

1 **COVID-ONE-humoral immune: The One-stop Database for COVID-19-specific Antibody**

2 **Responses and Clinical Parameters**

3 Zhaowei Xu ^{1,2, #,a}, Yang Li^{1, #,b}, Qing Lei^{3, #,c}, Likun Huang^{5, #,d}, Dan-yun Lai^{1, #,e}, Shu-juan

4 Guo^{1, #,f}, He-wei Jiang^{1,g}, Hongyan Hou^{4,h}, Yun-xiao Zheng^{1,i}, Xue-ning Wang^{1,j}, Jiaoxiang

5 Wu^{6,k}, Ming-liang Ma^{1,l}, Bo Zhang^{4,m}, Hong Chen^{1,n}, Caizheng Yu^{7,o}, Jun-biao Xue^{1,p},

6 Hai-nan Zhang^{1,q}, Huan Qi^{1,r}, Siqi Yu^{2,s}, Mingxi Lin^{2,t}, Yandi Zhang^{3,u}, Xiaosong Lin^{3,v},

7 Zongjie Yao^{3,w}, Huiming Sheng^{6,x}, Ziyong Sun^{3,y}, Feng Wang^{4, *,z}, Xionglin Fan^{3, *,aa},

8 Sheng-ce Tao^{1,*,ab}

9

10 ¹Shanghai Center for Systems Biomedicine, Key Laboratory of Systems Biomedicine
11 (Ministry of Education), Shanghai Jiao Tong University, Shanghai 200240, China

12 ²Key Laboratory of Gastrointestinal Cancer (Fujian Medical University), Ministry of
13 Education, School of Basic Medical Sciences, Fujian Medical University, Fuzhou 350122,
14 China

15 ³Department of Pathogen Biology, School of Basic Medicine, Tongji Medical College,
16 Huazhong University of Science and Technology, Wuhan 430074, China

17 ⁴Department of Clinical Laboratory, Tongji Hospital, Tongji Medical College, Huazhong
18 University of Science and Technology, Wuhan 430074, China

19 ⁵Fujian Key Laboratory of Crop Breeding by Design, Key Laboratory of Genetics,
20 Breeding and Multiple Utilization of Crops, Ministry of Education, Fujian Agriculture and
21 Forestry University, Fuzhou 350028, Fujian

22 ⁶Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025,
23 China

24 ⁷Department of Public Health, Tongji Hospital, Tongji Medical College, Huazhong
25 University of Science and Technology, Wuhan 430074, China

26 #These authors contribute equally to this work.

27 *Corresponding Authors, e-mail: taosc@sjtu.edu.cn (S.-C. Tao), xlfan@hust.edu.cn (X.-L.

28 Fan), fengwang@tjh.tjmu.edu.cn (F. Wang)

29

30 **Running title:** *Xu ZW et al / Database for COVID-19 Antibody Immune*

31 ^aORCID : 0000-0002-9134-1892

32 ^bORCID : 0000-0002-3182-6169

33 ^cORCID : 0000-0002-6679-1752

34 ^dORCID : 0000-0002-3377-8340

35 ^eORCID : 0000-0001-8719-6947

36 ^fORCID : 0000-0002-4252-0608

37 ^gORCID : 0000-0001-8700-7042

38 ^hORCID : 0000-0003-2337-0509

39 ⁱORCID : 0000-0001-8461-8884

40 ^jORCID : 0000-0001-8550-7166

41 ^kORCID : 0000-0002-2309-4164

42 ^lORCID : 0000-0002-0045-4876

43 ^mORCID : 0000-0002-1441-3100

44 ⁿORCID : 0000-0002-5412-7467

45 ^oORCID : 0000-0003-4583-9339

46 ^pORCID : 0000-0003-2029-7430

47 ^qORCID : 0000-0002-1058-0462

48 ^rORCID : 0000-0002-3800-0140

49 ^sORCID : 0000-0001-9373-4596

50 ^tORCID : 0000-0002-9745-9488

51 ^uORCID : 0000-0001-6268-5461

52 ^vORCID : 0000-0002-1792-9639

53 ^wORCID : 0000-0002-4383-1862

54 ^xORCID : 0000-0001-9382-3687

55 ^yORCID : 0000-0002-6443-9755

56 ^zORCID : 0000-0001-6324-9135

57 ^{aa}ORCID : 0000-0001-9754-372X

58 ^{ab}ORCID : 0000-0002-9210-1823

59

60 The manuscript includes 5011 words, 4 figures, 4 tables, 2 supplementary figures and 2
61 supplementary datasets. All of the 30 references in the manuscript are references from
62 2020. And there are 106 characters in article title, 42 characters in running title, 5
63 keywords, 180 words in abstract.

