medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Proteome-wide analysis of differentially-expressed SARS-CoV-2 antibodies in early COVID-19 infection

Xiaomei Zhang[1,3], Xian Wu [2,3], Dan Wang[1,3],Minya Lu[2,3], Xin, Hou [2] ,Hongye Wang[1], Te Liang[1], Jiayu Dai[1], Hu Duan[1], Yingchun Xu[2], Yongzhe Li[2,4], Xiaobo Yu[1,4]

[1] State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences-Beijing (PHOENIX Center), Beijing Institute of Lifeomics, Beijing, 102206, China.

[2] Department of Clinical Laboratory, Peking Union Medical College Hospital, Chinese Academy of Medical Science & Peking Union Medical College, Beijing 100730, China.

[3]These authors contributed equally to this work.

[4]Correspondence to yuxiaobo@mail.ncpsb.org, yongzhelipumch@126.com.

Abstract

Rapid and accurate tests that detect IgM and IgG antibodies to SARS-CoV-2 proteins are essential in slowing the spread of COVID-19 by identifying patients who are infected with COVID-19. Using a SARS-CoV-2 proteome microarray developed in our lab, we comprehensively profiled both IgM and IgG antibodies in forty patients with early-stage COVID-19, influenza, or non-influenza who had similar symptoms. The results revealed that the SARS-CoV-2 N protein is not an ideal biomarker for COVID-19 diagnosis

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

because of its low immunogenicity, thus tests that rely on this marker alone will have a high false negative rate. Our data further suggest that the S protein subunit 1 receptor binding domain (S1-RBD) might be the optimal antigen for IgM antibody detection, while the S protein extracellular domain (S1+S2ECD) would be the optimal antigen for both IgM and IgG antibody detection. Notably, the combination of all IgM and IgG biomarkers can identify 87% and 73.3% COVID-19 patients, respectively. Finally, the COVID-19-specific antibodies are significantly correlated with the clinical indices of viral infection and acute myocardial injury ($p \leq 0.05$). Our data may help understand the function of anti-SARS-CoV-2 antibodies and improve serology tests for rapid COVID-19 screening.

Text

Since December 2019, coronavirus disease 2019 (COVID-19) has become a worldwide pandemic. As of April 14, 2020, COVID-19 has spread to 185 countries with 1,929,922 confirmed cases and 120,450 deaths. COVID-19 symptoms range from a mild cough to pneumonia, and it is estimated that ~17.9% of patients have mild symptoms or no symptoms at all [1-4]. This subgroup of patients, although contagious, would not be selected for viral RNA testing and quarantined. However, diagnostic testing is critical to effective containment of SARS-CoV-2, which is the virus responsible for COVID-19 [5].

SARS-CoV-2 antibodies are produced within the first week of infection [6,7].

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Thus, tests to detect SARS-CoV-2 antibodies in serum, plasma, and whole blood have been developed to rapidly screen patients to ascertain infection status and possibly immunity to COVID-19 [8]. Current rapid antibody tests from different manufacturers rely on antibody binding to SARS-CoV-2 nucleocapsid (N) protein, spike (S) protein, or S protein fragments (i.e., subunit 1, S1; receptor binding domain, RBD). Both N and S proteins are structural proteins. Some manufacturers even use a combination of the N and S proteins and protein fragments, thus resulting in varying or inconsistent antibody tests [8,9]. The SARS-CoV-2 genome also encodes a polyprotein (the open reading frame 1a and 1b, Orf1ab), two additional structural proteins (envelope, E; membrane, M) and five accessary proteins (Orf3a, Orf6, Orf7a, Orf8, Orf10)[10]. The development of a rapid *in vitro* diagnostic serology test with high sensitivity and specificity relies on selecting the antigens with high immunogenicity. Furthermore, it important to consider the homology of the antigen with proteins from other viruses to decrease the number of false negative results [8,11].

In this study, we used a SARS-CoV-2 proteome microarray developed in our laboratory to perform a proteome-wide analysis of differential antibody responses to SARS-CoV-2 proteins in the serum of 40 patients displaying similar symptoms (i.e., fever, cough or muscle ache) with COVID-19, influenza, or non-influenza (Figure 1a, Table 1, Supplementary Figure 1)[12]. The early-stage COVID-19 patients {Onset of symptoms, 4.0 (1.0-20.0) days} were diagnosed according to the Diagnosis and Management Plan of Pneumonia

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

with New Coronavirus Infection (trial version 7). The non-influenza controls are the patients with similar symptoms as COVID-19, but excluded the infection of respiratory virus (Table 1).

The SARS-CoV-2 proteome microarray contained full-length N protein, full-length E protein, five S truncated proteins and 966 tiled peptides representing the SARS-CoV-2 proteome (Supplementary Table 1). Each peptide was 15 amino acids long with a 5 amino acid overlap. The intra- and inter-array correlation of serum IgM and IgG antibody detection were 0.9916 to 0.9992, respectively, demonstrating the high-reproducibility of the SARS-CoV-2 proteome microarray (Supplementary Figure 2). Using the microarray, we analyzed the serum of 15 COVID-19 confirmed patients, 13 influenza patients and 12 non-influenza patients, which generated a total of 77,840 antibody-antigen reactions that were screened.

First, we analyzed the elevated antibodies in early-stage COVID-19 patients compared to the influenza and non-influenza patients (Supplementary Figure 1). A comparison of the COVID-19 and non-influenza patient groups identified 76 IgM antigens and 17 IgG antigens that were differentially targeted (Supplementary Figure 3). The IgM antigens included the S protein (26 peptides, S1, S1RBD, S1+S2ECD, S2ECD), N protein (1 peptide, full-length protein), M protein (2 peptides), E protein (3 peptides, full-length protein), Orf1ab (10 peptides), Orf3a (18 peptides), Orf6 (5 peptides), Orf8 (1 peptide), and Orf10 (3 peptides). The IgG antigens included the S protein (5 peptides,

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

S1, S1RBD, S1+S2ECD, S2ECD), N protein (2 peptides, full-length protein), E protein (1 peptides), and Orf1ab (4 peptides). A comparison of the COVID-19 and influenza groups identified 86 IgM antigens and 64 IgG antigens (Supplementary Figure 4). The IgM antigens included the S protein (67 peptides, S1, S1RBD, S1+S2ECD, S2ECD), N protein (2 peptides), M protein (2 peptides), and Orf1ab (10 peptides). The IgG antigens included the S protein (10 peptides, S1RBD, S1+S2ECD, S2ECD), N protein (2 peptides, full-length protein), and Orf1ab (48 peptides) (Supplementary Table 2).

A Venn diagram analysis indicates that 27 IgM antigens and 9 IgG antigens were unique to COVID-19 compared to both influenza and non-influenza groups (Supplementary Figure 5). All of the IgM antigens were from the S protein (22 peptides, S1, S1RBD, S1+S2ECD, S2ECD). The IgG antigens were from the S protein (2 peptides, S1RBD, S1+S2ECD, S2ECD), N protein (2 peptides, full-length protein) and Orf1ab (1 peptide) (Supplementary Table 2). These results suggest that, during the early stage of COVID-19 infection, IgM antibodies targeting the S protein are produced. The IgG anti-SARS-CoV-2 antibodies, which are produced later during COVID-19 infection, bind to the S, N and Orf1ab proteins. These results make sense because the S protein is displayed across the virion surface, which is easily accessible to the humoral immune system. The N and Orf1ab proteins, on the other hand, is usually compartmentalized within the viral particle. As the virus replicates within the host cells, the N and Orf1ab proteins could be released

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

from the cells and later recognized by the humoral immune system[13].

