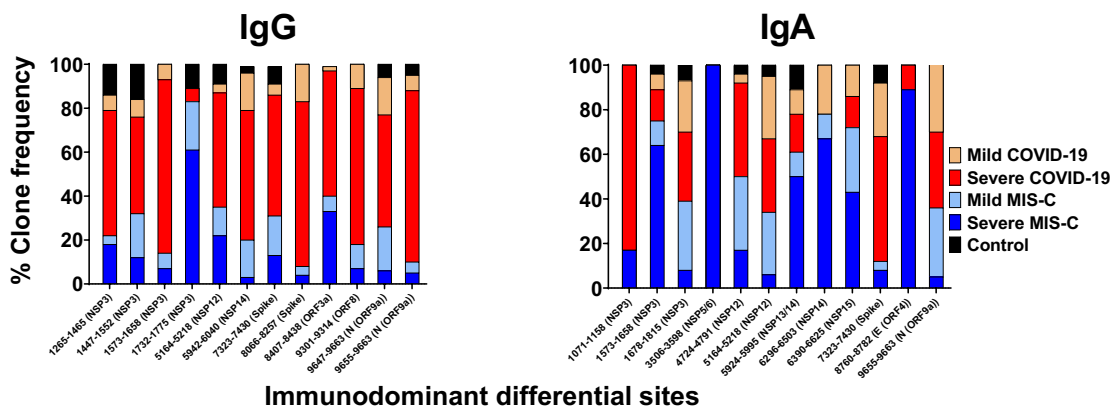
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | SARS-CoV-2 ORF proteins IgM, IgG and IgA epitope repertoires in MIS-C or COVID-19. IgM, IgG and IgA antibody epitope repertoire recognized in the SARS-CoV-2 infected sera from children with mild/moderate or severe COVID-19, or with mild/moderate or severe MIS-C across the SARS-CoV-2 encoded ORF proteins 3a to 10. The schematic on top shows various ORF proteins of SARS-CoV-2: ORF3a (3a), ORF4 **(e)**, ORF5 (M), ORF7a (7a), ORF8 (8), ORF9a (N), ORF9b (9b), ORF9c (9c) and ORF10 (10). Epitope coverage map shows alignment position (amino acid position) on the X-axis and coverage (phage clone frequency) of each position on the Y-axis defined by epitopes recognized by IgG, IgM and IgA in mild/moderate or severe COVID-19 (orange and red) or MIS-C (cyan and blue), respectively. Owing to higher frequency of selected phage clones, the original clone frequency (Y-axis) of ORF3a antigenic sites were scaled with a scaling ratio of 3.05 for ORF3a site (8407-8438) to aid in data visualization.