64

65

66

67

68

69 **Abstract**

70 Coronavirus disease 2019 (COVID-19), which is caused by SARS-CoV-2, varies with
71 regard to symptoms and mortality rates among populations. Humoral immunity plays
72 critical roles in SARS-CoV-2 infection and recovery from COVID-19. However, differences
73 in immune responses and clinical features among COVID-19 patients remain largely
74 unknown. Here, we report a database for COVID-19-specific IgG/IgM immune responses
75 and clinical parameters (COVID-ONE humoral immune). COVID-ONE humoral immunity

76 is based on a dataset that contains the IgG/IgM responses to 21 of 28 known
77 SARS-CoV-2 proteins and 197 spike protein peptides against 2,360 COVID-19 samples
78 collected from 783 patients. In addition, 96 clinical parameters for the 2,360 samples and
79 information for the 783 patients are integrated into the database. Furthermore,
80 COVID-ONE humoral immune provides a dashboard for defining samples and a one-click
81 analysis pipeline for a single group or paired groups. A set of samples of interest is easily
82 defined by adjusting the scale bars of a variety of parameters. After the “START” button is
83 clicked, one can readily obtain a comprehensive analysis report for further interpretation.
84 COVID-ONE-humoral immune is freely available at www.COVID-ONE.cn.

85

86 **KEYWORDS:** SARS-CoV-2; Protein microarray; Humoral immunity; One-stop tool; Shiny

87 Introduction

88 COVID-19 is an unprecedented global threat caused by severe acute respiratory
89 syndrome coronavirus 2 (SARS-CoV-2), which has already caused 188,843,580
90 infections and claimed 4,065,400 lives as of July 16, 2021
91 (<https://coronavirus.jhu.edu/map.html>) [1]. There is still no effective medicine [2, 3] for
92 treating COVID-19.

93 Most patients recover via their own immunity, including SARS-CoV-2-specific IgG
94 responses, especially neutralizing antibodies [4-6]. Overall, it is of great interest to
95 decipher SARS-CoV-2-specific IgG and IgM responses at a systems level and to correlate
96 responses to clinical parameters.

97 To understand how the human immune system responds to SARS-CoV-2, we
98 constructed a SARS-CoV-2 proteome microarray containing 18 of the 28 predicted
99 proteins and applied it to characterize IgG and IgM antibodies in the sera of 29
100 convalescent patients [7]. Recently, we upgraded the SARS-CoV-2 protein microarray,
101 and the new microarray contains 21 predicted SARS-CoV-2 proteins and 197 spike
102 protein peptides (with full coverage of spike) [8]. Using this microarray, we screened 2,360
103 serum samples from 783 COVID-19 patients, covering mild, severe and critical cases.
104 Thus, we compiled a dataset with comprehensive information on SARS-CoV-2-specific
105 humoral responses and rich in clinical parameters.

106 To share the dataset efficiently, in addition to the related research that we have already
107 published [9-13], we built a database for COVID-19-specific humoral immune responses
108 and clinical parameters, namely, COVID-ONE-humoral immune (www.covid-one.cn),
109 using Shiny. This database contains a comprehensive dataset of IgG and IgM responses
110 to the 21 predicted SARS-CoV-2 proteins and 197 spike protein peptides from a cohort of
111 783 COVID-19 patients. To bolster clinical relevance, 96 clinical parameters and basic

112 patient information were also included. COVID-ONE humoral immunity provides search,
113 data analysis, and visualization functions. In particular, COVID-ONE-humoral immune
114 integrates antibody response landscape analysis, correlation analysis, machine learning,
115 *etc.* In the data analysis module, users can easily define sample groups of interest by
116 adjusting scale bars, and the sample groups can be either one group or paired groups.
117 In-depth analysis is achieved by clicking a single button; optionally, the results can be
118 saved and downloaded as an independent package for further analysis.

119 To our knowledge, COVID-ONE humoral immune is the first database for
120 SARS-CoV-2-specific humoral immune responses. We believe that COVID-19 humoral
121 immune will be of broad interest and will facilitate understanding of immune responses in
122 COVID-19 to combat the pandemic.