Hierarchical cluster analyses show that 86.7 % (13/15) and 73.3% (11/15) of COVID-19 patients can be distinguished from influenza and non-influenza patients by measuring these differentially-expressed IgM and IgG antibodies, respectively (Figure 1b). These numbers are superior to the current serological antibody tests using Immunocolloidal Gold lateral flow, ELISA and Chemiluminescence technologies, which have positive rates (%) ranging from 7% to 53% from the same serum samples (Supplementary Figure 6). These results indicate the potential of our SARS-CoV-2 proteome microarray in identifying appropriate antibody biomarkers for early COVID-19 diagnosis.

In order to know which SARS-CoV-2 proteins are suitable for developing a diagnostic test, we compared the IgM and IgG antibodies of COVID-19, influenza, and non-influenza patients in regards to how they targeted two structural proteins, N and S (N, S1, S1RBD, S1+S2ECD and S2ECD). It is surprising that the N protein IgM antibodies cannot discriminate between COVID-19 and the other influenza and non-influenza patients. Instead, IgM antibodies targeting S1-RBD and S1+S2ECD could best discriminate COVID-19 patients from the other groups with a p-value less than 0.01 (COVID-19 vs. influenza) and 0.001 (COVID-19 vs. non-influenza) (Figure 2a). These results were validated using a peptide microarray, in which four antibody binding epitopes on the S protein subunit 2 (S2) were identified (Table 1, Supplementary Table 2). Furthermore, the structural analysis of the

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

trimerized S protein indicate that four epitopes are located within the extracellular domain of subunit 2 (S2ECD, Ser686-Pro1213). These epitopes included 756-920 amino acid residues (YGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILP DPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVL PPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQ NVLYENQ), residues 956-970 (AQALNTLVKQLSSNF), residues 1096-1100 (VSNGT), and residues 996-1000 (LITGR) (Figure 3a).

The concentration of IgG antibodies to N, S, S1+S2ECD, S1-RBD, and S2ECD proteins was significantly higher in COVID-19 patients than the other groups, with S2ECD having the best discriminative performance. The results were also validated using a peptide microarray, in which four peptides from N, S, and Orf1ab proteins were identified as the antibody binding epitopes (Table 2, Supplementary Table 2). Sequential alignment and structural analysis indicate that two epitopes on the S protein (residues 816-820, SFIED; residues 886-890, WTFGA) are located on the surface and inside of the S protein, respectively (Figure 3b). Notably, the one epitope (residues 86-100, YYRRATRRIRGGDGK) on the N protein is located within the RNA binding domain loop that is easily accessible to antibodies. These results suggest that using the appropriate control groups (e.g., non-COVID-19 patients with similar symptoms) is critical in selecting antibody biomarkers for an antibody-based diagnostic test (Supplementary Figures 3 and 4).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

The N protein is not an ideal antigen for the diagnostic test because the IgM antibodies that target this protein could not discriminate between the COVID-19 patients from the influenza patients (Figure 2a). The sensitivity of IgG antibodies for N protein is also low and only 30% (5/15) early COVID-19 patients can be discriminated (Figure 2b). Similar results were also observed in a recent study in which antibodies to the N protein were also detected in some lung cancer patients and healthy people [14]. The S1-RBD and S1+S2ECD proteins might be the optimal antigens for the IgM antibody test, whereas the S2ECD protein might be the optimal antigen for the IgG antibody test (Figure 2). Furthermore, the combination of different biomarkers could increase the test's sensitivity compared to using a single biomarker alone by distinguishing 86.7 % (13/15) and 73.3% (11/15) COVID-19 patients from the influenza and non-influenza patients (Figure 1b). Identification of multiple biomarkers or peptide epitopes with high immunogenicity is possible with our SARS-CoV-2 proteome array.

Finally, we performed a comprehensive circos correlation analysis of COVID-19 specific SARS-CoV-2 antibodies and 20 clinical indices. Statistically significant correlations ($p \leq 0.05$) are indicated in pink (Figure 4, Supplementary Table 3). The IgM antibodies to the S protein significantly correlated with eosinophil count (EOS#), hemoglobin (HGB), platelet count (PLT), albumin (Alb), creatine kinase MB mass (CKMB-mass) and N-terminal pro-brain natriuretic peptide (NT-proBNP) levels. On the other hand, the IgG antibodies

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

to the S protein significantly correlated with lymphocyte count and percentage (LY# and LY%), CKMB-mass and NT-proBNP levels. Notably, the levels of two biomarkers of cardiac injury, NT-proBNP and CKMB-mass, were significantly correlated with both IgM and IgG COVID-19 specific antibodies (Figure 4), suggesting that anti-SARS-CoV-2 antibodies contribute to acute myocardial injury[15].

A limitation of this study is that antibodies binding to conformational epitopes or post translational modifications could not be detected using our microarray comprised of chemically-synthesized, linear peptides [16]. In addition, a small number of patients were employed. These results should be validated using a large patient cohort with other respiratory infections[15].

Our work provides valuable information to help fight COVID-19. Understanding the humoral immune response to COVID-19 will help us understand the role of antibodies in COVID-19, identify potential targets for immunotherapy, and enable the production of accurate rapid serology tests for COVID-19 detection. Detection of IgM antibodies will help identify those with acute COVID-19 infection who need to be quarantined. Detection of IgG antibodies may help identify those who are recovering from or have gained immunity to COVID-19.

References

1. Mizumoto. K, Kagaya. K, Zarebski; A, Chowell. G. Estimating the Asymptomatic Proportion of 2019 Novel Coronavirus onboard the Princess Cruises Ship, 2020 *medRxiv* 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

2. Mahase E. China coronavirus: mild but infectious cases may make it hard to control outbreak, report warns. *BMJ* 2020; **368**: m325.

3. Phan LT, Nguyen TV, Luong QC, et al. Importation and Human-to-Human Transmission of a Novel Coronavirus in Vietnam. *N Engl J Med* 2020.

4. Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med* 2020.

5. Sharfstein JM, Becker SJ, Mello MM. Diagnostic Testing for the Novel Coronavirus. *JAMA* 2020.

6. Yu X, Bian X, Throop A, et al. Exploration of panviral proteome: high-throughput cloning and functional implications in virus-host interactions. *Theranostics* 2014; **4**(8): 808-22.

7. Zhu H, Hu S, Jona G, et al. Severe acute respiratory syndrome diagnostics using a coronavirus protein microarray. *Proc Natl Acad Sci U S A* 2006; **103**(11): 4011-6.

8. OKBA. N, Muller. M, Li. W, et al. SARS-CoV-2 specific antibody responses in COVID-19 patients. *medRxiv* 2020.

9. Long. Q, Deng. H, Chen. J, et al. Antibody responses to SARS-CoV-2 in COVID-19 patients: the perspective application of serological tests in clinical practice. *medRxiv* 2020.

10. Narayanan K, Huang C, Makino S. SARS coronavirus accessory proteins. *Virus Res* 2008; **133**(1): 113-21.

11. Evidence for network evolution in an Arabidopsis interactome map. *Science* 2011; **333**(6042): 601-7.

12. Wang. H, Hou. X, Wu. X, et al. SARS-CoV-2 proteome microarray for mapping COVID-19 antibody interactions at amino acid resolution. *BioRxiv* 2020.