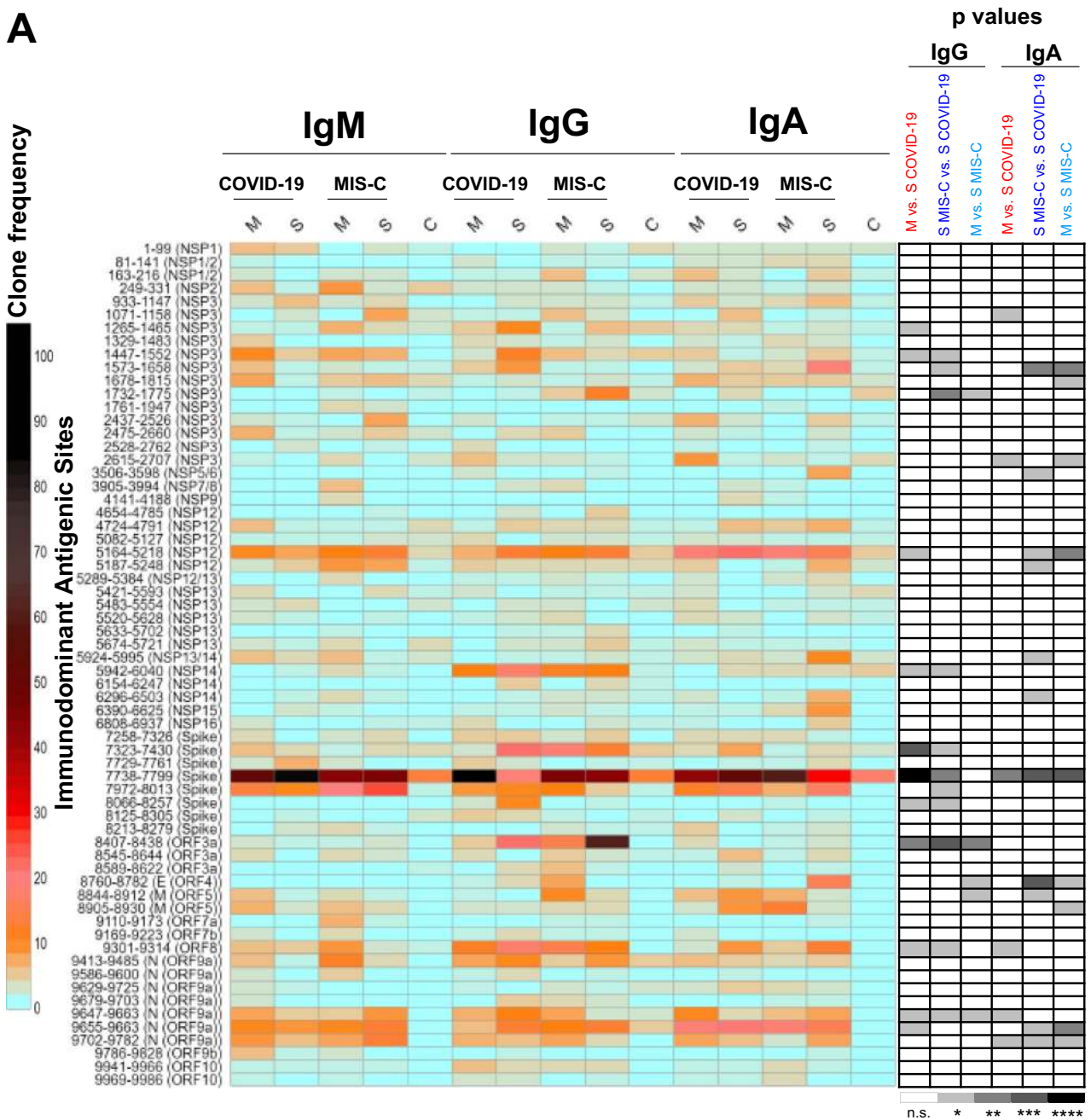
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | SARS-CoV-2 immunodominant epitope profile of children with COVID-19 or MIS-C. (a) Heat map of immunodominant sites (≥ 3 clonal frequency in at least one sample) on the SARS-CoV-2 proteome. IgM, IgG and IgA antibodies in mild (M) or severe (S) serum samples from COVID-19, MIS-C patients or severe control **(c)** samples identified using GFPDL analyses. The immunodominant sites on the left indicate amino acid residue of the antigenic sites and the region it encompasses on the proteome in parentheses. Color scale on the left represents range of clonal occurrences (frequency) of each site. Heat map was generated using R. Significant values for frequency of sites recognized by mild vs. severe COVID-19 and severe COVID-19 vs. severe MIS-C and mild vs. severe MIS-C samples for IgG and IgA were determined by Unpaired