123

124 **Materials and methods**

125 **Patients and samples.** All 783 COVID-19 cases were laboratory confirmed; the patients
126 were hospitalized at Tongji Hospital from 25 January 2020 to 28 April 2020. The criteria
127 for defining severity, *i.e.*, mild, severe and critically severe, referenced the Diagnosis and
128 Treatment Protocol for Novel Coronavirus Pneumonia (Trial Version 7), as released by
129 the National Health Commission & State Administration of Traditional Chinese Medicine.
130 For many of the patients, sera were collected during hospitalization at several time points.
131 Negative reference samples were obtained from the National Institutes for Food and Drug
132 Control. All serum samples were stored at -80°C until use.

133

134 **Peptide preparation.** In this study, the SARS-CoV-2 spike protein (1,273 aa) was divided
135 into 211 peptides of 12 aa, with 6 aa overlapping between adjacent peptides. After
136 cysteine was added to the N-terminus, these peptides were synthesized by GL Biochem,

137 Ltd. (Shanghai, China) and conjugated to BSA using Sulfo-SMCC (Thermo Fisher
138 Scientific, MA, USA). Briefly, BSA was activated by Sulfo-SMCC at a molar ratio of 1:30
139 and dialyzed against PBS buffer. A total of 197 soluble peptides were individually
140 conjugated with activated BSA in a w/w ratio of 1:1 and incubated for 2 h at room
141 temperature. Free peptides were removed by dialysis with a pore size of 10 kD. The
142 conjugates were assessed by SDS-PAGE.

143

144 **Protein preparation.** SARS-CoV-2 protein sequences were downloaded from GenBank
145 (Accession number: MN908947.3) and converted to *Escherichia coli* codon-optimized
146 gene sequences. The optimized genes were synthesized and cloned into pET32a or
147 pGEX-4T-1 by Sangon Biotech (Shanghai, China). Recombinant proteins were expressed
148 in *E. coli* BL21 by growing cells in 200 mL LB medium to OD₆₀₀ ~ 0.6 at 37 °C followed by
149 induction with 0.2 mM isopropyl-β-d-thiogalactoside (IPTG) overnight at 16 °C. For the
150 purification of 6xHis-tagged proteins, cell pellets were re-suspended in lysis buffer
151 containing 50 mM Tris-HCl, 500 mM NaCl, and 20 mM imidazole (pH 8.0) and lysed
152 using a high-pressure cell cracker (Union Biotech, Shanghai, China). After centrifugation
153 at 12,000 x g for 20 min at 4 °C, the lysates were incubated with Ni²⁺ Sepharose beads
154 (Senhui Microsphere Technology, Suzhou, China) for 1 h at 4 °C, washed 3 times with
155 lysis buffer and eluted with buffer containing 50 mM Tris-HCl, 500 mM NaCl, and
156 300 mM imidazole (pH 8.0). For the purification of GST-tagged proteins, cells were
157 harvested and lysed by a high-pressure cell cracker in lysis buffer containing 50 mM
158 Tris-HCl, 500 mM NaCl, and 1 mM DTT at pH 8.0. After centrifugation, the supernatant
159 was incubated with GST-Sepharose beads (Senhui Microsphere Technology, Suzhou,
160 China). The target proteins were washed with lysis buffer and eluted with 50 mM Tris-HCl,
161 500 mM NaCl, 1 mM DTT, and 40 mM glutathione at pH 8.0. The purified proteins were

162 quality checked by SDS-PAGE and Coomassie blue staining and stored at -80°C until
163 use.

164

165 **Protein microarray fabrication.** The SARS-CoV-2 proteome microarray used in this study
166 is an updated version of the original microarray[7], which contains 18 of the 28 predicted
167 SARS-CoV-2 proteins. Three more proteins, *i.e.*, ORF3a, ORF3b, and ORF7b, and 197
168 spike protein peptides were added to the updated version. Therefore, the protein
169 microarray used in this study contained 21/28 SARS-CoV-2 proteins and 197 peptides,
170 with full coverage of the spike protein. The proteins and spike protein peptides, along with
171 BSA and anti-human IgG/IgM (Jackson ImmunoResearch Laboratories, USA), were used
172 as negative and positive controls, respectively, and printed in triplicate on PATH substrate
173 slides (Grace Bio-Labs, Oregon, USA) to generate identical arrays in a 2 x 7 subarray
174 format using a Super Marathon printer (Arrayjet, UK). Anti-His (Millipore, USA), anti-GST
175 (Sigma, USA), and anti-BSA (Sangon Biotech, China) antibodies were used for quality
176 control of the SARS-CoV-2 proteome microarray. The protein microarrays were stored at
177 -80°C until use.

