medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

SARS-CoV-2 antibody signatures for predicting the outcome of COVID-19

Qing Lei[1,+], Cai-zheng Yu[2,+], Yang Li[3,+], Hong-yan Hou[4], Zhao-wei Xu[3], Zong-jie Yao[1], Yan-di Zhang[1], Dan-yun Lai[3], Jo-Lewis Banga Ndzouboukou[1], Bo Zhang[4], Hong Chen[3], Zhu-qing Ouyang[1], Jun-biao Xue[3], Xiao-song Lin[1], Yun-xiao Zheng[3], Xue-ning Wang[3], He-wei Jiang[3], Hai-nan Zhang[3], Huan Qi[3], Shu-juan Guo[3], Mei-an He[5], Zi-yong Sun[4], Feng Wang[4,*], Sheng-ce Tao[3,*], Xiong-lin Fan[1,*]

[1]Department of Pathogen Biology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[2]Department of Public Health, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[3]Shanghai Center for Systems Biomedicine, Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Jiao Tong University, Shanghai 200240, China

[4]Department of Laboratory Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.

[5]Department of Occupational and Environmental Health and State Key Laboratory of Environmental and Health for Incubating, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.

[+] These authors contributed equally to this paper.

[*]Corresponding authors: Xiong-lin Fan (xlfan@hust.edu.cn), Sheng-ce Tao (taosc@sjtu.edu.cn), Feng Wang (wangfeng@tjh.tjmu.edu.cn).

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Abstract

The COIVD-19 global pandemic is far from ending. There is an urgent need to identify applicable biomarkers for predicting the outcome of COVID-19. Growing evidences have revealed that SARS-CoV-2 specific antibodies remain elevated with disease progression and severity in COIVD-19 patients. We assumed that antibodies may serve as biomarkers for predicting disease outcome. By taking advantage of a newly developed SARS-CoV-2 proteome microarray, we surveyed IgM/IgG responses against 20 SARS-CoV-2 proteins in 1,034 hospitalized COVID-19 patients on admission, who were followed till 66 days. The microarray results were correlated with clinical information, laboratory test results and patient outcomes. Cox proportional hazards model was used to explore the association between SARS-CoV-2 specific antibodies and COVID-19 mortality. We found that high level of IgM against ORF7b at the time of hospitalization is an independent predictor of patient survival (p trend = 0.002), while levels of IgG responses to 6 non-structural proteins and 1 accessory protein, i. e., NSP4, NSP7, NSP9, NSP10, RdRp (NSP12), NSP14, and ORF3b, possess significant predictive power for patient death, even after further adjustments for demographics, comorbidities, and common laboratory markers for disease severity (all with p trend < 0.05). Spline regression analysis indicated that the correlation between ORF7b IgM, NSP9 IgG, and NSP10 IgG and risk of COVID-19 mortality is linear (p = 0.0013, 0.0073 and 0.0003, respectively). Their AUCs for predictions, determined by computational cross-validations (validation1), were 0.74 (cut-off = 7.59), 0.66 (cut-off = 9.13), and 0.68 (cut-off = 6.29), respectively. Further validations were conducted in the second and third serial samples of these cases (validation2A, n = 633, validation2B, n = 382), with high accuracy of prediction for outcome. These findings have important implications for improving clinical management, and especially for developing medical interventions and vaccines.

Keywords

COVID-19, SARS-CoV-2, predicting signature, outcome

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Highlights

1. IgM/IgG responses of 1,034 patients upon admission against 20 SARS-CoV-2 proteins were analyzed.
2. High level of IgM against ORF7b at the time of hospitalization is an independent predictor of patient survival.
3. IgG responses to NSP9 and NSP10 possess significant predictive power for patient death.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Introduction

The coronavirus disease 2019 (COVID-19), the emerging infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in December 2019, has quickly become the greatest crisis of global public health and economic development in our times [1]. As of October 11, 2020, there has been 36.75 million confirmed cases and 1.06 million patients have died from SARS-CoV-2 infection worldwide [2]. The crude mortality rate of COVID-19 was approximately 2.9% as estimated by the WHO [2], and there are no highly effective therapeutics or vaccines currently available for COVID-19 [3-5], highlighting the significance to understand the role of immunity in the progression and outcome of COVID-19 patients for improving clinical management and developing effective interventions and vaccines.

Similar to SARS-CoV and MERS-CoV, SARS-CoV-2 belongs to the betacoronavirus genus and its genome encodes four major structural proteins spike (S), envelope (E), membrane (M), and nucleocapsid (N), and 15 non-structural proteins (Nsp1-10 and Nsp12-16) and 9 accessory proteins [6]. Among these, the S protein, consisted of N-terminal S1 peptide with an important receptor binding domain (RBD) region and C-terminal S2 fragment, plays an essential role in viral attachment, and fusion, entry into the target cells which express the viral receptor angiotensin-converting enzyme 2 (ACE2) [7]. There is rapidly growing serological testing evidence that IgM, IgG, and IgA antibodies against spike (S) or nucleocapsid (N) proteins of SARS-CoV-2 evolve rapidly in the serum of both asympomatic and symptomatic COVID-19 infections within one week after infection or onset of symptoms [8-11]. More importantly, these antibodies remain elevated with disease progression and severity in symptomatic COIVD-19 patients [12]. Therefore, anti-SARS-CoV-2 specific antibodies may involve in the pathogenesis and affect the prognosis and outcome of COVID-19. However, little is known about humoral immune responses to other structural and non-structural proteins of SARS-CoV-2 during disease progression and outcome.

In this study, we assumed levels of anti-SARS-CoV-2 IgM and IgG antibodies might help predict disease prognosis and outcome in patients with COVID-19. To enable the global understanding of SARS-CoV-2 specific IgM/IgG responses, we constructed a proteome microarray with 20 out of the 28 predicted proteins of SARS-CoV-2 [13]. Clinical serum specimens were further analyzed

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

on the SARS-CoV-2 proteome microarray, which can provide a high-throughput assay for 12 samples on each microarray and a rapid turnaround time of assay results (within 5 h of sample collection).

1,034 patients hospitalized for confirmed COVID-19 at Tongji hospital from the day of hospitalization to the day of discharge or death were enrolled in this study and were classified into two groups, namely survivors and nonsurvivors based on the known clinical outcome. Serum IgM and IgG profiles for 1,034 patients with COVID-19 on admission were probed using the SARS-CoV-2 proteome microarray, which were further correlated with laboratory biomarkers of disease severity and comorbidities, and with death. We found that elevated ORF7b specific IgM serum levels at presentation is a useful predictor of survival, while high levels of IgG responses to most of non-structural proteins, especially NSP9 and NSP10 are powerful predictions of COVID-19 death. Our results indicate that the set of anti-SARS-CoV-2 antibody signatures are independent from other biomarkers of laboratory and clinical severity factors, which could be used to guide clinical management, vaccine developments, and interventional studies.

Materials

Patient information and data source

1,056 confirmed COVID-19 patients were recruited from Tongji Hospital, Wuhan, China, between 17 February 2020 and 28 April 2020. COVID-19 was diagnosed based on positive SARS-CoV-2 nucleic acid test from respiratory tract specimens or based on clinical diagnosis with clinical symptoms and imaging features of pneumonia on chest computed tomographic (CT) according to the fifth version of COVID-19 diagnostic and treatment guideline published by the National Health Commission of China (NHCC) [14]. Demographic information, medical history, comorbidities, signs and symptoms, chest CT, laboratory findings on admission, and clinical outcomes were collected from electronic medical records. Among these, laboratory biomarkers related with disease severity factors such as the blood routine (leucocytes, lymphocytes, platelets, and neutrophils), liver and kidney functions (aspartate aminotransferase, alanine aminotransferase, lactate dehydrogenase, and creatinine), coagulation function (D-dimer) and infection markers (C-reactive protein, procalcitonin) were performed by automated analyzers according to the

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

manufacturers' instructions. The level of IL-2R in serum was measured by automatic procedure of a solid-phase two-site chemiluminescent immunometric assay via IMMULITE 1000 Analyzer (Siemens). Serum IL-6 was measured by electro-chemiluminescence method (Roche Diagnostics). Serum specimens were collected from each patient on admission and were stored at -80 ºC until use. Serum detection based on proteome microarray and data analysis were performed during April 2020 to July 2020. After excluding individuals whose 23 anti-SARS-CoV-2 antibody indicators were missing more than three, a total of 1,034 eligible participants (524 females and 510 males) with available data from serum proteome microarray and clinical outcomes were enrolled for the final analysis. Among 1,034 eligible participants, some of whom had serial serum samples and collected for a total of 2,973 samples.

Ethical approval

The study was approved by the Ethical Committee of Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China (IRB ID:TJ-C20200128).

Protein microarray fabrication

The microarray used for serum IgM and IgG profiling was prepared as described previously [11, 13]. 20 proteins of SARS-CoV-2 with indicated concentrations, along with the negative (GST, Biotin-control, and eGFP) and positive controls (Human IgG, Human IgM, and ACE2-Fc), were printed in quadruplicate on PATH substrate slide (Grace Bio-Labs, Oregon, USA) to generate identical arrays in a 2×7 subarray format using Super Marathon printer (Arrayjet, UK). The prepared protein microarrays were incubated in blocking buffer (3% BSA in 1×PBS buffer with 0.1% Tween 20) for 3 h, and then stored at -80 °C until use.

Microarray-based serum analysis

The protein microarrays stored at -80 °C were warmed to room temperature before detection and were performed to probe all available seral samples. A 14-chamber rubber gasket was mounted onto each slide to create individual chambers for the 14 identical subarrays. Serum samples were diluted 1:200 in PBS containing 0.1% Tween 20 and a total of 200 μL of diluted serum or buffer

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

only (negative controls) was incubated with each subarray for 2h at 4° C. The arrays were washed with 1×PBST and bound antibodies were detected by incubating with Cy3-conjugated goat anti-human IgG and Alexa Fluor 647-conjugated donkey anti-human IgM (Jackson ImmunoResearch, PA, USA), which were diluted 1: 1,000 in 1×PBST, and incubated at room temperature for 1 h. The microarrays were then washed with 1×PBST and dried by centrifugation at room temperature and scanned by LuxScan 10K-A (CapitalBio Corporation, Beijing, China) with the parameters set as 95% laser power/ PMT 550 and 95% laser power/ PMT 480 for IgM and IgG, respectively. Data of fluorescent intensity (FI) from each microarray was extracted by GenePix Pro 6.0 software (Molecular Devices, CA, USA). The result of FI for each serum response to each protein was defined as the median of the foreground subtracted by the median of background for each spot and then averaged the triplicate spots for each protein. The result of the protein-specific antibody in the serum was expressed as $\log_2$(FI). IgG and IgM data were analyzed separately.

Statistical analysis

Shapiro-Wilk test was used to test data normality. Two-tailed t-test was conducted to test difference in means between survivor and nonsurvivor groups, Mann-Whitney U test was performed to test difference in skewed parameters. Chi-square tests or Fisher's exact test, when appropriate, was used for categorical variables. Cox proportional-hazards model was performed to estimate the hazard ratios (HRs) and 95% confidence intervals (CIs) of COVID-19 mortality for individual levels of virus-specific IgM and IgG responses categorized into tertiles according to distributions. The lowest tertiles were assigned to be the reference groups. Age and sex were included in Model 1. In Model 2, we further adjusted hypertension (yes/no), diabetes (yes/no), lymphopenia (<1.1, ≥1.1, ×10^9/L), increased alanine aminotransferase (<40, ≥41, U/L), and increased lactate dehydrogenase (<214, ≥214, U/L). Linear trend *p*-values were calculated by modeling the median value of each metal tertiles as a continuous variable in the adjusted models. Spearman's rank correlation analysis was performed to explore the correlations between virus-specific IgM/IgG responses and laboratory results in the study population. The principal component analysis (PCA) based on the 20 proteins of SARS-CoV-2 specific IgM/IgG responses

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

was used to optimize the type of data and extract principal components (PCs). Proteins of SARS-CoV-2 specific IgM /G responses with factor loadings over 0.7 on a particular PC were regarded as main contributors of it. Each PC was modeled into the Cox proportional-hazards models as tertiles to evaluate the association with SARS-CoV-2 specific IgG responses and COVID-19 mortality.