13. Jiang S, Hillyer C, Du L. Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends Immunol* 2020.

14. Market Trends for Biomarker-Based IVT Tests (2003-2014). Amplion Inc. (www.amplion.com) 2015.

15. Bonow RO, Fonarow GC, O'Gara PT, Yancy CW. Association of Coronavirus Disease 2019 (COVID-19) With Myocardial Injury and Mortality. *JAMA Cardiol* 2020.

16. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 2020.

17. Cheadle C, Vawter MP, Freed WJ, Becker KG. Analysis of microarray data using Z score transformation. *J Mol Diagn* 2003; **5**(2): 73-81.

Contributions

X. H., Y. L. X. W. provided the clinical samples. X. Z., H. W., D. W., J. D. prepared the microarrays. X. H., X. W. and X. Z. executed microarray experiments. D. W., X. Z., T. L. and X.Y. executed the statistical and structural analysis. X.Y., and Y. L. conceived the idea, designed experiments, analyzed the data and wrote the manuscript.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Acknowledgement

This work was supported by the State Key Laboratory of Proteomics (SKLP-C202001，SKLP-O201703 and SKLP-K201505), the Beijing Municipal Education Commission, National Natural Science Foundation of China (81671618，81871302, 81673040, 31870823), the National Program on Key Basic Research Project (2018YFA0507503, 2017YFC0906703 and 2018ZX09733003) and the CAMS Initiative for Innovative Medicine (2017-I2M-3-001 and 2017-I2M-B&R-01). We also thank Dr. Brianne Petritis for her critical review and editing of this manuscript.

Competing interests

None declared.

Supplementary information

The supplementary information includes the materials and methods, 3 supplementary tables and 6 supplementary figures.

Tables

Table 1. Clinical characteristics of early COVID-19 and suspected patients displaying similar symptoms

	Non-influenza patients	Influenza patients (n=13)	Covid-19 patients (n=15)

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

	(n=12)		
Demographics and clinical characteristics			
male	3 (25%)	8 (61.5%)	7 (46.7%)
female	9 (75%)	5 (38.5%)	8 (53.3%)
Age, year	36.5 (32.0-57.0)	51.0 (20.0-88.0)	38.0 (7.0-68.0)
2019-nCoV(+)	0	0	15 (100%)
FluA-RNA(+)	0	6 (46.2%)	0
FluB RNA(+)	0	3 (23%)	0
RSV RNA(+)	0	4 (30.8%)	0
Exposure history	2 (16.7%)	4 (30.8%)	14 (93.3%)
Fever	12 (100%)	13 (100%)	11 (73.3%)
Headache	3 (25%)	4 (30.8%)	1 (6.7%)
Cough	4 (33.3%)	6 (46.2%)	7 (46.7%)
Sputum	3 (25%)	3 (23%)	2 (13.3%)
Myalgia	1 (8.3%)	7 (53.8%)	2 (13.3%)
Fatigue	1 (8.3%)	6 (46.2%)	0
Diarrhea	0	1 (7.7%)	0
Dyspnea	0	1 (7.7%)	2 (13.3%)
Nausea or vomiting	0	0	0
Onset of symptoms, days	1.0 (0-5.0)	2.0 (1.0-22.0)	4.0 (1.0-20.0)
Imaging features			
Ground-glass opacity	8 (66.7%)	7(53.8%)	12 (80%)
Bilateral pulmonary infiltration	4 (33.3%)	6(46.2%)	10 (66.7%)

Data are median (IQR) or n (%).

Table 2. Binding epitopes of anti-SARS-CoV-2 antibodies

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

	IgM	IgG
S	756-YGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPI KDFGGFNFSQILPDPSKPSKR SFIEDLLFNKVTLADAGFIKQY GDCLGDIAARDLICAQKFNGL TVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQ MAYRFNGIGVTQNVLYENQ-9 20	816-SFIED-820
	956-AQALNTLVKQLSSNF-970	886-WTFGA-890
	986-KVEAE-990	
	996- LITGR-1000	
	1096- VSNGT-1100	
N		86-YYRRATRRIRGGDGK-110
Orf1ab		3366-FVRIQ-3370

Figure and legends

Figure 1. Proteome-wide analysis of differential antibody response to SARS-CoV-2 proteins using a SARS-CoV-2 proteome microarray. (a) The workflow of proteome-wide analysis of differential antibody response to SARS-CoV-2 proteins using a proteome microarray. (b) Hierarchical clustering analysis of differential antibody responses to SARS-CoV-2 proteome between patients with COVID-19, influenza, and non-influenza displaying similar symptoms. The false-colored rainbow color from blue to red correspond to low to high signal intensities on the array, respectively.

Figure 2. Box plot analysis of antibody responses to SARS-CoV-2 structural N and S proteins between confirmed and suspected COVID-19 patients. The statistical significance was calculated using Mann Whitney *U*-test with a p-value less than 0.05. The one, two and three asterisks

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

represent the p-value of 0.05, 0.01 and 0.001, respectively.

Figure 3. Structural analysis of immunogenic epitopes of SARS-CoV-2 proteins. Structural analyses of the (a) nucleocapsid protein's RNA binding domain (PDB ID: 6VYO) and (b) spike protein (PDB ID: 6VXX). The epitope is labeled in yellow and indicated with a red arrow.

Figure 4. Correlation network of COVID-19 specific antibodies and clinical indices. (A) IgM and (B) IgG antibody correlations between COVID-19 specific antibodies and clinical index using circos. Correlation with statistical significance ($p \leq 0.05$) are indicated in pink. Positive and negative correlations with non-significance ($p \geq 0.05$) are indicated in red and gray, respectively. (c) and (d) are the COVID-19 antibodies correlated with NT-proBNP and CKMBmass, respectively. The COVID-19 antibodies were selected with a Spearman coefficient correlation higher than 0.6 and a p-value less than 0.05.

Methods

Collection of clinical samples

Patients with COVID-19, influenza or non-influenza with similar symptoms were enrolled in the Outpatient department of Peking Union Medical College Hospital. All serum samples were collected under the approval of the intuitional review board (IRB) from Peking Union Medical College Hospital (Ethical number: ZS-2303) and Beijing Proteome Research Center. Written informed consent was waived due to the rapid emergence of this infectious disease. All

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

experiments were performed according to the standards of the Declaration of Helsinki. The COVID-19 patients with mild symptoms were diagnosed according to the Diagnosis and Management Plan of Pneumonia with New Coronavirus Infection (trial version 7).