178

179 **Microarray-based serum analysis.** A 14-chamber rubber gasket was mounted onto each
180 slide to create individual chambers for 14 identical subarrays. The microarray was used
181 for serum profiling as described previously, with minor modifications[14]. Briefly, arrays
182 stored at -80°C were warmed to room temperature and then incubated in blocking buffer
183 (3% BSA in 1×PBS buffer with 0.1% Tween 20) for 3 h. A total of 200 µL of diluted serum
184 or antibodies was incubated with each subarray for 2 h. For most samples, sera were
185 diluted to 1:200; for the competition experiment, free peptides were added at a
186 concentration of 0.25 mg/mL. For the enriched antibodies, 0.1-0.5 µg antibodies were

187 included in 200 μ L incubation buffer. The arrays were washed with 1 \times PBST, and the
188 bound antibodies were monitored by incubating with Cy3-conjugated goat anti-human IgG
189 and Alexa Fluor 647-conjugated donkey anti-human IgM (Jackson ImmunoResearch, PA,
190 USA) diluted 1:1,000 in 1 \times PBST at room temperature for 1 h. The microarrays were then
191 washed with 1 \times PBST, dried by centrifugation at room temperature and scanned using a
192 LuxScan 10K-A (CapitalBio Corporation, Beijing, China) with the parameters set as 95%
193 laser power/PMT 550 and 95% laser power/PMT 480 for IgM and IgG, respectively. The
194 fluorescence intensity was extracted with GenePix Pro 6.0 software (Molecular Devices,
195 CA, USA).

196

197 **Protein microarray data analysis.** IgG and IgM signal intensities were defined as
198 foreground medians (F) subtracted by background medians (B) for each spot, and the
199 signal intensity of a protein was averaged for triplicate spots. Block #14 of each slide was
200 incubated with SARS-CoV-2 immunopositive serum as the positive control. Data
201 normalization between slides was performed by a linear method according to the positive
202 control; specifically, a normalization factor for each slide was calculated by linear
203 regression according to the positive control. To reduce error among microarrays, the
204 signals of all the proteins from each slide were divided by its normalization factor.

205

206 **Quantification and statistical analysis.** To calculate the rate of antibody response for each
207 protein, the mean plus 2 times the standard deviation (SD) of the control serum was set as
208 the cut-off. R was used for most data analysis and drawing, *i.e.*, Pearson correlation
209 coefficient, ROC, T-test, cluster analysis and machine learning.

210

211 **Data collection.** Specific IgG/IgM immune response data were obtained by

212 microarray-based serum analysis. Blood parameters were collected from Tongji Hospital,
213 Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.

214

215 **Database architecture and web interface.** COVID-ONE-humoral immune is a Shiny-based
216 (1.5.0) database. Shinydashboard (0.7.1) and Shiny BS (0.61) were used to shape the UI,
217 and the package DT (0.15) was used to format data tables. For data analysis, dplyr (1.0.2),
218 tidyverse (1.3.0), randomForest (4.6-14), pROC (1.16.2), and umap (0.2.6.0) were
219 integrated into Shiny. Pheatmap (1.0.12) and ggplot2 (3.3.2) carry out plotting. For the
220 basic environment, the operation system is Ubuntu 20.04 LTS, and the version of R is
221 3.6.3.

222

223 **Ethics statement.** The study was approved by the Ethical Committee of Tongji Hospital,
224 Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
225 (ITJ-C20200128). Written informed consent was obtained from all participants enrolled in
226 this study.

227

228 **Results**

229 **The database framework and clinical information for the patients**

230 In this study, we collected 2,360 serum samples from 783 patients with an average age of
231 61.4 years and average onset time of 50 days. There were 387 males and 396 females
232 and 369 non-severe, 309 severe, 105 critical cases. Regarding outcome, there were 723
233 survivors and 60 deaths (**Fig. 1 A, Table 1, Supplementary dataset 1**).