In addition, we also conducted sensitivity analyses by exclusion of patients who died within the initial 7 days after hospitalization to avoid reverse causality. The associations of viral specific IgG and IgM responses with the risk of COVID-19 mortality were also evaluated using restricted cubic splines, with 3 konts defined at the 5th, 50th, and 90th percentiles of its distribution; the reference value (HR = 1.00) set at the 10th percentile. The measured level replaced with the mean level ± 3SD for all observations with measured concentrations above this value. The results of antibodies were classified as two groups of the high levels (≥ median) and low levels (< median) based on the medians of IgM and IgG responses to each protein of all involved patients and further correlated these results with on day 66 mortality of COVID-19 by Kaplan-Meier survival curve and log-rank test. The receiver operating characteristic curve was conducted for the prediction of COVID-19 survival and death, and 1,000 times computational cross-validations were conducted. For each cross-validation procedure, 477 survivors and 39 nonsurvivors were randomly selected as the training set. The rest of the samples were treated as the testing set (478 survivors and 40 nonsurvivors). Further validation was conducted using the second and third serial samples after hospitalization (validation2A, n = 633, validation2B, n = 382). Loess regression was used to establish the kinetics of SARS-CoV-2 specific antibodies. Cluster analysis was performed with pheatmap package of R. SAS (version 9.4), R (version 4.0.0), and SPSS (version 23.0) were used to conduct statistical analyses when applicably used. Two-sided statistical tests were considered to be significant at *p* values below 0.05.

Results

Characteristics of the study population

1,034 participants, having available serum microarray results and consisting of 955 survivors and 79 nonsurvivors, were enrolled in this study. Baseline characteristics of participated patients based

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

on electronic medical records were analyzed as **Table 1**. The median age of all enrolled patients was 63 years old (IQR, 51-71). The median intervals from onset of symptoms to hospital admission, from onset of symptoms to recovery, and from onset of symptoms to death were 13 days (IQR, 8-21), 41 days (IQR, 33-52), and 32 days (IQR, 25-39), respectively. The median length of all COVID-19 patients' hospital stay was 24 days (IQR, 15-35). 37% patients with COVID-19 had hypertension and 18.5% with diabetes. 30.7% patients had lymphopenia, and increased levels of lactate dehydrogenase and alanine aminotransferase were detected in 43% and 25.4% patients, respectively. Consistent with previous reports [15, 16], nonsurvivors were more likely to be male and older than survivors ($p < 0.001$). Our study also demonstrated that higher proportion of abnormal laboratory results and shorter hospitalization time were obtained in nonsurvivors than those of survivors ($p < 0.001$).

The magnitude of antibody responses correlates with disease outcome

To establish the association of anti-SARS-CoV-2 IgM and IgG antibodies with COVID-19 survival and death, serum collected from each involved patients on admission was used for microarray-based serum analysis. Based on the FI extracted from proteome microarray for each serum of 1034 patients, we first presented overall visualizations and quantitative data of IgM (**Figure S1** and **Table S1)** and IgG (**Figure S2** and **Table S2)** profiles against 20 proteins of SARS-CoV-2, respectively. We demonstrated that higher levels of both IgM and IgG responses against N, ORF3a, and ORF7b proteins were induced in survivors than those of nonsurvivors, apart from ORF6 specific IgM antibody ($p < 0.05$, **Figure 1a**, **Table S1** and **Table S2**). On the contrary, nonsurvivors elicited higher levels of NSP10 specific IgM antibody ($p < 0.05$, **Table S1**) and IgG responses against E, NSP1, NSP2, NSP4, NSP5, NSP7, NSP8, NSP9, NSP10, RNA-dependent RNA polymerase (RdRp or NSP12), NSP14, NSP15, NSP16, ORF3b, and ORF9b proteins than survivors ($p < 0.05$, **Figure 2a** and **Table S1**). The levels of ORF7b IgM, NSP9 IgG, and NSP10 IgG fluctuated with the days after symptoms onset, but they were not obvious (**Figure S4**). Our results strongly indicate that the magnitude of IgM or IgG responses against most of structural and non-structural proteins of SARS-CoV-2 might involve in the prognosis and outcome of COVID-19.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

High levels of ORF7b IgM antibody predict disease survival

To assess the relationship of the magnitude of IgM antibodies with the mortality risk of COVID-19, the HRs (95% CIs) for the mortality risk of COVID-19 associated with the levels of IgM responses against different proteins of SARS-CoV-2 were categorized into tertiles (**Table 2** and **Table S3**). We first analyzed the effects of age and gender on disease death as model 1. After adjusting for age and gender, we found that IgM responses to N, ORF3a, or ORF7b were significantly associated with COVID-19 mortality (all p trend < 0.05), while no significant association was observed among other protein-specific IgM responses and the death, respectively. Previous studies reported that comorbidities and laboratory biomarkers related with the function of important organs are also the risk factors resulting in the COVID-19 death [16, 17]. We further adjusted the association for hypertension, diabetes, lymphopenia, increased alanine aminotransferase and lactate dehydrogenase as shown in model 2. Interestingly, only the IgM response to ORF7b was significantly associated with the mortality risk of COVID-19 (T2 vs T1: HR = 0.86, 95% CI: 0.51-1.44; T3 vs T1: HR = 0.19, 95% CI: 0.07-0.55; p trend = 0.002, **Table 2**), even independently of the factor excluding patients who died within 7 days after admission (**Table S4**). Moreover, Kaplan-Meier survival curve also showed that COVID-19 patients with early detected high level of ORF7b specific IgM antibody ($\log_2$FI ≥ 7.5) on admission had lower risk of morality than the patients with low levels ($\log_2$FI < 7.5) during the following-up observation period of 66 days ($p < 0.001$, **Figure 1b**). The linear association between the levels of IgM response to ORF7b and the mortality risk of COVID-19 was further demonstrated by spline regression analysis ($p = 0.0013$, **Figure 1c**). Taken together, our results suggest that high levels of ORF7b IgM antibody upon admission are negatively correlated with the mortality risk of COVID-19.

High levels of IgG antibody against NSP7 or NSP9 predict disease death

To establish the associations between anti-SARS-CoV-2 IgG responses with risk of death, the relationship between the levels of IgG antibody against 20 proteins of SARS-CoV-2 with the mortality risk of COVID-19 was shown in **Table 2** and **Table S5**, respectively. After adjusting for

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

age and gender in model 1, we observed that the levels of IgG responses to N, NSP1, NSP4, NSP7, NSP8, NSP9, NSP10, RdRp (NSP12), NSP14, ORF3a, ORF3b, ORF7b, and ORF9b were significantly associated with the risk of COVID-19 mortality, respectively (all with p trend < 0.05). After further adjustment for potential confounders in model 2, the association of the multivariable adjusted HRs (95% CI) of COVID-19 mortality with these IgG responses remained statistically significant, except IgG responses to N, ORF3a and ORF7b (**Table 2**). After excluding patients who died within the first 7 days after admission, the associations between IgG responses to NSP4, NSP7, NSP9, NSP10, RdRp (NSP12), NSP14, and ORF3b and the mortality risk of COVID-19 remained statistically significant (**Table S6**). As shown in **Figure 2,** Kaplan-Meier survival curve also supported that COVID-19 patients with higher levels of specific IgG responses against NSP1 ($\log_2$FI ≥ 8.2), NSP4 ($\log_2$FI ≥ 7.9), NSP7 ($\log_2$FI ≥ 9.4,), NSP8 ($\log_2$FI ≥ 7.8), NSP9 ($\log_2$FI ≥ 8.7), NSP10 ($\log_2$FI ≥ 6.3), RdRp ($\log_2$FI ≥ 8.1), NSP14 ($\log_2$FI ≥ 7.4), ORF3b ($\log_2$FI ≥ 8.7), and ORF9b ($\log_2$FI ≥ 8.0) had higher risk of morality after admission, respectively (all $p < 0.05$). Most importantly, the linear association between high levels of IgG antibodies against only NSP9 or NSP10 with COVID-19 death was demonstrated by spline regression analysis with p values 0.0073 and 0.0003, respectively (**Figure 3**). Taken together, our results suggest that high levels of IgG responses to 6 non-structural proteins such as NSP4, NSP7, NSP9, NSP10, RdRp (NSP12), NSP14 and 1 accessory protein ORF3b are positively correlated with the mortality risk of COVID-19.

Principal component analyses support the predictive role of IgG antibody as disease death

To further establish the association among IgG responses to different proteins with the outcome of COVID-19, we further conducted principal component analyses (PCs) and screened hypothetical new variables that account for as much as possible of the variance, in order to reduce the dimension of data and the complexity of data with the least loss of original information. The HRs (95%CIs) for the COVID-19 mortality according to PCs tertiles are presented in **Table 3**. Four PCs with eigen values > 1 were extracted, accounting for 71.95% of the total variance. Of four PCs, we found that only PC1 had the statistical association with the COVID-19 mortality (p trend = 0.004, **Table 3**), whatever adjusting age and sex, or further for hypertension, diabetes,

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

lymphopenia, increased alanine aminotransferase and lactate dehydrogenase. Interestingly, IgG responses to NSP1, NSP2, NSP4, NSP7, NSP8, NSP9, NSP10, RdRp (NSP12), NSP14, NSP15, NSP16, ORF3b, and ORF9b remained main contributors of PC1 (**Table S7**), in line with our above findings.

IgG responses positively correlate with laboratory biomarker measurements related with severity factors

Previous studies established the associations between COVID-19 death with several laboratory biomarker measurements related with severity factors, such as lymphocyte count, procalcitonin, C-reactive protein, lactate dehydrogenase, D-dimer, IL-2R, and IL-6 [15-17]. Therefore, linear correlation among SARS-CoV-2 specific IgM/IgG responses and these biomarkers was further analyzed (**Table S8**). Interestingly, the biomarker of lymphocyte count was positively correlated with the ORF7b specific IgM antibody (r_s=0.21, $p < 0.01$) but negatively correlated with IgG responses to NSP1, NSP4, NSP7, NSP8, NSP9, NSP10, RdRp (NSP12), NSP14, ORF3b, and ORF9b, respectively (all $p < 0.01$). Moreover, IgM response to ORF7b was negatively correlated with pro-inflammatory factors such as procalcitonin, C-reactive protein, lactate dehydrogenase, D-dimer, IL-2R, and IL-6, respectively (all $p < 0.01$). However, these pro-inflammatory factors except IL-2R and IL-6 were positively correlated with all of these IgG responses. The levels of IL-2R were also positively correlated with these IgG responses except NSP9 and ORF9b specific IgG antibodies, while NSP8, NSP10, RdRp (NSP12) and NSP14 specific IgG antibodies were positively correlated with IL-6, respectively.

High prediction efficacy for clinical outcome using validation models

It is a common practice to validate "potential biomarker" by independent sample cohort. However, it is very difficult to collect more COVID-19 samples at this moment, because of very strict regulations of sample handling and very few COVID-19 patients in China. To assure the reliability of our finding, alternatively, we performed computational cross-validation based on the large sample cohort that we have already analyzed, by following a protocol that we have established previously [18]. ORF7b IgM, NSP9 IgG and NSP10 IgG were confirmed as three potential

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

biomarkers for predicting clinical outcome (**Figure 4a**). Through computational cross validation (validation1), the AUC of ORF7b IgM for predicting COVID-19 survival was determined as 0.74 (cut-off = 7.59, **Figure 4b**). The AUCs of NSP9 and NSP10 for predicting COVID-19 death were 0.66 and 0.68, respectively (cut-off = 9.13 for NSP9 and 6.29 for NSP10, **Figure 4b**). Furthermore, we evaluated the prognosis efficacy of these three potential biomarkers using the samples collected at the second time point (n = 633, Validation2A) and the third time point (n = 382, Validation2B) after hospitalization from the 1,034 patients. In validation 2A and validation 2B, the accuracies of ORF7b IgM (**Figure 4c-d**), NSP9 IgG, and NSP10 IgG (**Figure 4c-d**) were 0.69 and 0.66, 0.77 and 0.76, 0.76 and 0.77, respectively. Besides, we depicted the dynamic changes of ORF7b IgM (**Figure 5a**), NSP9 IgG (**Figure 5b**), and NSP10 IgG (**Figure 5c**) for 1,034 COVID-19 patients with 2,973 consecutive samples. They all maintained relatively stabilized levels after symptoms onset.

Discussion

In this study, we aimed to understand the role of anti-SARS-CoV-2 IgM and IgG responses on COVID-19 disease prognosis and outcome. Our findings have important indications for better control of COVID-19 pandemic.