Screening of serological antibodies in COVID-19 and suspected patients using SARS-CoV-2 proteome microarray

The SARS-CoV-2 proteome microarrays containing 966 peptides and full-length N, full-length E and truncated S proteins were prepared as described previously[12]. Prior to antibody detection, the microarrays were assembled in an incubation tray and blocked with 5% (w/v) milk in 1xPBS with 0.05% (v/v) Tween-20 (PBST) for 1 min at room temperature. After washing with PBST three times, the array was incubated with 1:100 diluted serum for 30 min at room temperature and washed again. The array was then incubated for 30 min with a mixture containing Cy3 Affinipure donkey anti-human IgG(H+L) and Alexa Fluor 647 Affinipure goat anti-human IgM FC5µ antibody (Jackson ImmunoResearch, USA) (2µg/mL). Finally, the resulting array was washed with PBST and water, dissembled from the tray and dried with centrifugation for 2 min at 2,000 rpm. The slide was scanned with a GenePix 4300A microarray scanner (Molecular Devices, Sunnyvale, CA, USA) at 532 and 635 nm to measure Cy3 and Alexa Fluor 647 fluorescence, respectively.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20064535. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

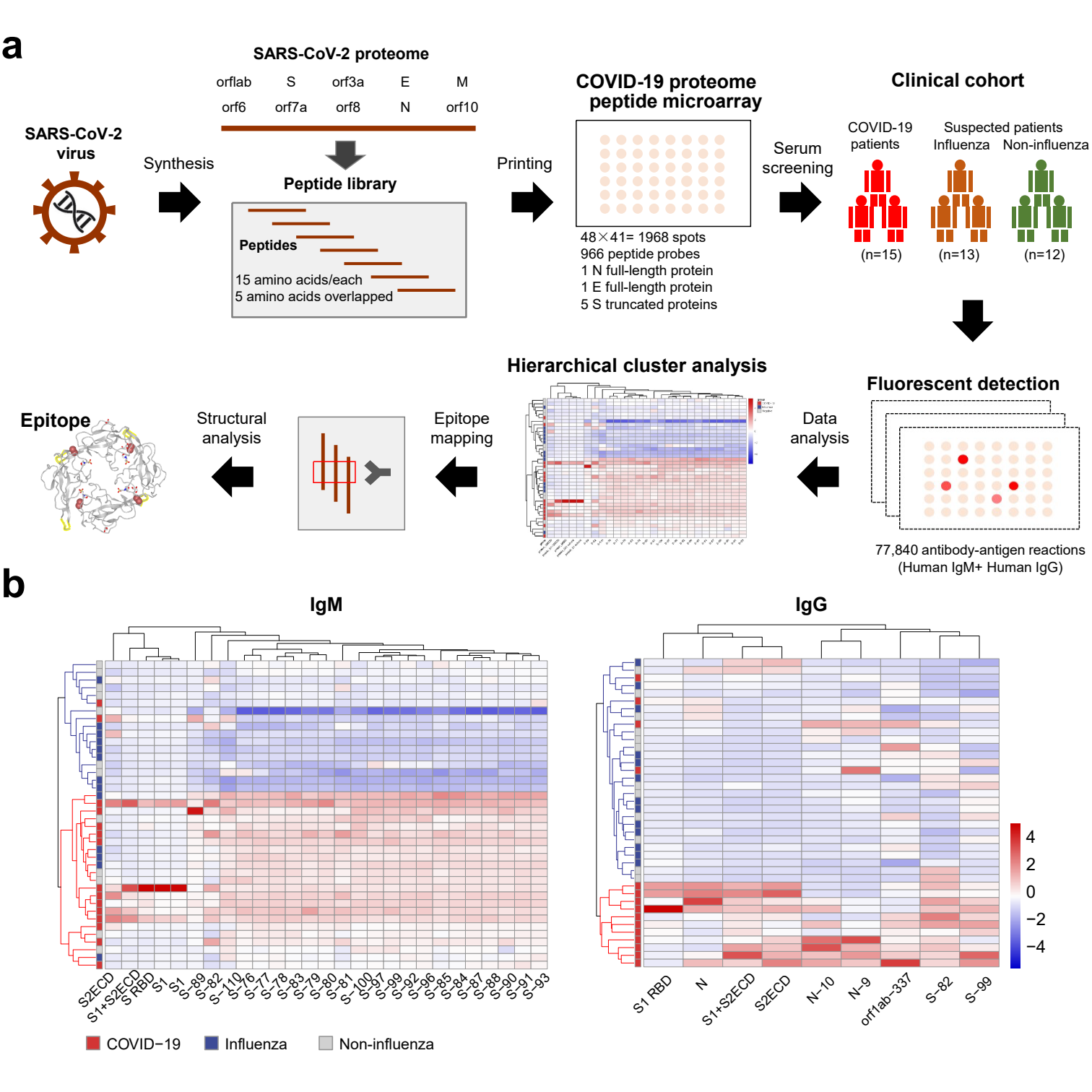
Median spot intensity minus background was extracted using GenePix Pro7 software (Molecular Devices, Sunnyvale, CA, USA).

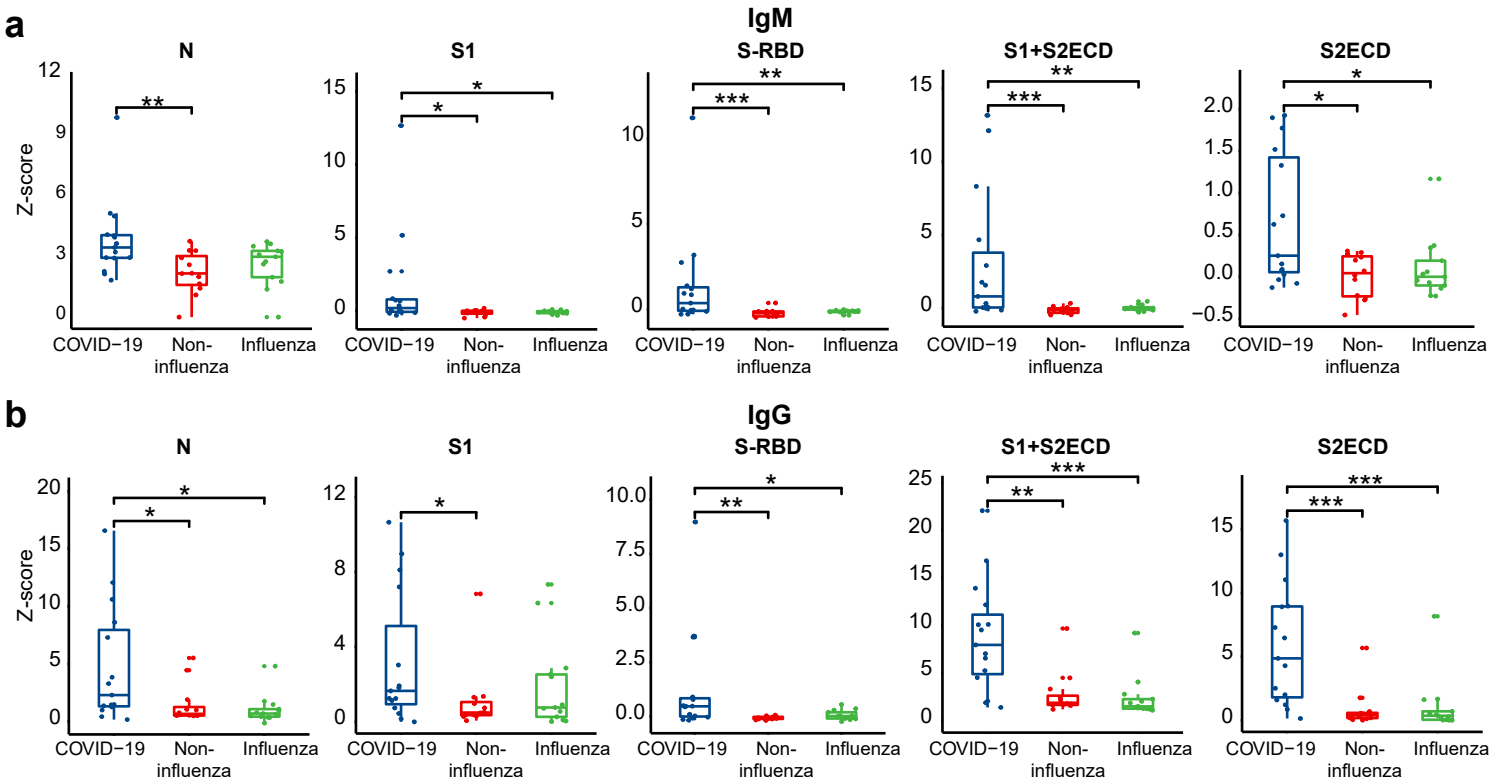
Detection of serological antibody in early COVID-19 patients using Immunocolloidal Gold Strip, ELISA and Chemiluminescence.