t-test and are color coded (white-n.s., light grey-$p < 0.05$, grey-$p < 0.01$, dark grey-$p < 0.001$, black-$p < 0.0001$). **(b)** Frequency of clones of each of statistically significant IgG or IgA differential sites between mild vs severe COVID-19 or MIS-C is shown for all disease cohorts.
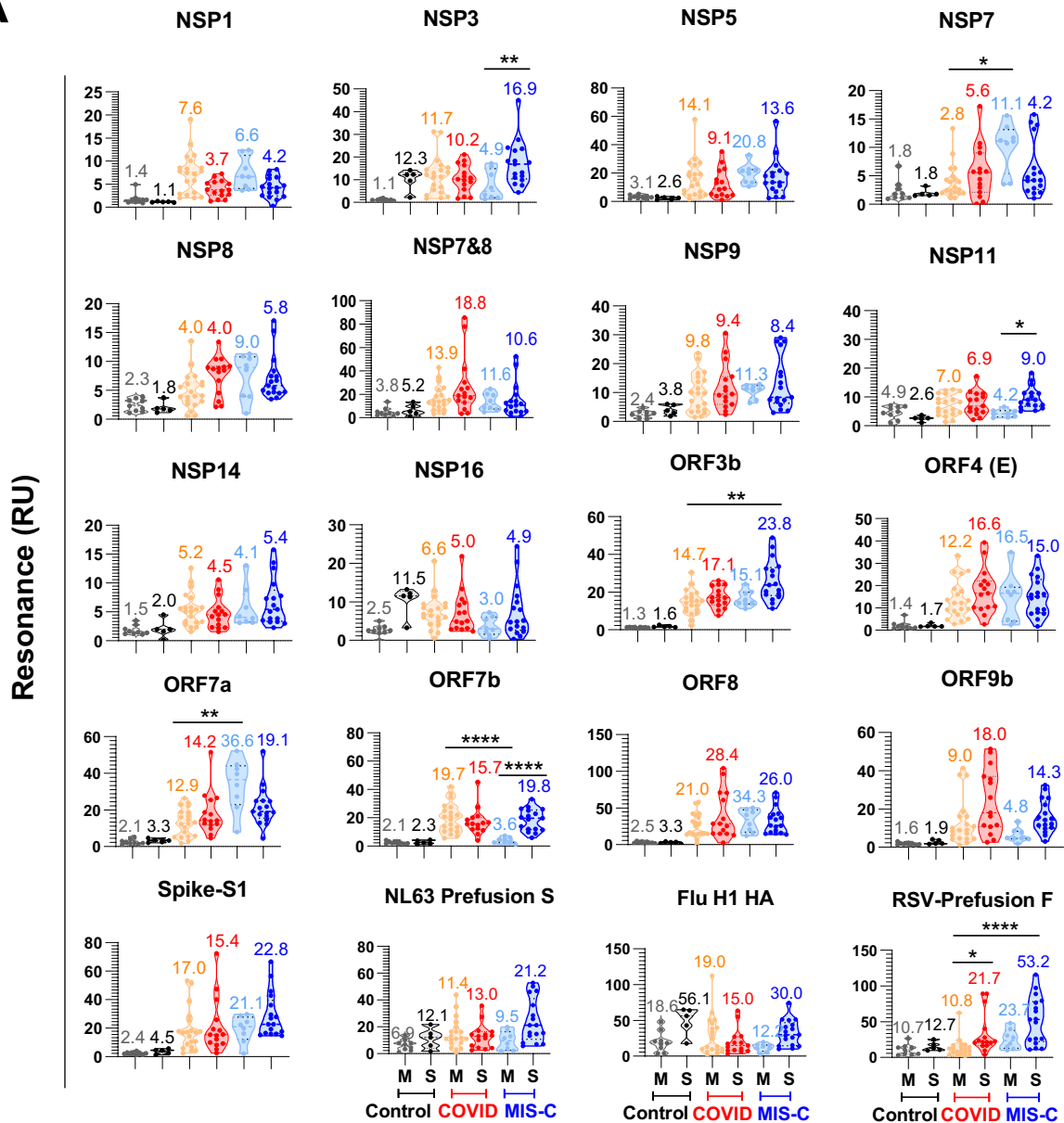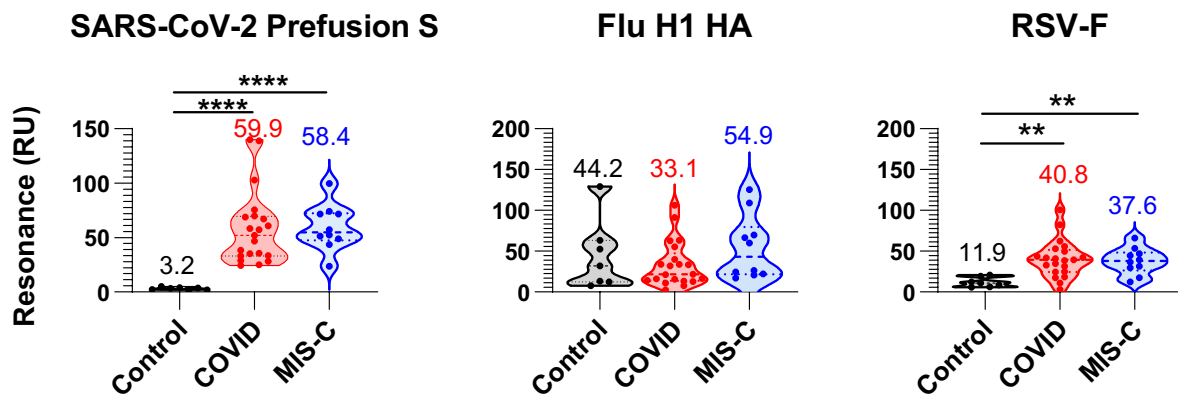
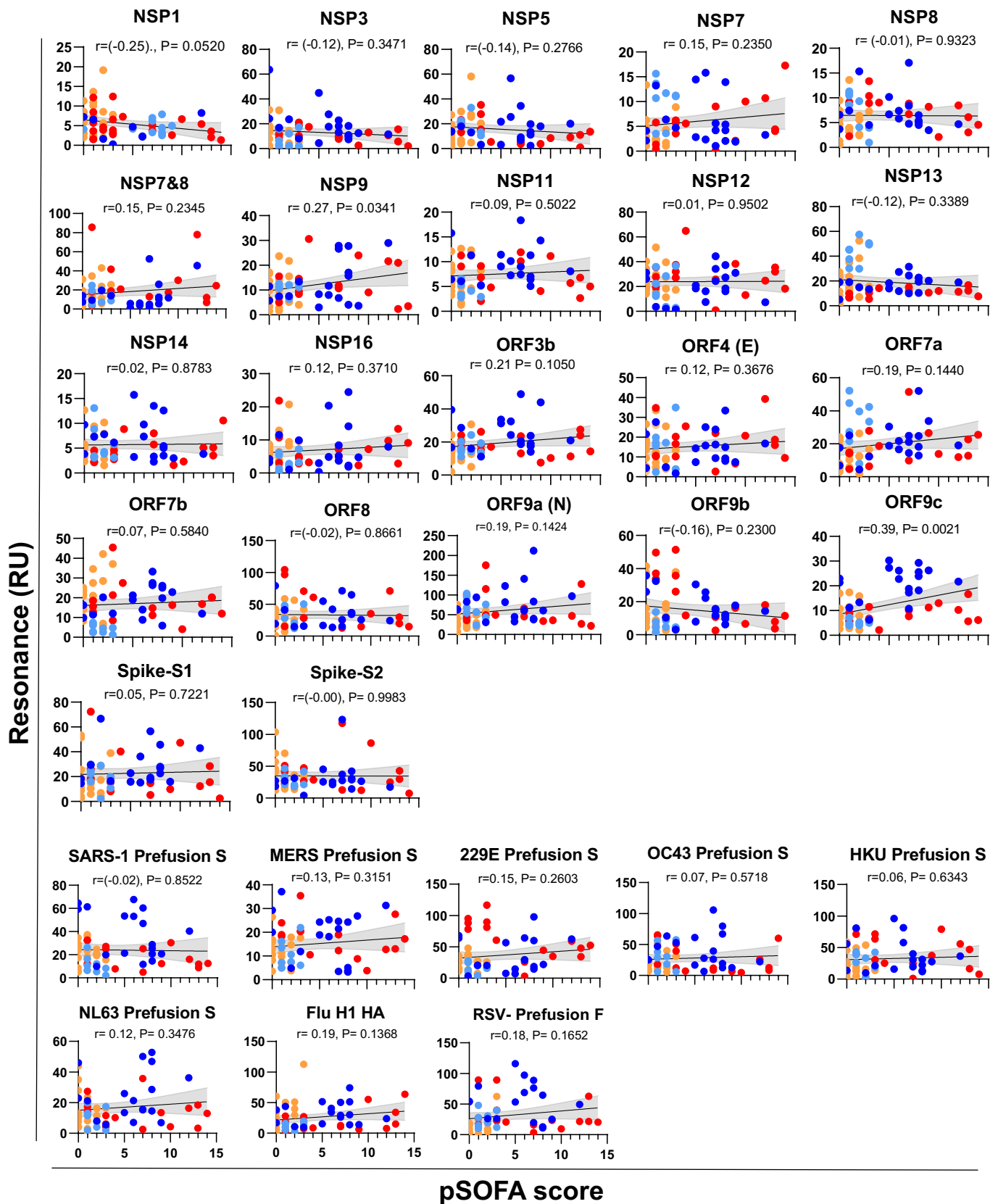**Extended Data Fig. 6 | See next page for caption.**

Extended Data Fig. 6 | SPR based analysis of serum antibodies to purified proteins of SARS-CoV-2 proteome. (a) Serial dilutions of first time-point drawn serum sample from each the controls and SARS-CoV-2 infected children were analyzed for antibody binding to 28 purified SARS-CoV-2 proteins and prefusion spike proteins from other coronaviruses or RSV prefusion-F or H1N1 influenza HA. Rest of the SPR data are shown in Fig. 3. The samples are coded by groups: grey, Mild/moderate Control (n=9); black, Severe control (n=5); orange, Mild/moderate COVID-19 (n=20); red, Severe COVID-19 (n=15); cyan, Mild/moderate MIS-C (n=8); blue, Severe MIS-C (n=17). Total antibody binding shown is observed maximum RU for 10-fold diluted serum sample. **(b)** SPR based analysis of SARS-CoV-2-seropositive children to SARS-CoV-2 prefusion spike, Flu-HA and RSV-F. Serum sample from hospitalized controls (n=7; black) and SARS-CoV-2 infected seropositive children (20 COVID-19, shown in red or 10 MIS-C, shown in blue) were analyzed for antibody binding to purified SARS-CoV-2 prefusion spike protein or RSV prefusion-F or H1N1 influenza HA. Total antibody binding shown is observed maximum RU for 10-fold diluted serum sample. Differences among groups were performed using R. The tests were two-sided using Tukey multiple group comparisons to reduce Type 1 error. The differences were considered statistically significant with a 95% confidence interval when the p value was less than 0.05. (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).