Firstly, we established a rapid and high-throughput assay platform based on proteome microarrays to measure IgM and IgG responses against 20 SARS-CoV-2 proteins in COVID-19 patients. After analyzing 1,034 hospitalized patients, we established that COVID-19 is associated with high levels of IgM and IgG responses to 11 non-structural proteins and 3 accessory proteins of SARS-CoV-2 at presentation. Importantly, our observations indicate that antibody patterns are predictive of COVID-19 survival and mortality, independently of demographics and comorbidities, but also of standard clinical biomarkers of disease severity. We found that OFR7b IgM response is independently the prognostic marker of survival, and IgG antibodies against 6 non-structural proteins NSP4, NSP7, NSP9, NSP10, RdRp (NSP12), NSP14, and 1 accessory protein ORF3b, especially NSP9 and NSP10 are predictors of death after adjusting for the demographic features and comorbidities. Early antibody measurements based on our established serum proteome microarray analysis as predictors of survival and death, therefore, raise the importance of using

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

antibody levels for rapidly improving clinical management, treatment decisions and rational allocation of medical resources in short supply during the process of dealing with COVID-19 pandemic.

ORF7b is an accessory protein of SARS-CoV-2 with a length of 43 amino acids, which is also highly similar to the SARS-CoV ORF7b but absent from MERS-CoV [6]. Previous studies reported that the SARS-CoV ORF7b is not only an accessory protein but also a structural component of the virion, which is a viral attenuation factor during early phase of infection [19-21]. Our results implicate that anti-SARS-CoV-2 ORF7b specific antibody might play a protective role against COVID-19 infection and disease, further supported by the evidence that there was no significant association between levels of ORF7b IgG response and the risk of COVID-19 mortality. Therefore, ORF7b might be a promising target antigen for vaccine development. Unfortunately, high mutation rate was observed for current prevalent SARS-CoV-2 strains worldwide (**Table S9**). The effects of ORF7b and its mutations on immunization and host-virus interaction deserve further exploration. Although the function of each non-structural proteins of the SARS-CoV-2 is not yet clear, their protein sequences are highly similar to that of SARS-CoV. Most non-structural proteins always locate in the core of virion and play important roles in the pathogenesis. For example, RdRp, also called NSP12 of SARS-CoV, can catalyze the synthesis of viral RNA and plays an important role in the replication and transcription cycle of the virus [22, 23]. RdRp itself performs the polymerase reaction with limited efficiency, whereas NSP7 and NSP8 as co-factors can significantly stimulate its polymerase activity [22]. Previous studies based on cryogenic electron microscopy (cryo-EM) indicated that the viral polymerase RdRp-NSP7-NSP8 complex might be an excellent target for new therapeutics of SARS and COVID-19 [23, 24]. NSP1 of the SARS-CoV may promote viral gene expression and immune escape by affecting interferon-mediated signal transduction [25]. NSP4 is a multichannel membrane protein, which is an essential protein for viral replication [26]. NSP9 plays a role of dimeric ssRNA binding protein during viral replication [27, 28]. NSP10 interacts with NSP14 and regulates ribose-2'-O-MTase activities involved in mRNA capping [28-30]. Therefore, the relationship between these IgG responses and COVID-19 mortality indicates that IgG antibodies against these non-structural proteins might involve in the pathogenesis of SARS-CoV-2.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Even if not detected in preclinical and clinical trials, memory B cell responses to our reported non-structural proteins might be induced after a healthy individual vaccinated with inactivated or attenuated SARS-CoV-2 vaccine candidates. After exposure to or reinfection with SARS-CoV-2, IgG antibodies to these proteins might evolve in a higher, faster and stronger fashion in recipients of immunization. However, beneficial evidences from current animal models vaccinated with inactivated SARS-CoV-2 vaccine candidates against infection or reinfection [31, 32], attention should be paid to the risk of immunization with these kinds of vaccines to increase SARS-CoV-2 vulnerability, as demonstrated by a previous study that a double-inactivated SARS-CoV vaccine could elicit eosinophilic and immunoenhancing pathology, as well as poor protection, especially in aged animals upon challenge with virulent strains [33]. In addition, a newly reported COVID-19 case in USA with secondary infection with SARS-CoV-2 had more serious illness [34], also indicating the potential risk of the preexisting immunity. Although antibody-dependent enhancement (ADE) of SARS-CoV-2 infection has not been issued as yet [35-37], decay of antibodies was found to be related with the loss of Bcl-6-expressing T follicular helper cells and germinal centers influencing memory B cells in critically ill patients with COVID-19 [38], which might significantly decrease the risk of administration with convalescent plasma containing IgG antibodies to the above non-structural proteins to treat critically ill patients with COVID-19, especially when used our microarray to screen suitable donors of convalescent plasma before treatment.

In conclusion, we provided a novel application of SARS-CoV-2 proteome microarray to detect serum IgM/IgG responses for early predicting COVID-19 survival and death. Our results demonstrate that high level of IgM antibody against ORF7b at the time of hospitalization is an independent predictor of patient survival, while IgG responses to NSP9 and NSP10 possess significant predictive power for patient death. Our research might improve clinical management and guide the development of effective medical interventions and vaccines by enhancing the further understanding of pathogenesis of COVID-19.

Acknowledgments

We thank Prof. H. Eric Xu (Shanghai Institute of Materia Medica) for providing RdRp protein. We

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

also thank Healthcode Co., Ltd., Hangzhou Bioeast biotech Co., Ltd. and Vacure Biotechnology Co.,Ltd. for providing the proteins.

Author Contributors

X-L.F., S-C.T., and F.W. performed experiments and designed the study. Q.L., C-Z.Y., and Y.L. performed the experiments. H-Y.H., Z-Y.S., and B.Z. collect specimens. D-Y.L., J-B.X., Z-Q.O., J-L.B., and Z-W.X. prepared the reagents. Y-D.Z., Z-J.Y., H.C., Y-X.Z., and X-S.L. analyzed the data. M-A.H, X-N.W., H-W.J., H-N.Z., H.Q., and S-J.G. took responsibility for the accuracy of the data analysis. X-L.F., Q.L., and C-Z.Y. wrote the manuscript with suggestions from other authors.

Conflict of Interest Disclosures: The authors declare no conflicts of interest.

Funding/Support: This work was supported by grants from the National Mega-Projects of Science Research for the 13th Five-year Plan of China (No. 2018ZX10302302002-001), the Natural Science Foundation of China (No. 81971909), and the Fundamental Research Funds for the Central Universities (HUST COVID-19 Rapid Response Call No. 2020kfyXGYJ040). This work was partially supported by National Key Research and Development Program of China Grant (No. 2016YFA0500600), National Natural Science Foundation of China (No. 31970130, 31670831, 31370813 and 31501054)

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

References

[1] Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med 2020; 382: 727-33.

[2] WHO. Coronavirus disease (COVID-2019) situation reports. 2020; Available at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports. Accessed 20 June, 2020.

[3] Sempowski GD, Saunders KO, Acharya P, et al. Pandemic Preparedness: Developing Vaccines and Therapeutic Antibodies For COVID-19. Cell 2020; 181: 1458-63.

[4] Wiersinga WJ, Rhodes A, Cheng AC, et al. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. JAMA 2020.

[5] Dhama K, Khan S. Coronavirus Disease 2019-COVID-19. Clin Microbiol Rev 2020; 33.

[6] Wu A, Peng Y, Huang B, et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. Cell Host Microbe 2020; 27: 325-8.

[7] Walls AC, Park YJ, Tortorici MA, et al. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell 2020; 181: 281-92.e6.

[8] Wu J, Liang B, Chen C, et al. SARS-CoV-2 infection induces sustained humoral immune responses in convalescent patients following symptomatic COVID-19. 2020; Preprint at medRxiv: https://www.medrxiv.org/content/medrxiv/early/2020/07/24/2020.07.21.20159178.full.pdf.

[9] Vogelzang EH, Loeff FC, Derksen NI, et al. Development of a SARS-CoV-2 total antibody assay and the dynamics of antibody response over time in hospitalized and non-hospitalized patients with COVID-19. 2020; Preprint at medRxiv: https://www.medrxiv.org/content/medrxiv/early/2020/06/19/2020.06.17.20133793.full.pdf.

[10] Li K, Wu M, Huang B, et al. The Dynamic Changes of Antibodies against SARS-CoV-2 during the Infection and Recovery of COVID-19. 2020; Preprint at medRxiv: https://www.medrxiv.org/content/medrxiv/early/2020/05/21/2020.05.18.20105155.full.pdf.

[11] Lei Q, Li Y, Hou H, et al. Antibody dynamics to SARS-CoV-2 in asymptomatic COVID-19 infections. 2020; Preprint at: https://www.medrxiv.org/content/medrxiv/early/2020/08/04/2020.07.09.20149633.full.pdf.

[12] Lynch KL, Whitman JD, Lacanienta NP, et al. Magnitude and kinetics of anti-SARS-CoV-2 antibody responses and their relationship to disease severity. Clin Infect Dis 2020.

[13] Jiang H, Li Y, Zhang H, et al. SARS-CoV-2 proteome microarray for global profiling of COVID-19 specific IgG and IgM responses. Nature communications 2020; 11: 3581.

[14] Government of the People's Republic of China. New coronavirus pneumonia diagnosis and treatment plan (Fifth Edition). 2020; Available at: http://www.nhc.gov.cn/yzygj/s7653p/202002/d4b895337e19445f8d728fcaf1e3e13a.shtml. Accessed May 18, 2020.

[15] Chen T, Wu D, Chen H, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. BMJ 2020; 368: m1091.

[16] Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020; 395: 1054-62.

[17] Yu C, Lei Q, Li W, et al. Clinical Characteristics, Associated Factors, and Predicting COVID-19 Mortality Risk: A Retrospective Study in Wuhan, China. Am J Prev Med 2020; 59: 168-75.

[18] Yang L, Wang J, Li J, et al. Identification of Serum Biomarkers for Gastric Cancer Diagnosis Using a Human Proteome Microarray. Mol Cell Proteomics 2016; 15: 614-23.

[19] Schaecher SR, Diamond MS, Pekosz A. The Transmembrane Domain of the Severe Acute Respiratory Syndrome Coronavirus ORF7b Protein Is Necessary and Sufficient for Its Retention in the Golgi Complex. Journal of Virology 2008; 82: 9477-91.

[20] Schaecher SR, Mackenzie JM, Pekosz A. The ORF7b protein of severe acute respiratory syndrome coronavirus (SARS-CoV) is expressed in virus-infected cells and incorporated into SARS-CoV particles. J Virol 2007; 81: 718-31.

[21] Pfefferle S, Krähling V, Ditt V, et al. Reverse genetic characterization of the natural genomic deletion in SARS-Coronavirus strain Frankfurt-1 open reading frame 7b reveals an attenuating function of the 7b protein in-vitro and in-vivo. Virol J 2009; 6: 131.

[22] Subissi L, Posthuma CC, Collet A, et al. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc Natl Acad Sci U S A 2014; 111: E3900-9.

[23] Yan G, Liming Y, Yucen H, et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science 2020; 368: eabb7498.