Serological antibodies in patients with COVID-19, influenza and non-influenza suspected patients were analyzed using Immunocolloidal Gold Strip, ELISA and Chemiluminescence according to the manufacturers' instructions.

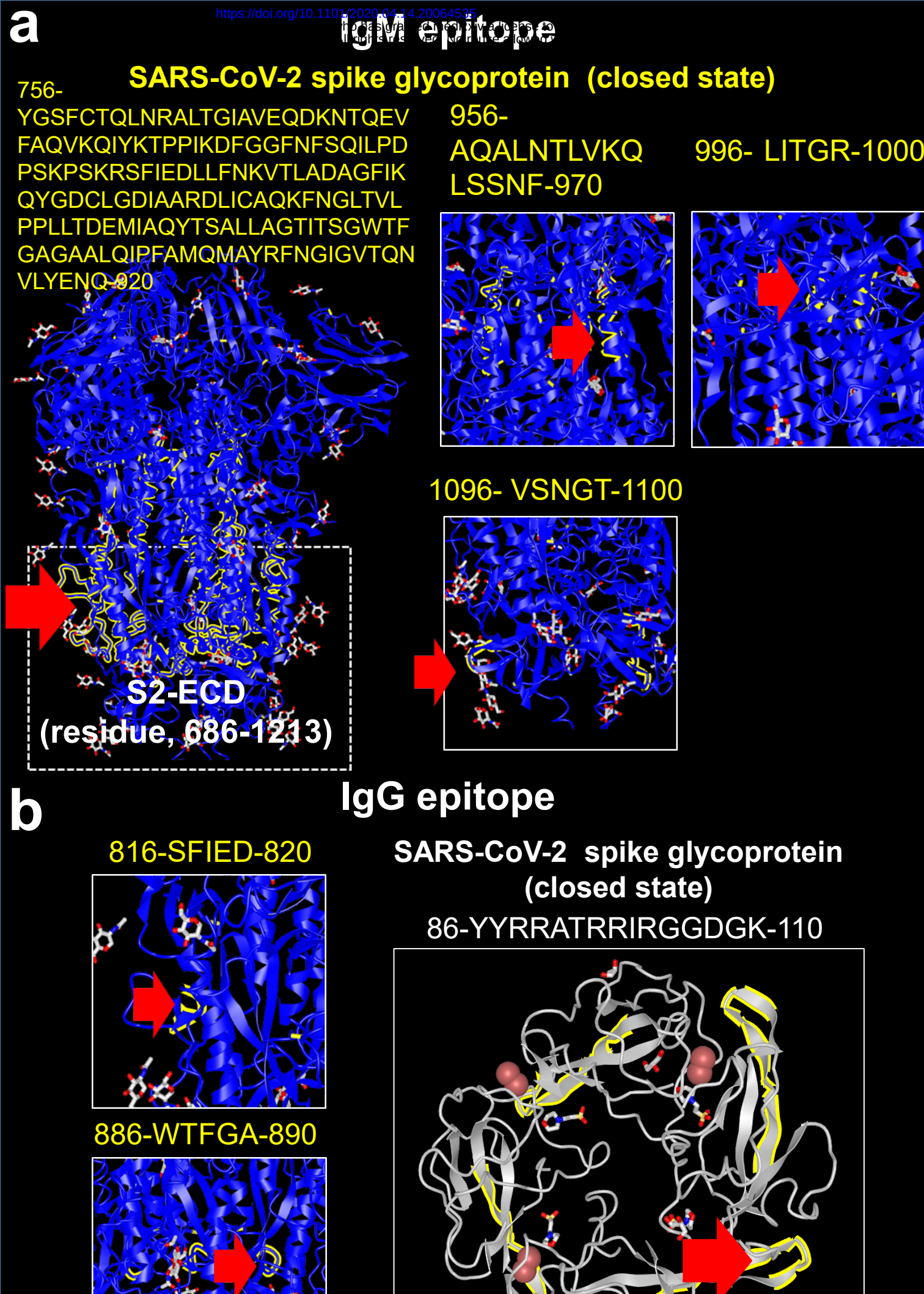
Statistical analysis

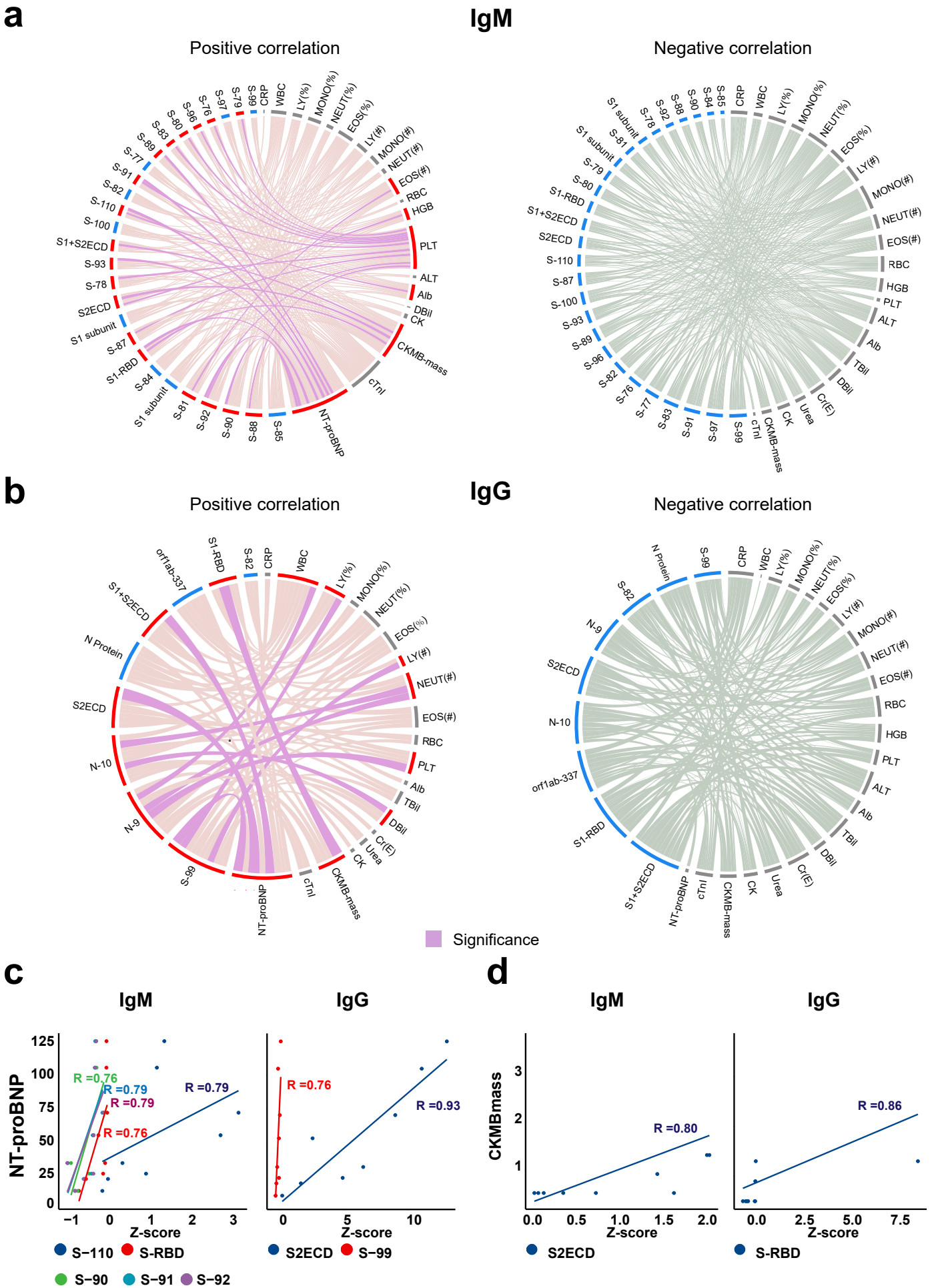
All microarray signals were normalized with a Z-score prior to statistical analyses[17]. Differentially-expressed SARS-CoV-2 antibodies were identified using Mann Whitney *U*-test with a p-value of 0.05. The hierarchical cluster analysis of serological antibody response to peptides was performed using the R pheatmap. The circos plot was made using circos (http://circos.ca/).