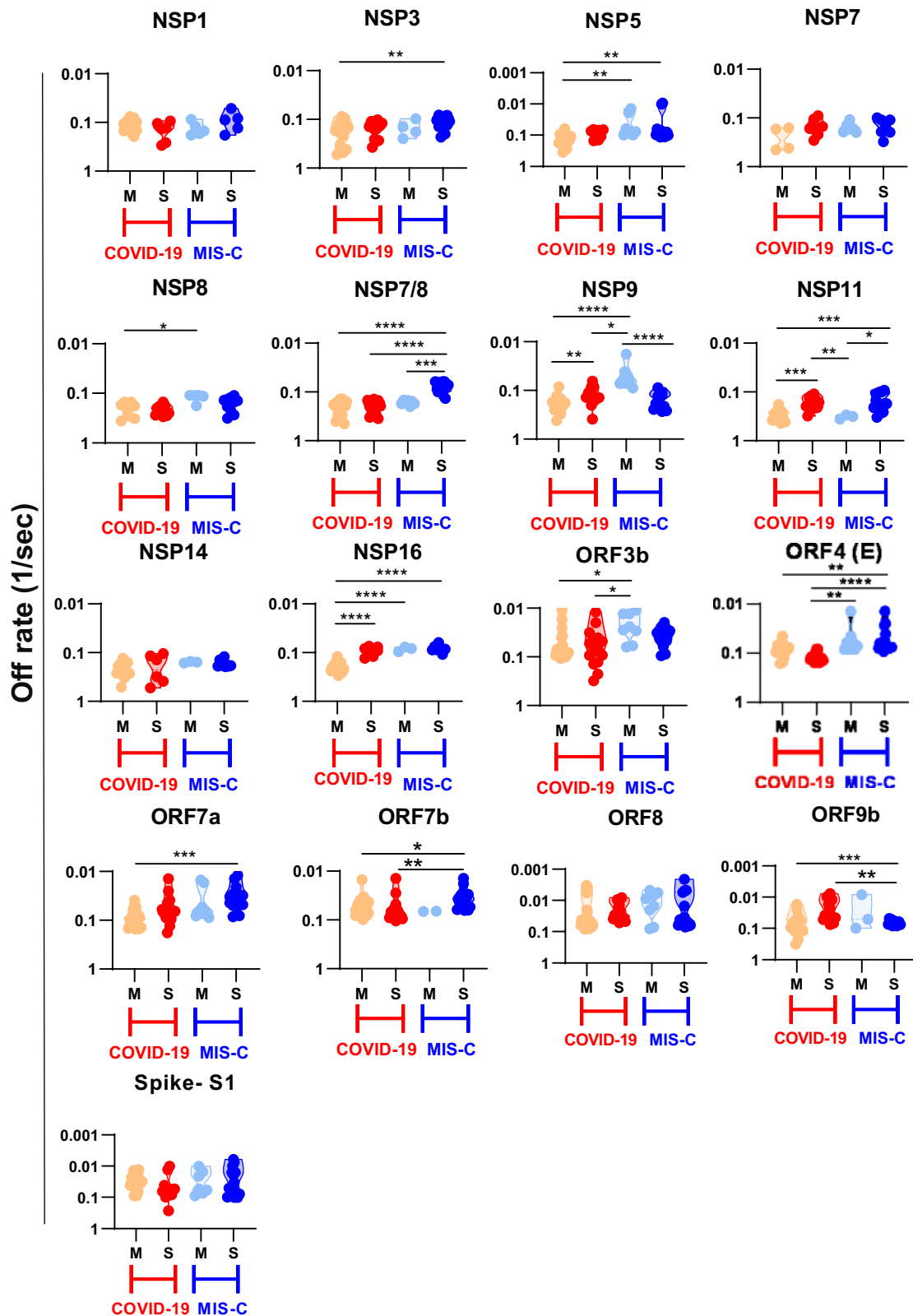
Extended Data Fig. 7 | Serum antibody avidity to purified proteins of SARS-CoV-2 proteome. Serial dilutions of serum sample from each of the SARS-CoV-2 infected children were analyzed for antibody avidity to 28 purified SARS-CoV-2 proteins and prefusion spike proteins from other coronaviruses or RSV prefusion-F or H1N1 influenza HA. Rest of the antibody affinity data is shown in Fig. 3. The samples are color coded by groups: orange, Mild/moderate COVID-19 (n = 20); red, Severe COVID-19 (n = 15); cyan, Mild/moderate MIS-C (n = 8); blue, Severe MIS-C (n = 17). Antibody off-rates as surrogate of antibody affinity was determined against each protein using serially diluted serum sample by SPR in the dissociation phase only for the sensorgrams with Max RU in the range of 5–150 RU. Differences among groups were performed using R. The tests were two-sided using Tukey multiple group comparisons to reduce Type 1 error. The differences were considered statistically significant with a 95% confidence interval when the p value was less than 0.05. (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$).