[24] Kirchdoerfer RN, Ward AB. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. Nat Commun 2019; 10: 2342.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

[25] Jauregui AR, Savalia D, Lowry VK, et al. Identification of residues of SARS-CoV nsp1 that differentially affect inhibition of gene expression and antiviral signaling. PLoS One 2013; 8: e62416.
[26] Angelini MM, Akhlaghpour M, Neuman BW, et al. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. mBio 2013; 4.
[27] Miknis ZJ, Donaldson EF, Umland TC, et al. Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth. J Virol 2009; 83: 3007-18.
[28] Snijder EJ, Decroly E, Ziebuhr J. The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. Adv Virus Res 2016; 96: 59-126.
[29] Bouvet M, Debarnot C, Imbert I, et al. In vitro reconstitution of SARS-coronavirus mRNA cap methylation. PLoS Pathog 2010; 6: e1000863.
[30] Bouvet M, Imbert I, Subissi L, et al. RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proc Natl Acad Sci U S A 2012; 109: 9372-7.
[31] Wang H, Zhang Y, Huang B, et al. Development of an Inactivated Vaccine Candidate, BBIBP-CorV, with Potent Protection against SARS-CoV-2. Cell 2020; 182: 713-21.e9.
[32] Deng W, Bao L. Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. Science 2020; 369: 818-23.
[33] Bolles M, Deming D, Long K, et al. A double-inactivated severe acute respiratory syndrome coronavirus vaccine provides incomplete protection in mice and induces increased eosinophilic proinflammatory pulmonary response upon challenge. J Virol 2011; 85: 12201-15.
[34] Tillett RL, Sevinsky JR, Hartley PD, et al. Genomic evidence for a case of reinfection with SARS-CoV-2. 2020; Preprint at SSRN: https://ssrn.com/abstract=3681489.
[35] Arvin AM, Fink K, Schmid MA, et al. A perspective on potential antibody-dependent enhancement of SARS-CoV-2. Nature 2020.
[36] Bournazos S, Gupta A, Ravetch JV. The role of IgG Fc receptors in antibody-dependent enhancement. Nat Rev Immunol 2020; 1-11.
[37] Eroshenko N, Gill T. Implications of antibody-dependent enhancement of infection for SARS-CoV-2 countermeasures. Nat Biotechnol 2020; 38: 789-91.
[38] Kaneko N, Kuo HH, Boucau J, et al. Loss of Bcl-6-Expressing T Follicular Helper Cells and Germinal Centers in COVID-19. Cell 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table 1. Baseline characteristics of participated patients with COVID-19

	All patients	**Survivors**	**Nonsurvivors**	***p* value**
N	1,034	955	79	
Age, median (IQR)	63(51-71)	62(51-70)	68(59-78)	<0.001
Female, n (%)	524(50.7)	491(51.4)	33(41.8)	0.10
Time from onset to admission, Median (IQR), days	13(8-21)	13(8-22)	11(5-19)	0.03
Length of hospital stay, Median (IQR), days	24(15-35)	25(16-35)	18(9-26)	<0.001
Time from onset to outcome, Median (IQR), days	40(33-52)	41(33-52)	32(25-39)	<0.001
Comorbidity, n (%)				
Hypertension	383(37.0)	355(37.2)	28(35.4)	0.76
Diabetes	191(18.5)	173(18.1)	18(22.8)	0.30
Coronary heart disease	68(6.6)	57(6.0)	11(13.9)	0.006
Chronic obstructive pulmonary disease	6(0.6)	3(0.3)	3(3.8)	0.007
Cerebrovascular disease	44(4.3)	37(3.9)	7(8.9)	0.07
Chronic liver disease	21(2.0)	19(2.0)	2(2.5)	0.67
Chronic renal disease	23(2.2)	20(2.1)	3(3.8)	0.41
Cancer	45(4.4)	35(3.7)	10(12.7)	0.001
Laboratory results, n (%)				
Lymphopenia, <1.1×10^9/L	294(30.7)	234(26.4)	60(83.3)	<0.001
Neutrophilia, ≥6.3×10^9/L	181(18.9)	125(14.1)	56(77.8)	<0.001
Thrombocytopenia, ≥350×10^9/L	64(6.7)	62(7.0)	2(2.7)	0.16
Leukocytosis, ≥9.5×10^9/L	146(15.2)	98(11.1)	48(65.8)	<0.001
Increased lactate dehydrogenase, ≥214 U/L	405(43.0)	342(39.3)	63(88.7)	<0.001
Increased alanine aminotransferase, ≥41 U/L	239(25.4)	217(24.9)	22(31.0)	0.26
Increased aspartate aminotransferase, ≥40 U/L	129(13.7)	101(11.6)	28(40.0)	<0.001
Increased creatinine, ≥104 μmol/L	57(6.3)	39(4.7)	18(26.1)	<0.001
Increased C-reactive protein, ≥3mg/L	330(45.9)	289(42.7)	41(97.6)	<0.001
Increased procalcitonin, ≥0.05 ng/ml	159(29.3)	122(24.3)	37(92.5)	<0.001
Increased D-dimer, ≥0.5 mg/L	361(59.4)	302(55.1)	59(98.3)	<0.001
Increased IL2R, >710 U/mL	67(16.2)	57(14.4)	10(55.6)	<0.001
Increased IL6, >7 ng/L	98(23.5)	82(20.6)	16(88.9)	<0.001

Data are shown as medians (IQR) or number (%), respectively. IQR, inter-quartile ranges.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table 2. Hazard ratio (95%CI) for COVID-19 mortality according to tertiles of anti-SARS-CoV-2 specific IgM/IgG responses

Proteins	Tertile of proteins [log_2(FI)]			*p* trend	FDR
	T1	T2	T3		
IgM response					
ORF7b					
Model 1	1.00	0.67(0.42-1.09)	0.12(0.04-0.34)	**<0.001**	**0.003**
Model 2	1.00	0.86(0.51-1.44)	0.19(0.07-0.55)	**0.002**	**0.03**
IgG response					
NSP1					
Model 1	1.00	3.05(1.38-6.71)	3.76(1.77-8.03)	0.0006	0.02
Model 2	1.00	2.84(1.21-6.63)	3.10(1.38-6.99)	0.02	0.10
NSP4					
Model 1	1.00	1.90(0.98-3.66)	2.06(1.09-3.90)	0.03	0.17
Model 2	1.00	2.74(1.29-5.85)	2.60(1.24-5.46)	0.03	0.16
NSP7					
Model 1	1.00	4.43(1.85-10.62)	4.94(2.10-11.64)	0.0003	0.009
Model 2	1.00	4.01(1.52-10.53)	4.28(1.67-10.98)	0.008	0.06
NSP8					
Model 1	1.00	2.71(1.23-5.98)	3.91(1.84-8.32)	0.0002	0.006
Model 2	1.00	2.34(0.99-5.52)	3.20(1.42-7.21)	0.009	0.07
NSP9					
Model 1	1.00	1.92(0.92-4.01)	3.28(1.65-6.54)	**0.0003**	**0.009**
Model 2	1.00	1.40(0.64-3.07)	2.69(1.29-5.61)	**0.005**	**0.045**
NSP10					
Model 1	1.00	3.55(1.46-8.59)	5.36(2.28-12.60)	**<0.0001**	**0.003**
Model 2	1.00	3.19(1.22-8.38)	4.89(1.92-12.46)	**0.0005**	**0.02**
RdRp					
Model 1	1.00	2.17(1.00-4.69)	3.57(1.74-7.32)	0.0002	0.006
Model 2	1.00	2.31(1.02-5.20)	2.80(1.30-6.02)	0.02	0.10
NSP14					
Model 1	1.00	1.75(0.79-3.85)	3.49(1.70-7.14)	0.0001	0.003
Model 2	1.00	1.37(0.59-3.19)	2.65(1.23-5.71)	0.007	0.06
ORF3b					
Model 1	1.00	1.63(0.77-3.43)	3.20(1.66-6.17)	0.0001	0.003
Model 2	1.00	1.68(0.76-3.70)	2.69(1.34-5.38)	0.02	0.13
ORF9b					
Model 1	1.00	1.66(0.87-3.15)	2.02(1.11-3.68)	0.02	0.13
Model 2	1.00	1.72(0.86-3.43)	2.11(1.11-4.04)	0.03	0.14

Model 1: Adjusted for age and sex.
Model 2: Additional adjustment for hypertension, diabetes, lymphopenia, increased alanine aminotransferase, and increased lactate dehydrogenase.
FI: Fluorescence Intensity.
FDR: False discovery rate adjusted *p*-values.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table 3. Hazard ratio (95%CI) for COVID-19 mortality according to tertiles of principal components of anti-SARS-CoV-2 specific IgG responses

Proteins	Tertile of principal components			*p* trend
	T1	T2	T3	
PC1				
Model 1	1.00	2.17(1.05-4.51)	2.79(1.40-5.59)	**0.004**
Model 2	1.00	1.66(0.76-3.65)	2.24(1.07-4.68)	**0.03**
PC2				
Model 1	1.00	0.70(0.43-1.13)	0.31(0.16-0.61)	<0.001
Model 2	1.00	0.89(0.53-1.51)	0.62(0.31-1.25)	0.20
PC3				
Model 1	1.00	0.69(0.42-1.14)	0.48(0.26-0.88)	0.01
Model 2	1.00	0.82(0.47-1.41)	0.72(0.38-1.39)	0.30
PC4				
Model 1	1.00	0.70(0.40-1.21)	0.98(0.59-1.65)	0.91
Model 2	1.00	0.94(0.52-1.72)	1.24(0.71-2.16)	0.47

PC: principal component. The main contributors are NSP1, NSP2, NSP4, NSP7, NSP8, NSP9, NSP10, RdRp, NSP14, NSP15, NSP16, ORF3b, and ORF9b for PC1; S1, N, N-Nter, and N-Cter for PC2; ORF7b for PC3.

Model 1: Adjusted for age and sex.

Model 2: Additional adjustment for hypertension, diabetes, lymphopenia, increased alanine aminotransferase, and increased lactate dehydrogenase.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

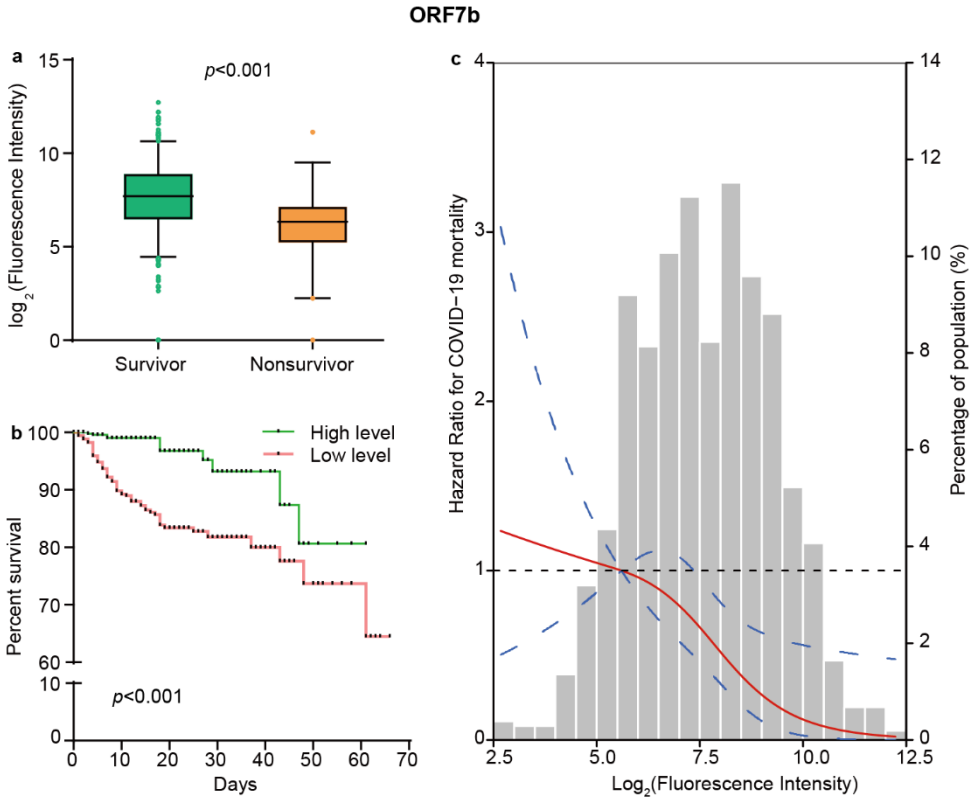


Figure 1. The levels of ORF7b IgM responses independently predict survival of COVID-19. (a) Comparison of the levels of IgM response to ORF7b between 955 survivors and 79 nonsurvivors. The boxplots show medians (middle line) and the third and the first quartiles (boxes), while the tentacles show 97.5 and 2.5 percentiles of the upper and lower parts of the box. (b) Kaplan-Meier survival curves of patients with different levels of IgM antibody against ORF7b. Based on the median level of ORF7b specific IgM responses, patients were classified as both high and low level groups. (c) The restricted cubic spline for the association between ORF7b IgM and risk of COVID-19 mortality. The lines represent adjusted hazard ratios (HRs) based on restricted cubic splines for the levels of ORF7b IgM in Cox regression model. Knots were placed at the 5th, 50th, and 95th percentiles of the distribution of ORF7b specific IgM levels, and the reference value was set at the 10th percentile. Age, sex, diabetes, hypertension, lymphopenia, increased alanine aminotransferase, and increased lactate dehydrogenase were used as adjustment factors.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