a
SARS-CoV-2 proteome
orflab S orf3a E M
orf6 orf7a orf8 N orf10
SARS-CoV-2 virus
Synthesis
Peptide library
Peptides
15 amino acids/each
5 amino acids overlapped
Printing
COVID-19 proteome peptide microarray
48×41= 1968 spots
966 peptide probes
1 N full-length protein
1 E full-length protein
5 S truncated proteins
Serum screening
Clinical cohort
COVID-19 patients
Suspected patients
Influenza
Non-influenza
(n=15)
(n=13)
(n=12)
Fluorescent detection
77,840 antibody-antigen reactions
(Human IgM+ Human IgG)
Data analysis
Hierarchical cluster analysis
Epitope mapping
Structural analysis
Epitope
b
IgM
IgG
S2ECD
S1+S2ECD
S RBD
S1
S1
S-89
S-82
S-110
S-76
S-77
S-78
S-83
S-79
S-80
S-81
S-100
S-97
S-99
S-92
S-96
S-85
S-84
S-87
S-88
S-90
S-91
S-93
S1 RBD
N
S1+S2ECD
S2ECD
N-10
N-9
orf1ab-337
S-82
S-99
4
2
0
−2
−4
COVID−19
Influenza
Non-influenza


a
IgM
N
S1
S-RBD
S1+S2ECD
S2ECD
Z-score
COVID−19
Non-influenza
Influenza
b
IgG
N
S1
S-RBD
S1+S2ECD
S2ECD
Z-score
COVID−19
Non-influenza
Influenza

https://doi.org/10.1101/2020.04.14.20064535




a
IgM
Positive correlation
Negative correlation
b
IgG
Positive correlation
Negative correlation
Significance
c
IgM
IgG
NT-proBNP
Z-score
R =0.76
R =0.79
R =0.79
R =0.76
R =0.79
R =0.76
R =0.93
S−110
S-RBD
S−90
S−91
S−92
S2ECD
S−99
d
IgM
IgG
CKMBmass
Z-score
R =0.80
R =0.86
S2ECD
S-RBD

- Supplementary Information-

Proteome-wide analysis of differentially-expressed SARS-CoV-2 antibodies in early COVID-19 infection

Xiaomei Zhang[1,3], Xian Wu [2,3], Dan Wang[1,3],Minya Lu[2,3], Xin, Hou [2] ,Hongye Wang[1], Te Liang[1], Jiayu Dai[1], Hu Duan[1], Yingchun Xu[2], Yongzhe Li[2,4], Xiaobo Yu[1,4]

[1] State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences-Beijing (PHOENIX Center), Beijing Institute of Lifeomics, Beijing, 102206, China.

[2] Department of Clinical Laboratory, Peking Union Medical College Hospital, Chinese Academy of Medical Science & Peking Union Medical College, Beijing 100730, China.

[3]These authors contributed equally to this work.

[4]Correspondence to yuxiaobo@mail.ncpsb.org, yongzhelipumch@126.com.

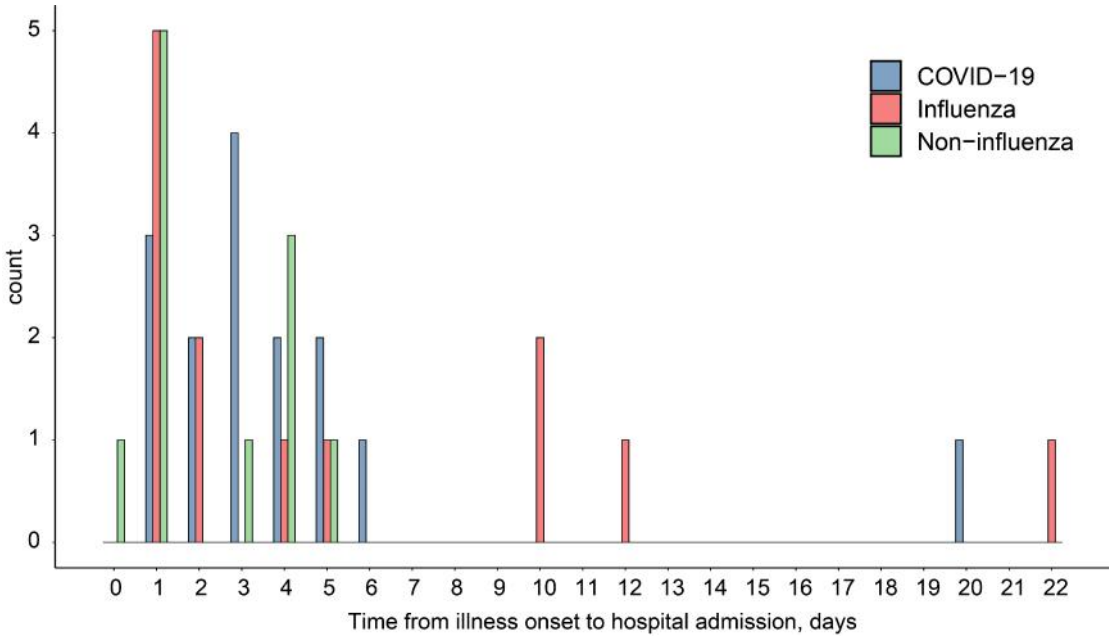
Supplementary Tables

Supplementary Table 1. The layout of the SARS-CoV-2 proteome microarray.