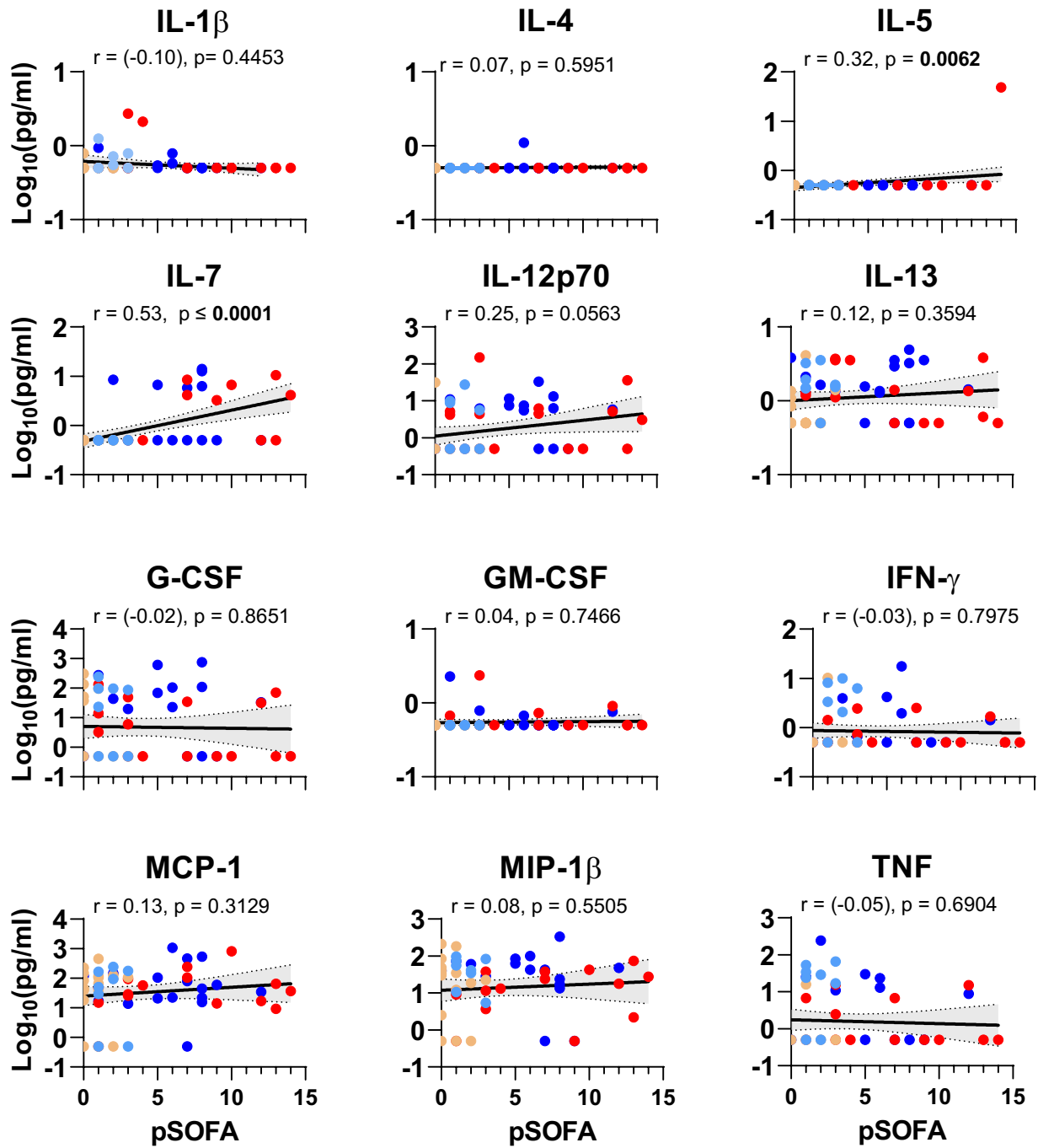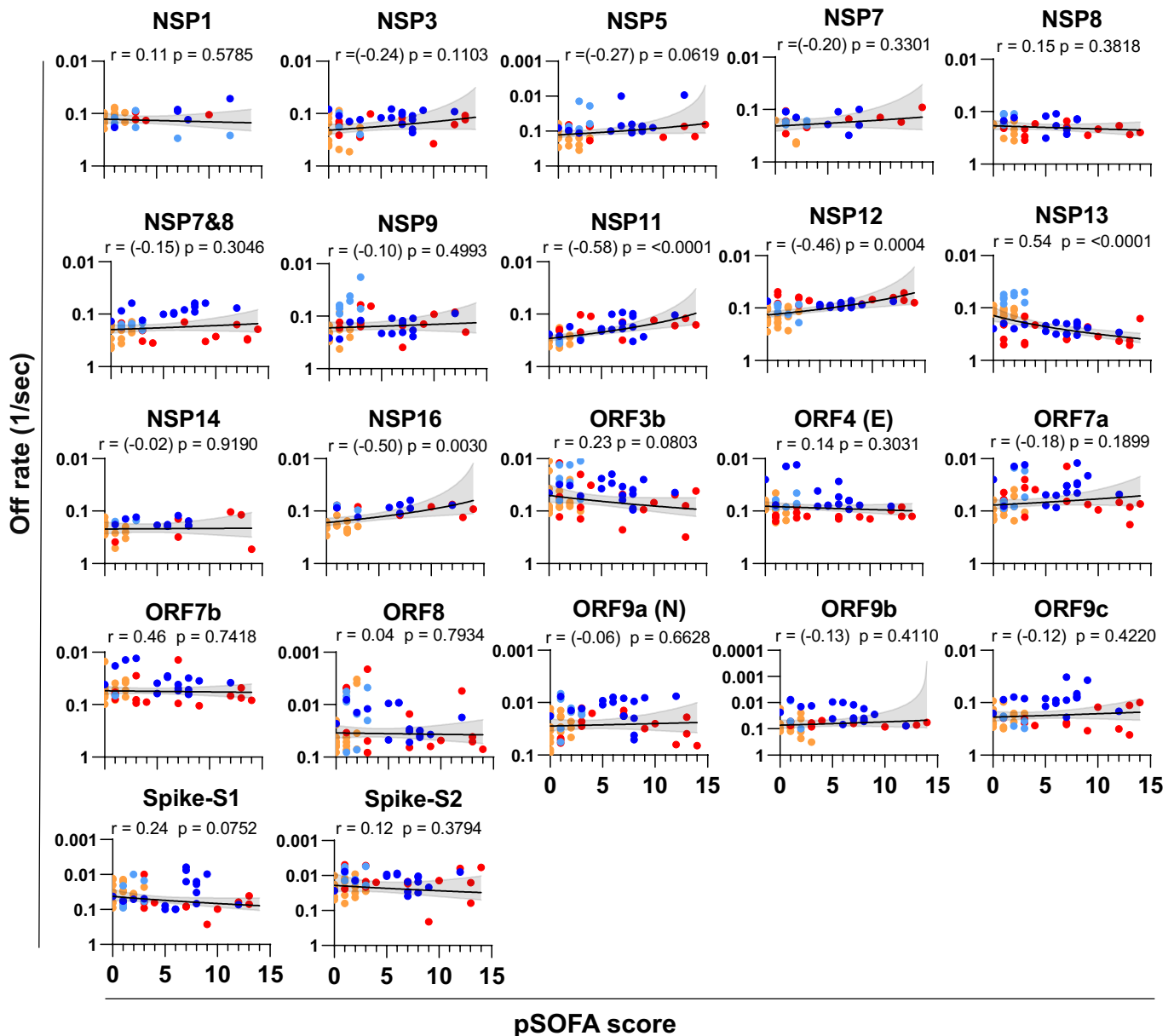
Extended Data Fig. 8 | Relationship of Cytokine/chemokine levels with SARS-CoV-2 clinical disease score. Correlation and linear regression of cytokine/chemokine levels (log10 concentration) for serum sample from each patient were evaluated with corresponding maximum pediatric Sequential Organ Failure Assessment (pSOFA) score for each SARS-CoV-2 infected child on all the days post-enrollment, regardless of disease severity. Each dot represents a unique patient and associated pSOFA score; orange, Mild/moderate COVID-19 (n=20); red, Severe COVID-19 (n=15); cyan, Mild/moderate MIS-C (n=8); blue, Severe MIS-C (n=17). The linear regression were performed across all SARS-CoV-2 infected children. Black line indicates the regression line for all patients. Associated Pearson's correlation coefficients (r) and linear regression significance (p) are shown. The trend line is depicted as solid black line with the error bands representing 95% confidence interval shown as shaded area.

Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Relationship of SARS-CoV-2 serum antibody binding and clinical symptom scores. Total binding antibody (Max RU) of serum sample from SARS-CoV-2 infected children against the 28 purified SARS-CoV-2 protein and prefusion spike proteins from other coronaviruses or RSV prefusion-F or H1N1 influenza HA on all children sera with COVID-19 and MIS-C, regardless of disease severity were plotted with corresponding pSOFA score. Each dot represents a unique patient antibody binding and associated pSOFA score: orange, Mild/moderate COVID-19 (n=20); red, Severe COVID-19 (n=15); cyan, Mild/moderate MIS-C (n=8); blue, Severe MIS-C (n=17). Linear regression was performed across all SARS-CoV-2 infected children, regardless of disease severity. Black line indicates the regression line for all patients. Associated Pearson's correlation coefficients (r) and linear regression significance (p) are shown. The trendline is depicted as solid black line with the error bands representing 95% confidence interval shown as shaded area.