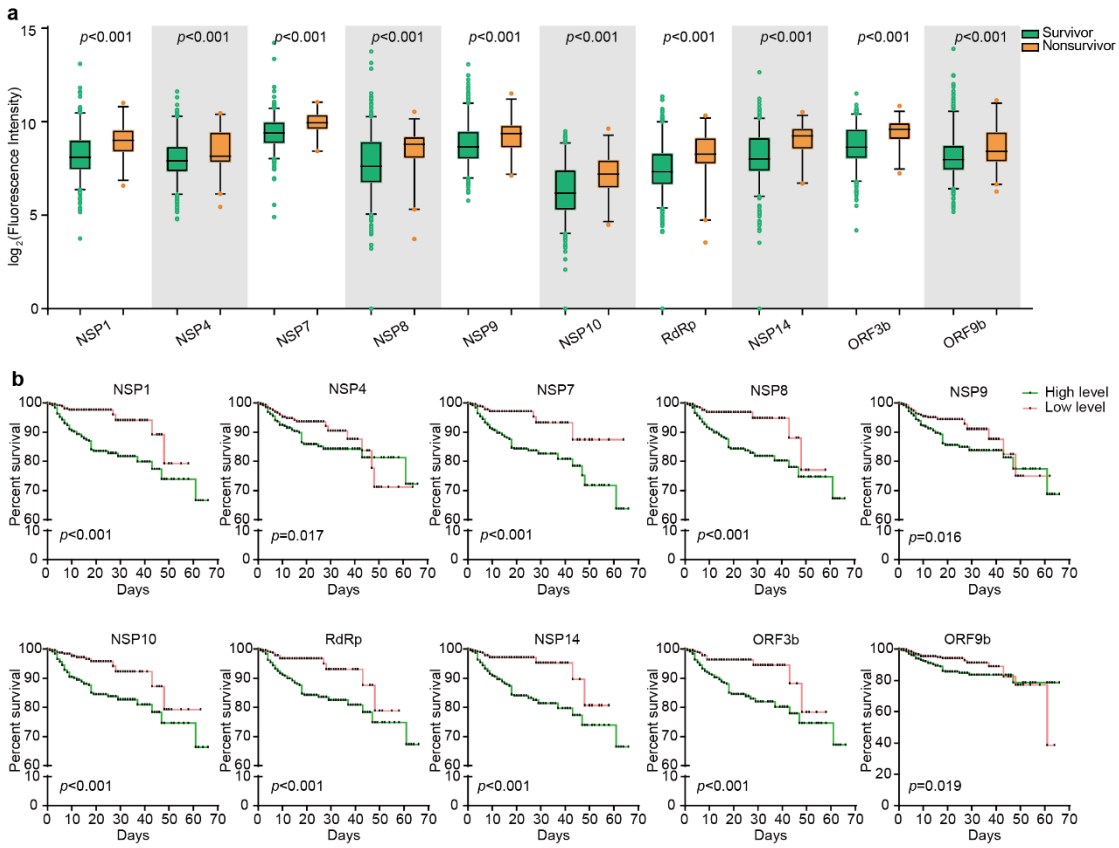


Figure 2. The levels of IgG responses to non-structural proteins correlate with and independently predict the death. (a) Comparison of the levels of IgG response to **non-structural proteins** from the proteome of SARS-CoV-2 between 955 survivors and 79 nonsurvivors. The boxplots show medians (middle line) and the third and the first quartiles (boxes), while the tentacles show 97.5 and 2.5 percentiles of the upper and lower parts of the box. (b) Kaplan-Meier survival curves of patients with high and low levels of IgG antibodies to each protein. Based on the median level of IgG responses to each protein, patients were classified as both high and low level groups.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

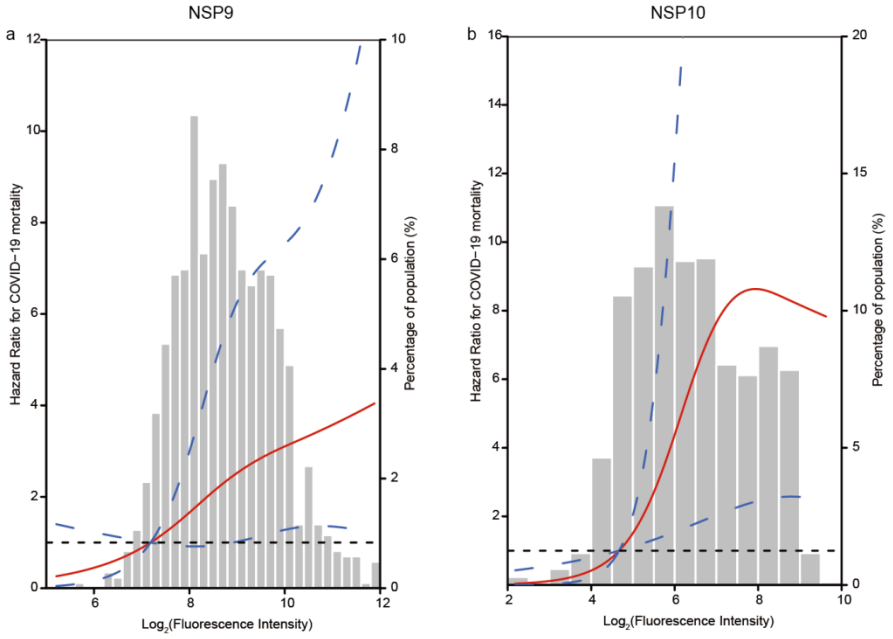
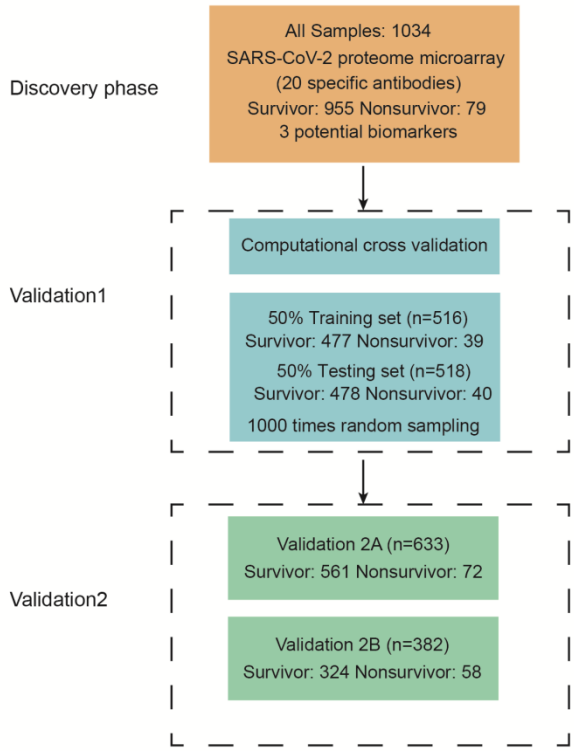


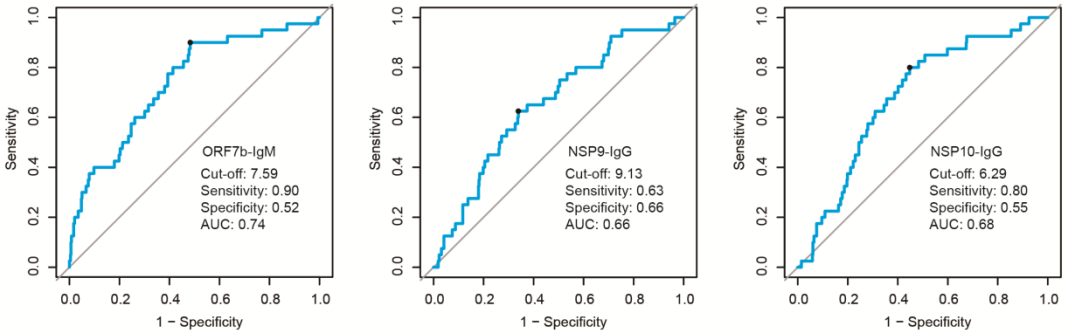
Figure 3. The associations of NSP9 and NSP10 IgG responses with risk of COVID-19 mortality. The lines represent adjusted hazard ratios based on restricted cubic splines for the levels of NSP9 and NSP10 IgG antibodies in Cox regression model. Knots were placed at the 5th, 50th, and 95th percentiles of the distribution of NSP9 and NSP10 IgG antibodies, and the reference value was set at the 10th percentile. Age, sex, hypertension, diabetes, lymphopenia, increased alanine aminotransferase and lactate dehydrogenase were used as adjustment factors.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

a



b Validation1



c Validation2A(n=633)

	Clinical outcome	
ORF7b-IgM	Survivor	Nonsurvivor
⩾7.59(Survivor)	387	18
<7.59(Nonsurvivor)	174	54

Accuracy=0.69

	Clinical outcome	
NSP9-IgG	Nonsurvivor	Survivor
⩾9.13(Nonsurvivor)	30	103
<9.13(Survivor)	42	458

Accuracy=0.77

	Clinical outcome	
NSP10-IgG	Nonsurvivor	Survivor
⩾6.29(Nonsurvivor)	49	129
<6.29(Survivor)	23	432

Accuracy=0.76

d Validation2B(n=382)

	Clinical outcome	
ORF7b-IgM	Survivor	Nonsurvivor
⩾7.59(Survivor)	217	21
<7.59(Nonsurvivor)	107	37

Accuracy=0.66

	Clinical outcome	
NSP9-IgG	Nonsurvivor	Survivor
⩾9.13(Nonsurvivor)	20	53
<9.13(Survivor)	38	271

Accuracy=0.76

	Clinical outcome	
NSP10-IgG	Nonsurvivor	Survivor
⩾6.29(Nonsurvivor)	30	59
<6.29(Survivor)	28	265

Accuracy=0.77

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Figure 4. Validation for ORF7b IgM, NSP9 IgG, and NSP10 IgG. (a) The workflow of validation. Firstly, three potential biomarkers were confirmed using 1,034 samples. Secondly, the three candidates were then computationally validated by 1,000 times random sampling. Finally, the accuracy of prognosis efficacy of the three biomarkers was evaluated in two other sample sets, *i. e.*, 633 samples and 382 samples. (b) The prognosis efficacy determined by computational cross-validation. The receiver operating characteristic curve was conducted for the prediction of COVID-19 survival and death, and 1,000 times computational cross-validations were conducted. For each cross-validation procedure, 477 survivors and 39 non-survivors were randomly selected as the training set. The rest of the samples were treated as the testing set (478 survivors and 40 non-survivors). The average cutoff values were shown. (c) The prognosis efficacy in Validation 2A. The second set of samples (n = 633) were analyzed to evaluate the prognosis efficacy of ORF7b IgM, NSP9 IgG, and NSP10 IgG. The accuracy was calculated. (d) The prognosis efficacy in Validation 2B. The third sample set (n = 388) were analyzed to evaluate the prognosis efficacy of ORF7b IgM, NSP9 IgG, and NSP10 IgG. The accuracy was calculated.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

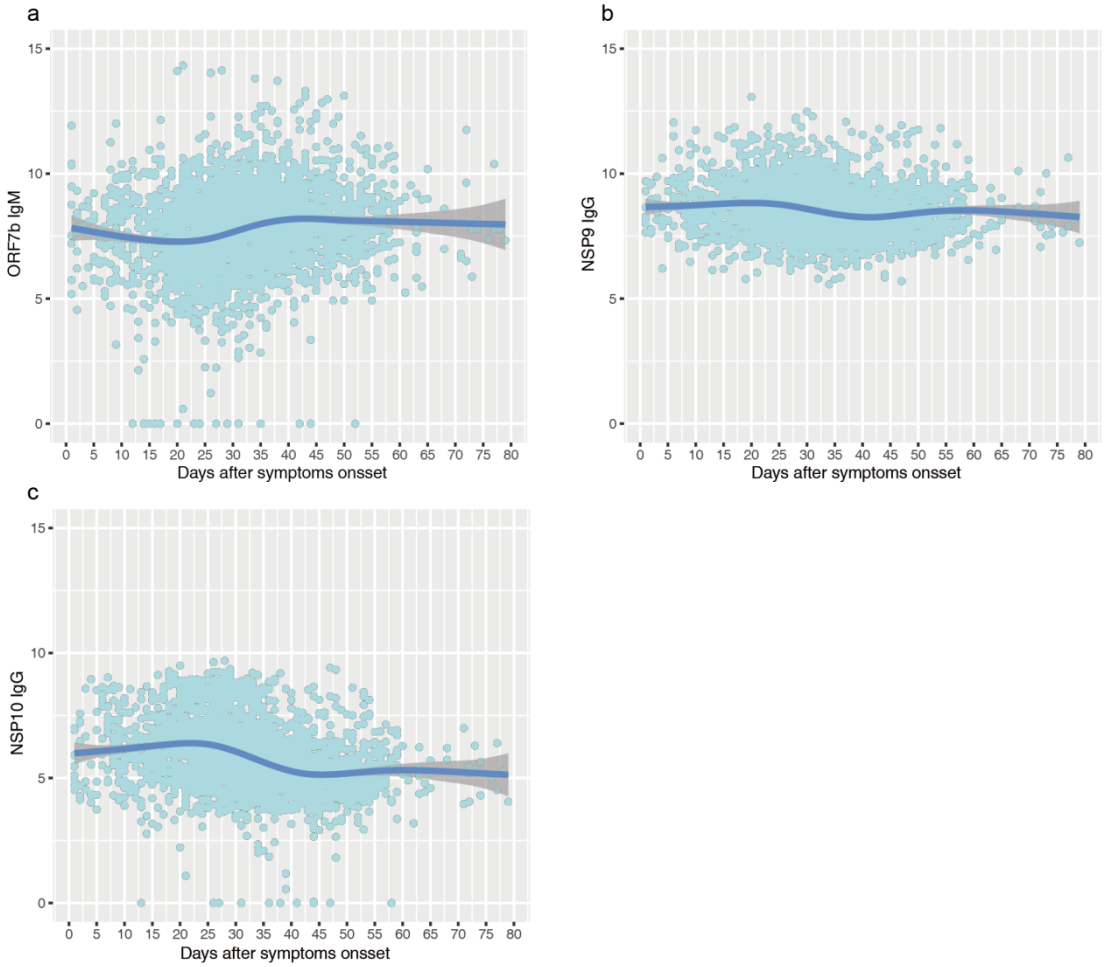


Figure 5. Dynamic changes of ORF7b IgM, NSP9 IgG, and NSP10 IgG. Dynamic changes of ORF7b IgM (a), NSP9 IgG (b), and NSP10 IgG (c) for 1034 COVID-19 patients (2,973 samples) with the days after symptoms onset. The lines show the mean value expected from a Loess regression model, the ribbons indicate the 95% confidence interval.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table S1. Comparison of SARS-CoV-2 specific IgM responses ([log2(FI)]) between survivors and nonsurivivors