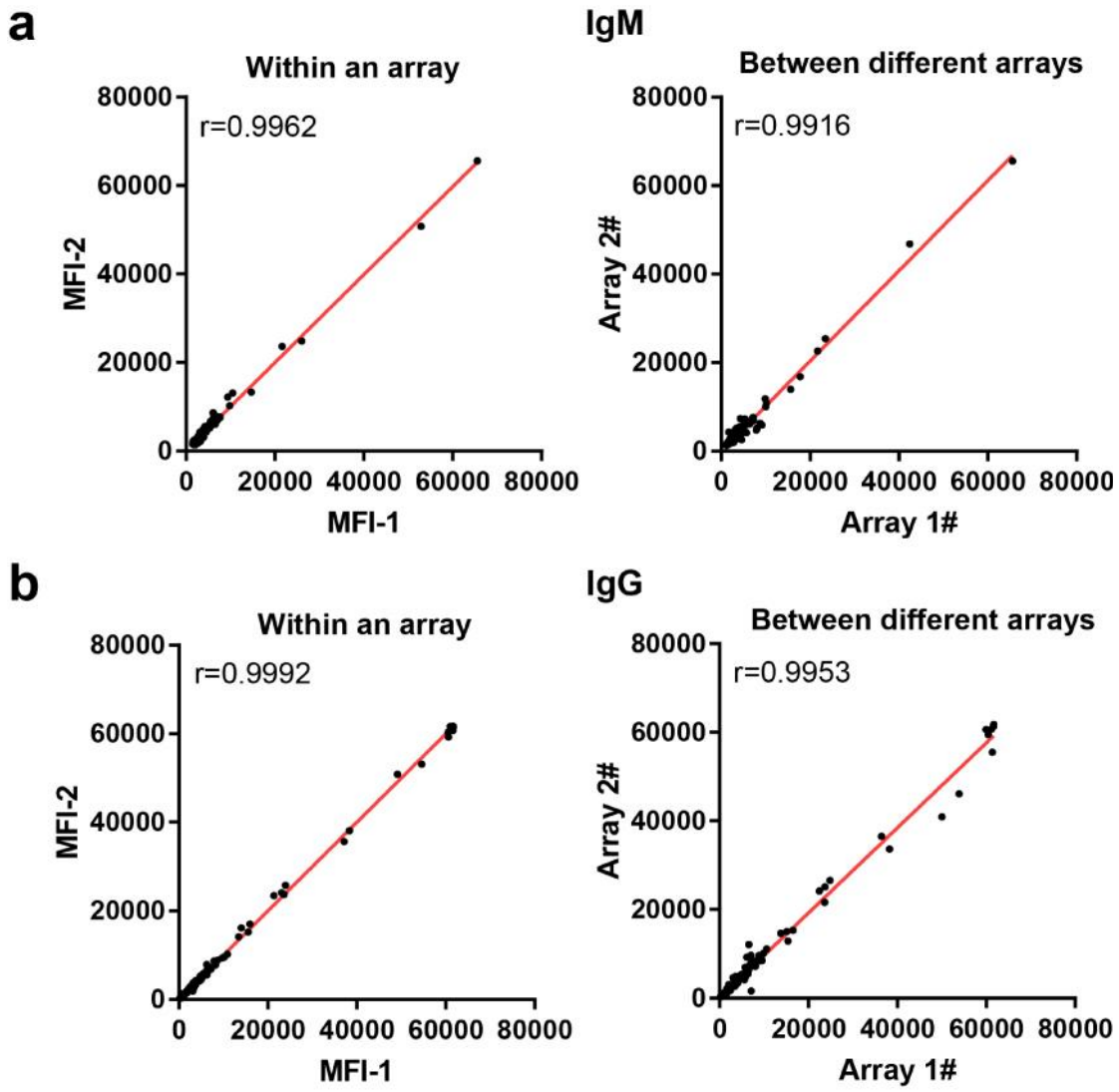
Supplementary Table 2. Differentially-expressed antibodies between COVID-19, influenza and non-influenza patients.

Supplementary Table 3. Correlation between COVID-19 specific SARS-CoV-2 antibodies and clinical indices.