Extended Data Fig. 10 | Relationship of SARS-CoV-2 serum antibody avidity and clinical disease severity. Antibody avidity (off rates; 1/sec) of serum sample from children against the 28 purified SARS-CoV-2 protein and prefusion spike proteins from other coronaviruses or RSV prefusion-F or H1N1 influenza HA on all the days post-enrollment, regardless of disease severity were plotted with corresponding pSOFA score. Each dot represents a unique patient antibody off-rate and associated pSOFA score: Orange, Mild/moderate COVID-19 (n = 20); Red, Severe COVID-19 (n = 15); cyan, Mild/moderate MIS-C (n = 8); Blue, Severe MIS-C (n = 17). The linear regression analysis was performed across all SARS-CoV-2 infected children, regardless of disease severity. Black line indicates the regression line for all patients. Associated Pearson's correlation coefficients (r) and linear regression significance (p) are shown. The trendline is depicted as solid black line with the error bands representing 95% confidence interval shown as shaded area.

Corresponding author(s): Surender Khurana

Last updated by author(s): Sep 1, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BioRad ProteON Manager software (Version 3.1.0) to collect antibody binding data from SPR machine (www.Biorad.com)
Data analysis	BioRad ProteON Manager software (Version 3.1.0) for antibody binding analysis from SPR machine (www.Biorad.com). Statistical analysis were performed using GraphPad prism version 8 (Graph Pad software Inc, San Diego, CA). Sequence alignments were performed using MacVector (MacVector Inc, Apex, NC) with Assembler (v 12.7.5) and CLC Genomics Workbench 20 (v 20.0.4). Structures were visualized and annotated using Chimera version 1.11.2 (https://www.cgl.ucsf.edu/chimera/). Multiple alignment of the sequences was performed using MAFFT v 7 (Madeira et al., Nucleic Acids Research, 2019; Katoh K., Standley D.M., Molecular Biology and Evolution, 2013). Sequence Identity was calculated using BioEdit version 7.2.5 (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Statistical analysis were performed using R and all information is provided the statistical methods section. The R code and data output is available and has been provided to the editors and reviewers.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data are shown in the manuscript figures and supplementary information. The source data are provided in Supplementary tables 1-8. Antigenic sites were

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All available samples were analyzed in this study
Data exclusions	No data was excluded
Replication	All experiments including Neutralization, cytokines, GFPDL and SPR based antibody analysis were performed twice by independent researchers in the lab. All replications were successful. The variation in duplicate experimental runs was <5%.
Randomization	All samples from the COVID-19 patients were analyzed in this study. The study was non-randomized performed during the pandemic on hospitalized COVID-19 patients. Initially, no patient information was provided, and all the immune analyses were conducted blindly. The participants were assigned in each experimental group based on their hospitalization status and clinical outcome.
Blinding	Experiments were performed by different investigators, who were blinded to sample identity.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	HRP-conjugated anti-human IgA + IgG + IgM (Cat No #109-035-064) and HRP-conjugated anti-human IgG-Fc specific antibody (Cat no #709-005-098) were purchased from Jackson Immuno Research.
Validation	These are secondary antibodies, which were characterized by the manufacturer (https://www.jacksonimmuno.com/)

Eukaryotic cell lines

Cell line source(s)	293 and Calu-3 cells were obtained from ATCC. 293-ACE2 cell line was obtained from BEI Resources (NR-52511). FreeStyle293F cells were obtained from ThermoFisher.
Authentication	Cell lines were checked for expression of ACE2 by FACS analysis. None of the cell lines were authenticated by karyotyping or other genomic techniques.
Mycoplasma contamination	Negative for Mycoplasma
Commonly misidentified lines (See ICLAC register)	No misidentified cell lines were used in the study.

Animals and other organisms

Laboratory animals	Female New Zealand white rabbits (4-6 weeks old) were obtained from Charles River labs.
Wild animals	Not applicable
Field-collected samples	Not applicable
Ethics oversight	All animal experiments were approved by the U.S. FDA Institutional Animal Care and Use Committee (IACUC) under Protocol #2008-10. The animal care and use protocol meets National Institutes of Health guidelines.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Population characteristics	Participants in this study were aged 0-21 years old. Individuals of any gender that were hospitalized with COVID-19 disease were eligible for the study.
Recruitment	All patients hospitalized with COVID-19 disease were eligible without any specific selection criteria, without any bias. Samples were collected from patients who provided informed consent and assent as appropriate for age to participate in the study during the pandemic.
Ethics oversight	Children and adolescents, 0 to 21 years of age, hospitalized at Children's Healthcare of Atlanta (CHOA) with confirmed or suspected COVID-19 or MIS-C were enrolled into an IRB-approved specimen collection protocol (Emory University IRB protocols STUDY00022371 and STUDY00000723) following informed consent and assent as appropriate for age. Samples were tested in different antibody assays with approval from the U.S. Food and Drug Administration's Research Involving Human Subjects Committee (FDA-RIHSC) under exemption protocol '252-Determination-CBER-2020-04-02'.

Note that full information on the approval of the study protocol must also be provided in the manuscript.