Proteins	All	Survivors	Nonsurvivors	*p*
S1	12.1(11.1-13.1)	12.1(11.1-13.1)	12.1(10.1-13.1)	0.25
S2	9.1(8.4-9.8)	9.1(8.4-9.8)	8.9(8.4-9.6)	0.22
N	**7.8(6.9-8.5)**	**7.8(7.0-8.6)**	**7.3(5.8-8.3)**	**0.002**
N-Nter	10.4(9.6-11.0)	10.4(9.6-11.0)	10.1(9.0-11.1)	0.31
N-Cter	11.4(10.5-12.3)	11.4(10.6-12.3)	11.2(10.2-12.5)	0.54
E	7.5(6.6-8.4)	7.5(6.6-8.4)	7.5(6.4-8.3)	0.61
NSP1	8.2(7.5-8.9)	8.2(7.5-8.9)	8.3(7.6-8.8)	0.46
NSP2	8.2(7.4-9.0)	8.2(7.4-9.0)	8.1(7.3-8.9)	0.48
NSP4	8.4(7.7-9.2)	8.4(7.7-9.2)	8.5(7.7-9.2)	0.89
NSP5	6.8(6.0-7.8)	6.8(6.0-7.8)	6.7(5.6-7.5)	0.10
NSP7	9.6(8.8-10.4)	9.6(8.8-10.4)	9.4(8.5-10.5)	0.22
NSP8	9.0(8.0-10.0)	9.0(8.0-10.0)	9.0(7.7-10.0)	0.84
NSP9	8.8(8.2-9.3)	8.7(8.2-9.3)	8.9(8.4-9.3)	0.07
NSP10	**6.4(5.7-7.3)**	**6.4(5.7-7.3)**	**6.7(5.9-7.6)**	**0.03**
RdRp	8.7(8.0-9.4)	8.7(8.0-9.4)	8.6(8.2-9.4)	0.85
NSP14	8.7(7.9-9.4)	8.6(7.9-9.4)	8.7(7.9-9.5)	0.60
NSP15	8.3(7.6-9.1)	8.3(7.6-9.1)	8.2(7.6-9.1)	0.94
NSP16	8.8(8.1-9.6)	8.8(8.1-9.6)	8.9(7.9-9.6)	0.74
ORF3a	**5.3(2.9-6.8)**	**5.4(3.5-6.8)**	**2.7(0.0-4.9)**	**<0.001**
ORF3b	8.6(7.9-9.3)	8.6(7.9-9.2)	8.6(7.9-9.3)	0.94
ORF6	**7.0(6.0-8.2)**	**7.1(6.0-8.3)**	**6.6(5.5-7.6)**	**0.01**
ORF7b	**7.5(6.3-8.8)**	**7.7(6.5-8.9)**	**6.3(5.2-7.1)**	**<0.001**
ORF9b	7.7(7.0-8.4)	7.7(7.0-8.4)	7.9(7.2-8.5)	0.12

FI: Fluorescence Intensity.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table S2. Comparison of SARS-CoV-2 specific IgG responses ([log2(FI)]) between survivors and nonsurivivors

Proteins	All	Survivors	Nonsurvivors	*p*
S1	13.9(13.0-14.4)	13.9(13.0-14.4)	13.6(12.0-14.6)	0.30
S2	9.1(8.4-9.6)	9.1(8.4-9.6)	9.0(8.1-9.7)	0.34
N	**10.3(9.2-11.1)**	**10.3(9.2-11.2)**	**9.8(7.9-10.6)**	**<0.001**
N-Nter	13.2(12.3-13.8)	13.2(12.3-13.8)	13.1(11.6-13.7)	0.10
N-Cter	13.4(12.5-14.0)	13.4(12.6-14.0)	13.3(11.8-14.2)	0.68
E	**5.5(4.6-6.8)**	**5.5(4.6-6.8)**	**5.8(4.9-7.6)**	**0.04**
NSP1	**8.2(7.5-9.1)**	**8.1(7.4-9.0)**	**9.0(8.4-9.6)**	**<0.001**
NSP2	**6.6(5.6-7.8)**	**6.5(5.6-7.7)**	**7.1(6.0-8.2)**	**0.01**
NSP4	**7.9(7.4-8.7)**	**7.9(7.3-8.7)**	**8.2(7.8-9.4)**	**<0.001**
NSP5	**5.5(4.9-6.2)**	**5.5(4.9-6.2)**	**5.8(5.1-6.7)**	**0.01**
NSP7	**9.4(8.8-10.0)**	**9.4(8.8-10.0)**	**9.9(9.6-10.4)**	**<0.001**
NSP8	**7.8(6.8-9.0)**	**7.6(6.7-8.9)**	**8.8(8.0-9.2)**	**<0.001**
NSP9	**8.7(8.0-9.5)**	**8.7(8.0-9.5)**	**9.4(8.6-9.8)**	**<0.001**
NSP10	**6.3(5.3-7.6)**	**6.2(5.3-7.4)**	**7.2(6.5-8.0)**	**<0.001**
RdRp	**8.1(7.4-9.3)**	**8.0(7.4-9.2)**	**9.2(8.5-9.6)**	**<0.001**
NSP14	**7.4(6.7-8.4)**	**7.3(6.6-8.3)**	**8.3(7.7-9.1)**	**<0.001**
NSP15	**7.1(6.2-8.4)**	**7.1(6.1-8.3)**	**7.7(6.6-9.1)**	**0.02**
NSP16	**7.1(6.3-8.2)**	**7.0(6.3-8.2)**	**7.7(6.6-8.9)**	**0.004**
ORF3a	**5.2(4.0-6.6)**	**5.3(4.0-6.6)**	**4.6(3.4-5.7)**	**0.001**
ORF3b	**8.7(8.0-9.6)**	**8.6(8.0-9.6)**	**9.6(9.1-9.9)**	**<0.001**
ORF6	3.7(0.0-4.9)	3.7(0.0-4.9)	3.4(0.0-4.7)	0.30
ORF7b	**6.4(5.4-7.2)**	**6.4(5.5-7.2)**	**5.6(4.8-6.8)**	**<0.001**
ORF9b	**8.0(7.5-8.8)**	**8.0(7.4-8.7)**	**8.4(7.9-9.5)**	**<0.001**

FI: Fluorescence Intensity.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table S3. Hazard ratio (95%CI) for COVID-19 mortality according to tertiles of SARS-CoV-2 specific IgM response

Proteins	Tertile of proteins [log2(FI)]			*p* trend	FDR
	T1	T2	T3		
S1					
Model 1	1.00	0.54(0.31-0.94)	0.69(0.41-1.16)	0.14	0.36
Model 2	1.00	0.67(0.37-1.22)	0.87(0.50-1.52)	0.61	0.78
S2					
Model 1	1.00	0.82(0.48-1.39)	0.88(0.51-1.51)	0.62	0.80
Model 2	1.00	0.79(0.44-1.42)	0.96(0.54-1.72)	0.88	0.92
N					
Model 1	1.00	0.52(0.30-0.90)	0.55(0.33-0.94)	0.02	0.11
Model 2	1.00	0.68(0.38-1.23)	0.87(0.49-1.53)	0.53	0.73
N-Nter					
Model 1	1.00	0.41(0.22-0.74)	0.73(0.44-1.20)	0.23	0.49
Model 2	1.00	0.53(0.27-1.03)	0.93(0.55-1.58)	0.79	0.89
N-Cter					
Model 1	1.00	0.70(0.40-1.22)	0.94(0.56-1.59)	0.82	0.92
Model 2	1.00	0.71(0.39-1.29)	0.96(0.55-1.68)	0.90	0.94
E					
Model 1	1.00	0.77(0.44-1.35)	1.02(0.61-1.72)	0.94	0.95
Model 2	1.00	0.71(0.39-1.31)	1.20(0.69-2.09)	0.53	0.73
NSP1					
Model 1	1.00	1.17(0.66-2.05)	1.20(0.69-2.08)	0.53	0.73
Model 2	1.00	0.95(0.51-1.77)	1.27(0.71-2.28)	0.39	0.62
NSP2					
Model 1	1.00	0.78(0.46-1.33)	0.75(0.44-1.29)	0.29	0.57
Model 2	1.00	0.87(0.49-1.54)	0.84(0.47-1.48)	0.53	0.73
NSP4					
Model 1	1.00	1.01(0.59-1.74)	0.86(0.49-1.50)	0.58	0.77
Model 2	1.00	1.06(0.58-1.91)	1.07(0.59-1.94)	0.82	0.92
NSP5					
Model 1	1.00	0.67(0.40-1.14)	0.74(0.43-1.27)	0.23	0.49
Model 2	1.00	0.77(0.44-1.36)	0.88(0.49-1.58)	0.62	0.80
NSP7					
Model 1	1.00	0.69(0.41-1.17)	0.65(0.38-1.12)	0.11	0.32
Model 2	1.00	0.83(0.46-1.47)	0.82(0.46-1.45)	0.48	0.69
NSP8					
Model 1	1.00	1.11(0.65-1.89)	0.91(0.52-1.59)	0.73	0.85
Model 2	1.00	1.29(0.72-2.30)	1.07(0.59-1.96)	0.84	0.92
NSP9					
Model 1	1.00	1.54(0.87-2.72)	1.43(0.80-2.55)	0.26	0.51
Model 2	1.00	1.66(0.88-3.11)	1.61(0.87-3.00)	0.15	0.37
NSP10					
Model 1	1.00	1.06(0.58-1.91)	1.44(0.84-2.48)	0.16	0.41
Model 2	1.00	0.86(0.45-1.65)	1.42(0.79-2.52)	0.18	0.42
RdRp					
Model 1	1.00	0.97(0.56-1.68)	1.09(0.63-1.86)	0.77	0.88

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Model 2	1.00	1.00(0.55-1.79)	1.14(0.64-2.04)	0.66	0.82
NSP14					
Model 1	1.00	0.63(0.35-1.12)	0.97(0.59-1.62)	0.95	0.95
Model 2	1.00	0.64(0.35-1.20)	1.02(0.59-1.75)	0.91	0.93
NSP15					
Model 1	1.00	0.92(0.54-1.58)	0.85(0.49-1.45)	0.54	0.74
Model 2	1.00	0.92(0.52-1.65)	0.87(0.49-1.54)	0.62	0.80
NSP16					
Model 1	1.00	0.69(0.40-1.20)	0.78(0.46-1.31)	0.33	0.59
Model 2	1.00	0.68(0.38-1.23)	0.79(0.45-1.38)	0.39	0.62
ORF3a					
Model 1	1.00	0.45(0.26-0.78)	0.31(0.16-0.59)	<0.001	0.003
Model 2	1.00	0.58(0.32-1.05)	0.54(0.28-1.06)	0.03	0.18
ORF3b					
Model 1	1.00	0.98(0.57-1.68)	0.89(0.51-1.55)	0.68	0.83
Model 2	1.00	1.20(0.67-2.15)	0.96(0.53-1.73)	0.87	0.92
ORF6					
Model 1	1.00	1.00(0.60-1.66)	0.72(0.40-1.30)	0.32	0.58
Model 2	1.00	1.39(0.81-2.39)	1.13(0.59-2.15)	0.58	0.77
ORF9b					
Model 1	1.00	1.05(0.60-1.84)	1.05(0.61-1.82)	0.87	0.92
Model 2	1.00	0.92(0.50-1.70)	1.18(0.66-2.11)	0.54	0.73

Model 1: Adjusted for age and sex.