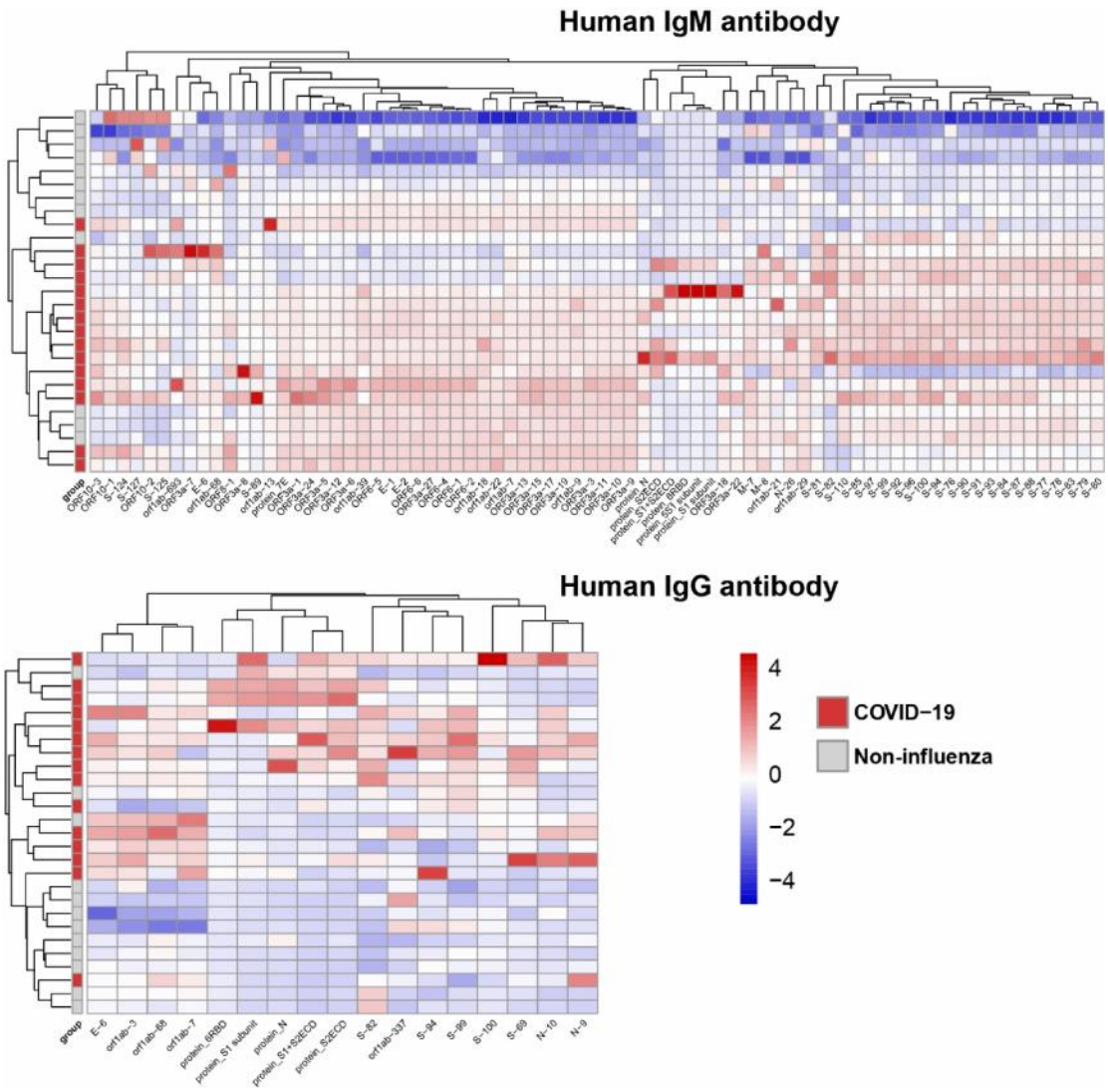
Supplementary figures



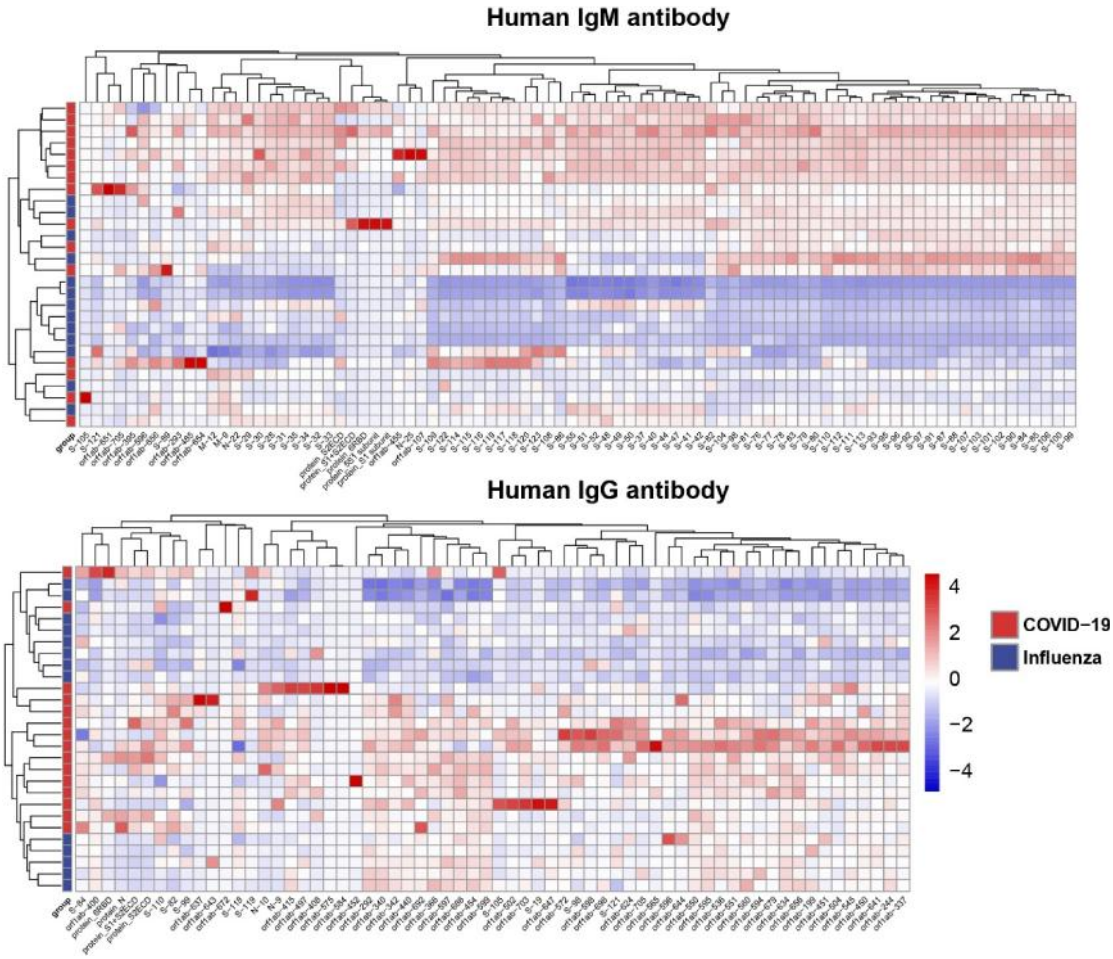
Supplementary Figure 1. Histogram analysis of the time from illness onset to hospital admission in COVID-19, influenza and non-influenza patients.



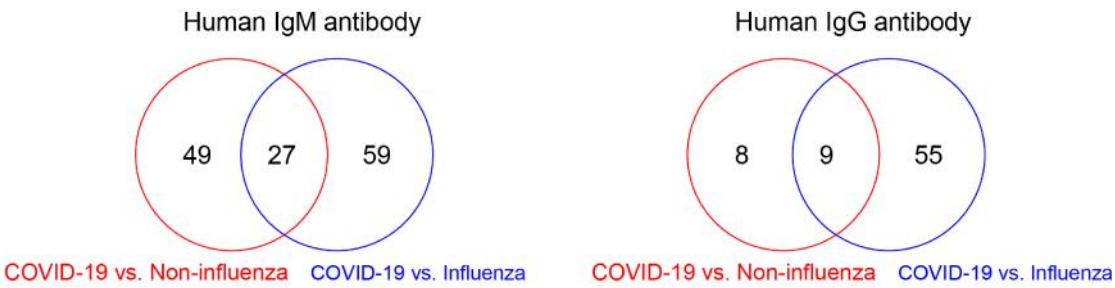
Supplementary Figure 2. Reproducibility of serological antibody detection using the SARS-CoV-2 proteome microarray. (a) and (b) are the correlation of serological IgM and IgG antibody detection within an array and between different arrays, respectively.



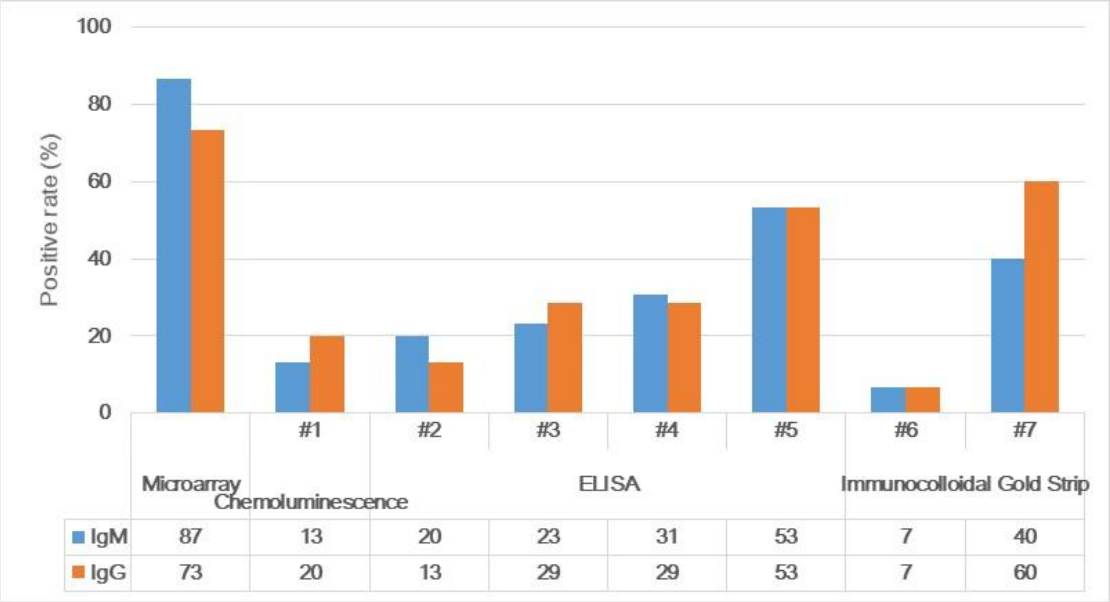
Supplementary Figure 3. Comparison of differential antibody responses to the SARS-CoV-2 proteome between the COVID-19 and non-influenza groups. The false-colored rainbow color from blue to red corresponds to the Z-score from low to high, respectively.



Supplementary Figure 4. Comparison of differential antibody responses to the SARS-CoV-2 proteome between the COVID-19 and influenza groups. The false-colored rainbow color from blue to red corresponds to the Z-score from low to high, respectively.



Supplementary Figure 5. Venn diagram analysis of antigens identified in the COVID-19, influenza and non-influenza patient groups.



Supplementary Figure 6. Comparison of the positive rate (%) for the detection of early COVID-19 patients using different methods. Positive rate (%) refers to the test accuracy, or the % of the time in which an accurate diagnosis is made. #1-#7 are the serum antibody tests from different manufacturers.

.