Model 2: Additional adjustment for hypertension, diabetes, lymphopenia, increased alanine aminotransferase, and increased lactate dehydrogenase.

FI: Fluorescence Intensity. False discovery rate (FDR) adjusted *p*-values.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table S4. Hazard ratio (95%CI) for COVID-19 mortality according to tertiles of SARS-CoV-2 specific IgM response after exclusion of patients who died within the first 7 days

Proteins	Tertile of proteins [log_2(FI)]			*p* trend
	T1	T2	T3	
N				
Model 1	1.00	0.30(0.13-0.70)	0.60(0.31-1.16)	0.08
Model 2	1.00	0.32(0.13-0.79)	0.80(0.40-1.60)	0.34
NSP10				
Model 1	1.00	1.00(0.45-2.19)	1.46(0.72-2.96)	0.25
Model 2	1.00	0.94(0.40-2.22)	1.54(0.73-3.24)	0.22
ORF3a				
Model 1	1.00	0.50(0.25-1.01)	0.38(0.17-0.87)	0.01
Model 2	1.00	0.54(0.25-1.16)	0.53(0.22-1.24)	0.08
ORF6				
Model 1	1.00	1.12(0.57-2.21)	1.03(0.49-2.16)	0.91
Model 2	1.00	1.67(0.81-3.47)	1.44(0.64-3.28)	0.32
ORF7b				
Model 1	1.00	0.65(0.34-1.24)	0.18(0.06-0.61)	0.002
Model 2	1.00	0.79(0.39-1.57)	0.26(0.08-0.88)	0.03

Model 1: Adjusted for age and sex.

Model 2: Additional adjustment for hypertension, diabetes, lymphopenia, increased alanine aminotransferase, and increased lactate dehydrogenase.

FI: Fluorescence Intensity.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table S5. Hazard ratio (95%CI) for COVID-19 mortality according to tertiles of SARS-CoV-2 specific IgG response

Proteins	Tertile of proteins [log2(FI)]			*p* trend	FDR
	T1	T2	T3		
S1					
Model 1	1	0.46(0.25-0.83)	0.61(0.37-1.00)	0.05	0.23
Model 2	1.00	0.52(0.27-0.99)	0.72(0.42-1.23)	0.25	0.50
S2					
Model 1	1	0.55(0.31-0.99)	0.81(0.49-1.35)	0.42	0.65
Model 2	1.00	0.74(0.40-1.37)	1.12(0.65-1.94)	0.69	0.84
N					
Model 1	1.00	0.63(0.38-1.05)	0.40(0.22-0.73)	0.002	0.02
Model 2	1.00	0.79(0.46-1.34)	0.73(0.39-1.37)	0.52	0.73
N-Nter					
Model 1	1.00	0.68(0.40-1.16)	0.59(0.34-1.00)	0.05	0.23
Model 2	1.00	0.91(0.51-1.62)	0.81(0.45-1.44)	0.46	0.68
N-Cter					
Model 1	1.00	0.79(0.45-1.39)	0.87(0.51-1.46)	0.60	0.77
Model 2	1.00	0.96(0.52-1.78)	1.06(0.61-1.84)	0.83	0.92
E					
Model 1	1.00	1.07(0.59-1.92)	1.25(0.72-2.20)	0.41	0.63
Model 2	1.00	1.11(0.59-2.09)	1.25(0.68-2.29)	0.56	0.76
NSP2					
Model 1	1.00	0.96(0.53-1.75)	1.30(0.75-2.26)	0.30	0.57
Model 2	1.00	0.76(0.39-1.45)	1.18(0.66-2.11)	0.64	0.80
NSP5					
Model 1	1.00	1.08(0.60-1.95)	1.48(0.85-2.57)	0.15	0.38
Model 2	1.00	1.15(0.60-2.22)	1.79(0.98-3.27)	0.07	0.25
NSP15					
Model 1	1.00	1.03(0.56-1.90)	1.40(0.80-2.45)	0.20	0.44
Model 2	1.00	0.85(0.44-1.65)	1.23(0.68-2.22)	0.49	0.71
NSP16					
Model 1	1.00	0.91(0.49-1.70)	1.52(0.87-2.64)	0.09	0.30
Model 2	1.00	0.71(0.36-1.39)	1.40(0.78-2.50)	0.28	0.55
ORF-3a					
Model 1	1.00	1.03(0.63-1.68)	0.50(0.27-0.92)	0.04	0.20
Model 2	1.00	1.35(0.79-2.29)	0.69(0.35-1.33)	0.53	0.73
ORF6					
Model 1	1.00	0.82(0.48-1.40)	0.75(0.44-1.30)	0.30	0.57
Model 2	1.00	0.93(0.52-1.65)	0.94(0.53-1.68)	0.83	0.92
ORF-7b					
Model 1	1.00	0.60(0.35-1.03)	0.45(0.26-0.81)	0.005	0.04
Model 2	1.00	0.79(0.45-1.39)	0.71(0.39-1.30)	0.20	0.44

Model 1: Adjusted for age and sex.

Model 2: Additional adjustment for hypertension, diabetes, lymphopenia, increased alanine aminotransferase, and increased lactate dehydrogenase.

FI: Fluorescence Intensity.

FDR: False discovery rate adjusted *p*-values.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table S6. Hazard ratio (95%CI) for COVID-19 mortality according to tertiles of anti-SARS-CoV-2 specific IgG response after exclusion of patients who died within the initial 7 days after hospitalization

Proteins	Tertile of proteins [log_2(FI)]			*p* trend
	T1	T2	T3	
N				
Model 1	1.00	0.60(0.31-1.19)	0.52(0.25-1.10)	0.067
Model 2	1.00	0.75(0.36-1.54)	0.87(0.39-1.92)	0.62
E				
Model 1	1.00	1.09(0.50-2.40)	1.39(0.67-2.91)	0.35
Model 2	1.00	1.22(0.52-2.86)	1.65(0.74-3.68)	0.21
NSP1				
Model 1	1.00	2.36(0.92-6.03)	2.40(0.98-5.88)	0.08
Model 2	1.00	2.59(0.93-7.26)	2.47(0.93-6.56)	0.106
NSP2				
Model 1	1.00	0.91(0.42-1.96)	1.13(0.55-2.33)	0.68
Model 2	1.00	0.75(0.33-1.73)	1.17(0.55-2.47)	0.57
NSP4				
Model 1	1.00	1.92(0.79-4.65)	2.22(0.95-5.20)	0.08
Model 2	1.00	3.54(1.18-10.58)	3.87(1.32-11.35)	0.02
NSP5				
Model 1	1.00	0.93(0.43-1.98)	1.29(0.64-2.60)	0.44
Model 2	1.00	1.33(0.58-3.06)	1.72(0.80-3.70)	0.16
NSP7				
Model 1	1.00	2.21(0.80-6.07)	3.05(1.17-7.94)	0.02
Model 2	1.00	2.43(0.79-7.53)	3.19(1.10-9.29)	0.03
NSP8				
Model 1	1.00	2.07(0.81-5.32)	2.50(1.02-6.11)	0.048
Model 2	1.00	2.19(0.78-6.14)	2.51(0.95-6.66)	0.074
NSP9				
Model 1	1.00	2.36(0.87-6.41)	3.28(1.25-8.59)	0.01
Model 2	1.00	2.28(0.75-6.92)	3.61(1.24-10.51)	0.01
NSP10				
Model 1	1.00	3.72(1.08-12.77)	5.38(1.63-17.80)	0.003
Model 2	1.00	4.76(1.08-20.94)	7.45(1.76-31.59)	0.002
RdRp				
Model 1	1.00	3.15(1.06-9.40)	3.61(1.25-10.42)	0.02
Model 2	1.00	4.15(1.20-14.29)	3.92(1.17-13.16)	0.049
NSP14				
Model 1	1.00	1.73(0.61-4.93)	3.20(1.23-8.35)	0.007
Model 2	1.00	1.67(0.53-5.30)	3.27(1.13-9.47)	0.01
NSP15				
Model 1	1.00	1.01(0.46-2.21)	1.19(0.57-2.49)	0.60
Model 2	1.00	0.99(0.43-2.30)	1.20(0.55-2.61)	0.61
NSP16				
Model 1	1.00	0.94(0.42-2.11)	1.44(0.70-2.99)	0.26
Model 2	1.00	0.75(0.31-1.80)	1.45(0.68-3.11)	0.22

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

ORF3a				
Model 1	1.00	1.05(0.56-1.98)	0.50(0.22-1.13)	0.13
Model 2	1.00	1.28(0.65-2.53)	0.53(0.22-1.28)	0.24
ORF3b				
Model 1	1.00	1.21(0.45-3.26)	2.65(1.16-6.03)	0.008
Model 2	1.00	1.35(0.47-3.93)	2.64(1.09-6.43)	0.02
ORF7b				
Model 1	1.00	0.75(0.37-1.51)	0.72(0.35-1.45)	0.32
Model 2	1.00	0.94(0.44-1.99)	0.95(0.45-2.01)	0.89
ORF9b				
Model 1	1.00	1.95(0.83-4.57)	2.13(0.95-4.78)	0.08
Model 2	1.00	2.13(0.86-5.27)	2.32(0.97-5.53)	0.070

Model 1: Adjusted for age and sex.

Model 2: Additional adjustment for hypertension, diabetes, lymphopenia, increased alanine aminotransferase, and increased lactate dehydrogenase.

FI: Fluorescence Intensity.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table S7. Factor loadings of 20 proteins of anti-SARS-CoV-2 specific IgG responses among the study participants

Variables	PC1	PC2	PC3	PC4
S1	0.26	**0.87**	-0.15	-0.06
S2	0.36	0.59	-0.05	0.27
N	0.15	**0.87**	0.07	-0.07
N-Nter	0.26	**0.90**	-0.08	-0.07
N-Cter	0.39	**0.83**	-0.12	-0.08
E	0.67	-0.07	0.45	-0.29
NSP1	**0.87**	-0.13	-0.13	0.09
NSP2	**0.78**	-0.04	0.27	-0.17
NSP4	**0.87**	-0.10	-0.06	0.11
NSP5	0.65	0.01	0.49	-0.12
NSP7	**0.78**	-0.05	-0.18	0.14
NSP8	**0.79**	-0.16	-0.21	0.13
NSP9	**0.72**	-0.13	-0.22	0.25
NSP10	**0.77**	-0.19	-0.28	0.19
RdRp	**0.81**	-0.12	-0.21	0.00
NSP14	**0.89**	-0.05	0.07	-0.20
NSP15	**0.78**	-0.09	0.15	-0.17
NSP16	**0.81**	-0.04	0.29	-0.21
ORF3a	-0.20	0.29	0.50	0.47
ORF3b	**0.85**	-0.13	-0.13	0.09
ORF6	0.17	0.04	0.23	0.67
ORF7b	0.18	-0.02	**0.72**	0.19
ORF9b	**0.78**	-0.05	-0.10	0.03
Eigen values	9.95	3.609	1.79	1.198
Total variance (%)	43.263	15.691	7.784	5.207
Cumulative variance (%)	43.263	58.954	66.737	71.945

Bold values denote factor loading > 0.7 are deemed to be statistically significant.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table S8. Correlations between the levels of anti-SARS-CoV-2 specific IgM/IgG responses and other laboratory biomarkers related with severity factors

	PCT	CRP	LYMPH	LDH	DD	lL-2R	IL-6
ORF7b_IgM							
r_s	-0.19**	-0.22**	0.21**	-0.26**	-0.35**	-0.27**	-0.17**
n	542	718	958	940	607	414	416
NSP1_IgG							
r_s	0.19**	0.21**	-0.16**	0.17**	0.31**	0.18**	0.09
n	541	718	958	940	607	413	416
NSP4_IgG							
r_s	0.09*	0.14**	-0.09**	0.10**	0.21**	0.10*	0.02
n	542	719	959	941	608	414	417
NSP7_IgG							
r_s	0.19**	0.22**	-0.17**	0.19**	0.31**	0.14**	0.08
n	542	719	959	941	608	414	417
NSP8_IgG							
r_s	0.12**	0.19**	-0.15**	0.16**	0.31**	0.11*	0.12*
n	542	719	959	941	608	414	417
NSP9_IgG							
r_s	0.12**	0.17**	-0.09**	0.12**	0.17**	0.07	0.07
n	541	717	957	939	606	413	416
NSP10_IgG							
r_s	0.12**	0.21**	-0.15**	0.15**	0.31**	0.16**	0.13**
n	541	718	958	940	607	413	416
RdRp_IgG							
r_s	0.17**	0.19**	-0.15**	0.14**	0.31**	0.13**	0.11*
n	542	718	958	940	608	414	417
NSP14_IgG							
r_s	0.15**	0.17**	-0.15**	0.16**	0.27**	0.17**	0.11*
n	541	718	958	940	607	413	416
ORF3b_IgG							
r_s	0.15**	0.18**	-0.14**	0.16**	0.29**	0.12*	0.06
n	541	717	957	939	606	413	415
ORF9b_IgG							
r_s	0.12**	0.12**	-0.07*	0.11**	0.19**	0.04	0.01
n	542	719	959	941	608	414	417

Spearman's rank correlation coefficients were shown in the table. *p <0.05, **p <0.01. PCT: procalcitonin; CRP: C-reactive protein; LYMPH: lymphocyte count; LDH: lactate dehydrogenase; DD: D-dimer; IL-2R: interleukin-2 receptor; IL-6: interleukin-6.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table S9. Frequently mutated amino acids in ORF7b of SARS-CoV-2

NO	aa	variation(%)				Total variation(%)
		variation1	variation2	variation3	variation4	
1	M	M→-(23%)				23%
2	I	I→-(22%)	I→X(2%)			24%
3	E	E→-(22%)	E→X(3%)			25%
4	L	L→-(21%)	L→X(7%)			28%
5	S	S→-(22%)	S→X(7%)			29%
6	L	L→-(20%)	L→X(9%)			29%
7	I	I→-(20%)	I→X(10%)	I→T(1%)		31%
8	D	D→-(19%)	D→X(13%)	D→N(1%)	D→Y(1%)	34%
9	F	F→-(18%)	F→X(16%)			34%
10	Y	Y→X(24%)	Y→-(18%)			42%
11	L	L→X(27%)	L→-(17%)			44%
12	C	C→X(33%)	C→-(17%)			50%
13	F	F→X(34%)	F→-(17%)			51%
14	L	L→X(33%)	L→-(12%)			45%
15	A	A→X(28%)	A→-(8%)	A→T(1%)		37%
16	F	F→X(30%)	F→-(5%)	F→I(1%)		36%
17	L	L→X(17%)	L→-(3%)			20%
18	L	L→X(19%)	L→-(1%)			20%
19	F	F→X(3%)	F→L(1%)			4%
20	L	L→X(2%)				2%
21	V	V→X(3%)				3%
22	L	L→X(3%)				3%
23	I	I→X(2%)				2%
24	M	M→X(2%)				2%
25	L	L→X(2%)	L→F(1%)			3%
26	I	I→X(1%)	I→T(1%)			2%
27	I	I→X(1%)				1%
28	F	F→X(3%)	F→Y(1%)			4%
29	W	W→X(1%)				1%
30	F	F→L(2%)	F→X(1%)			3%
31	S	S→L(3%)	S→X(1%)			4%
32	L	L→F(1%)	L→X(1%)			2%
33	E					0%
34	L					0%
35	Q	Q→X(1%)				1%
36	D					0%
37	H					0%
38	N	N→X(1%)				1%
39	E	E→X(2%)				2%
40	T	T→I(1%)				1%
41	C	C→X(2%)	C→F(1%)	C→S(1%)		4%
42	H	H→Y(1%)	H→X(1%)			2%
43	A	A→X(2%)	A→T(1%)	A→V(1%)		4%

100 SARS-CoV-2 strains were included and analyzed with BLAST. Data were obtained in the National Center for Biotechnology Information (NCBI), https://www.ncbi.nlm.nih.gov/.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

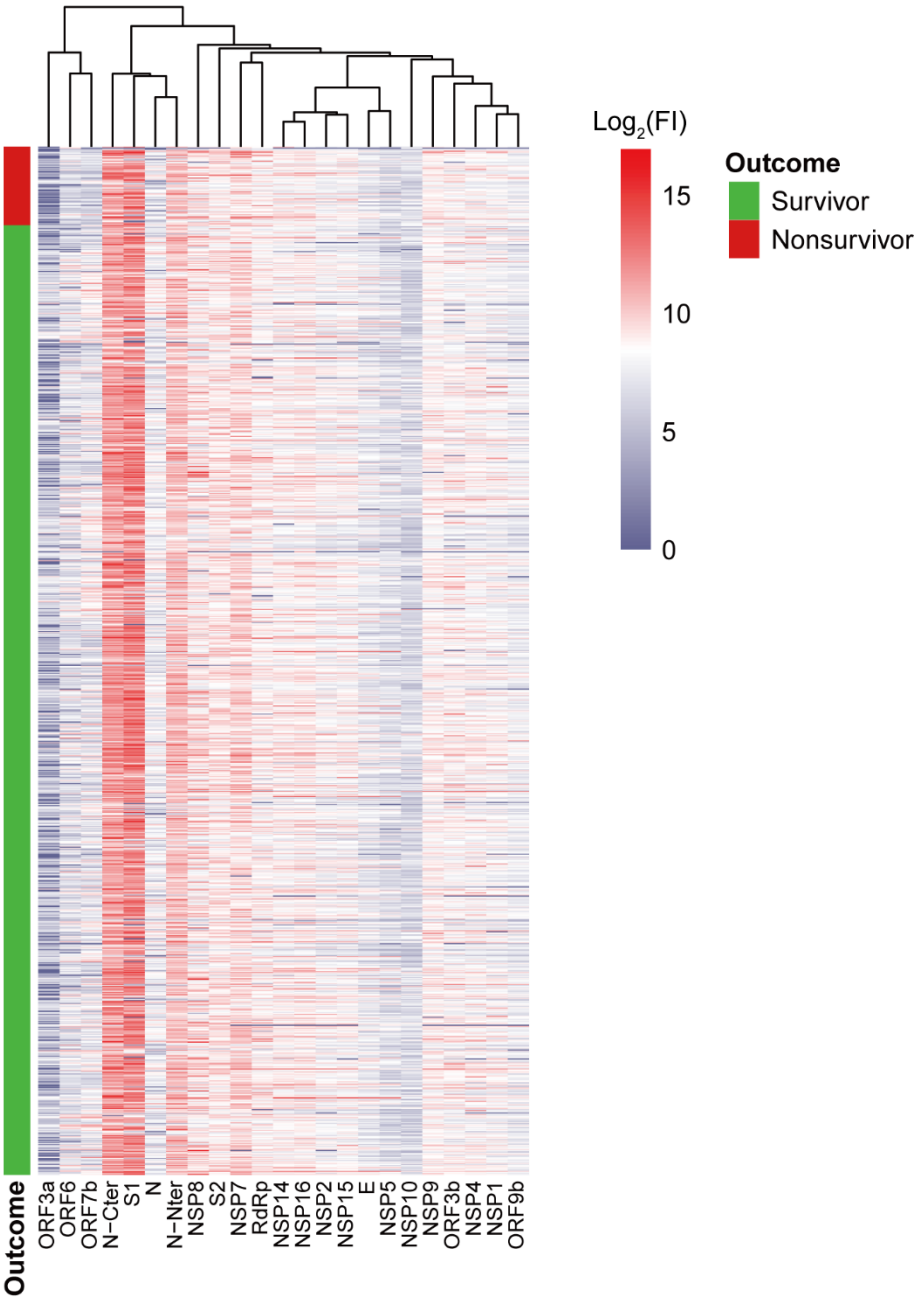


Figure S1. The overall SARS-CoV-2 specific IgM profiles. Heatmap of IgM antibody responses of 955 survivors and 79 nonsurvivors. Each square indicates the IgM antibody response against the protein (column) in the serum (row). The proteins were arranged with clustering. FI: Fluorescence Intensity.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20228890; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

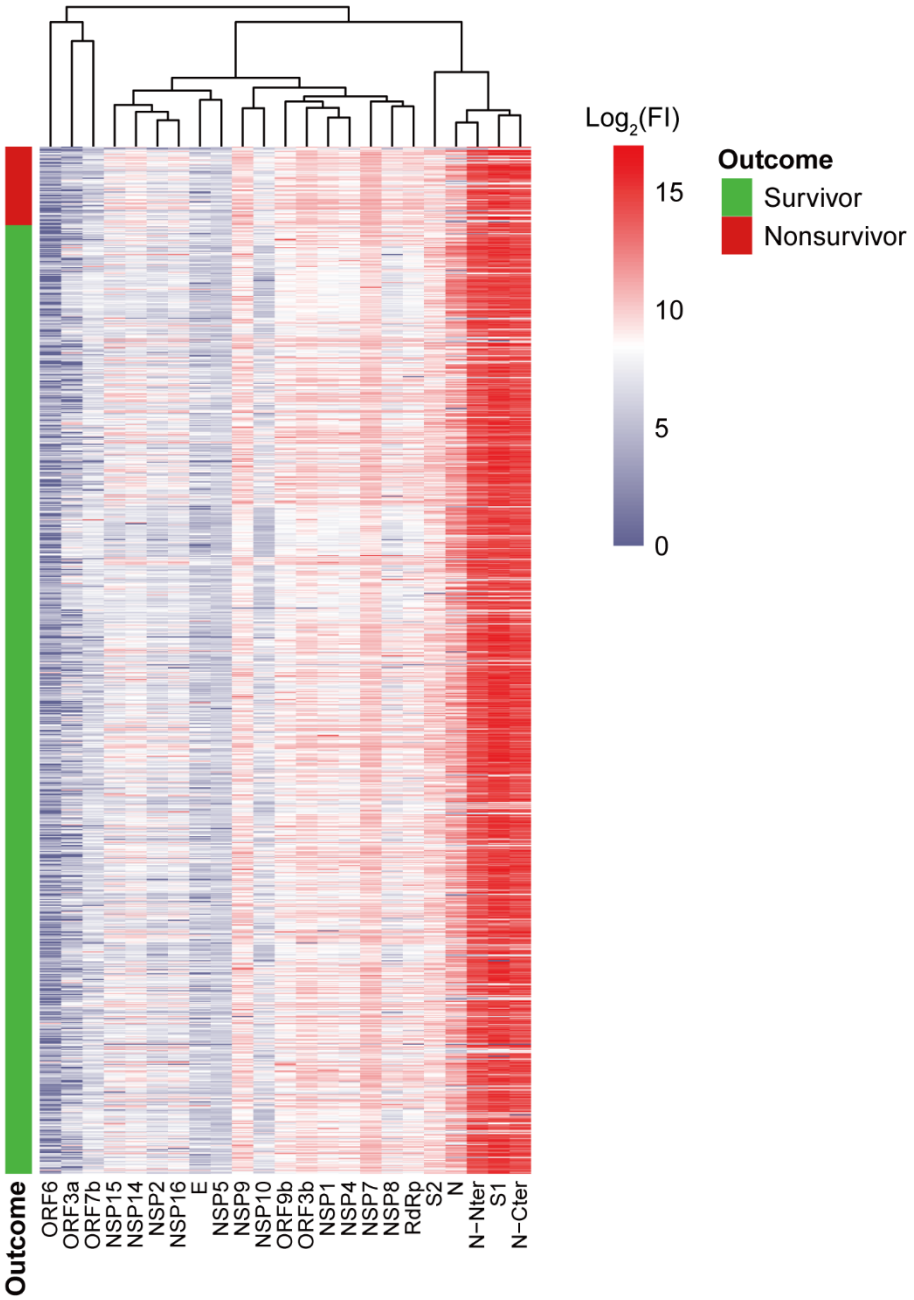


Figure S2. The overall SARS-CoV-2 specific IgG profiles. Heatmap of IgG antibody responses of 955 survivors and 79 nonsurvivors. Each square indicates the IgM antibody response against the protein (column) in the serum (row). The proteins were arranged with clustering. FI: Fluorescence Intensity.