234 To systematically analyse immune responses to SARS-CoV-2 infection, we screened
235 2,360 serum samples using a COVID-19 protein microarray that contains 21 proteins and
236 197 spike protein peptides. Additionally, we analysed 89 blood parameters for the 2,360

237 serum samples, *i.e.*, complete blood count, blood chemistry study and blood enzyme tests.
238 Hence, we obtained a comprehensive dataset that contains SARS-CoV-2-specific
239 humoral responses and is rich in clinical parameters.

240 By combining clinical information, IgG/IgM immune responses and blood parameters,
241 we established a database (COVID-ONE humoral immune) that provides a one-stop
242 analysis pipeline for COVID-19-specific immune responses and clinical parameters (Fig. 1
243 B). To allow users to obtain more COVID-19 serum profiling data, we set up a page on the
244 COVID-ONE humoral immune website, named “More studies”, to archive other highly
245 related data of COVID-19 serum profiling (protein/peptide microarrays/phage display
246 [15-20]. In addition, a healthy control dataset was added to the HELP page, which
247 contains the IgG and IgM responses for 528 healthy people to the 21 proteins and spike
248 protein peptides (Supplementary dataset 2).

249 The following steps are included in the analysis module:

- 250 ● Users select a set of samples in the panel of patient information and click START.
- 251 ● COVID-ONE humoral immune filters candidate samples according to the given
252 parameters.
- 253 ● COVID-ONE-humoral immune conducts analysis and provides results on the
254 webpage.

255 To demonstrate how to use COVID-ONE humoral immunity for analysis, we provide 2
256 datasets for a single group and paired groups as examples.

257

258 **Case □: Antibody responses and clinical parameters of non-survivors of COVID-19**

259 To study features of COVID-19 non-survivors, we selected the “death” parameter of
260 outcome in a single-group analysis module. This cohort contained 392 serum samples
261 and 60 cases, with an average age of 69.6 years and sex (38 male, 22 female) (Table 2).

262 The IgG response landscape analysis of SARS-CoV-2 proteins showed positive rates for
263 the S and N proteins and ORF3b of 95%, 93% and 87%, respectively, consistent with
264 previous studies [21, 22] (**Fig. 2A**). Interestingly, non-structural protein 7 (NSP7) had an 88%
265 IgG positive rate, which suggests that NSP7 may play an important role in COVID-19 (**Fig.**
266 **2A**). Spike peptide S1-45 had the highest positive rate (87%) for the IgM response,
267 indicating that the region including S1-45 may play an important role in IgM immunity (**Fig.**
268 **S1**).

269 Correlation analysis of clinical parameters showed that the neutrophil count had a
270 negative correlation with the monocyte count and lymphocyte ratio (**Fig. 2B**). In addition,
271 correlation analysis of antibody IgG responses showed a high correlation for IgG
272 responses of the S1 and N proteins, but not for S2, with all non-structural protein IgG
273 responses having no or very weak correlations (**Fig. 2C**). To study influencing factors of
274 S1 antibody production, we analysed correlation between the S1 IgG response and
275 clinical parameters and found the response to correlate with globulin in patients with
276 critical COVID-19 (**Fig. 2D**).

277

278 **Case □: Differences in IgG/IgM immune responses and clinical parameters associated**
279 **with sex**

280 Previous studies have shown that sex has a considerable effect on the outcome of
281 COVID-19 [23, 24] and is associated with underlying differences in immune responses to
282 infection [25]. To study differences in IgG/IgM immune responses and clinical parameters
283 between the sexes, we defined Group A as female and Group B as male for severe
284 patients, with 231 males at an average age of 64.3 and 183 females at an average age of
285 68.1. Consistent with previous studies [26], males had a higher risk of severe disease
286 than females (231/377 vs 183/379, $p < 0.001$) (**Table 3, Table 4**).

287 UMAP results showed no overall difference in IgG immunity between males and
288 females (**Fig. 3A**). To explore the disease mechanism in the sexes, we performed in-depth
289 analysis for antibody response and blood parameters using COVID-ONE. The antibody
290 response landscape shows that male patients have a higher positive rate than females for
291 ORF-9b IgG, RdRp IgG, NSP1 IgG, *etc.* (**Fig. 3B**). Moreover, longitudinal antibody
292 dynamic analysis showed a stronger ORF-9b IgG response in males during the whole
293 period of symptom onset, with a stronger NSP1 IgG response during the early stage of
294 symptom onset; however, there was no significant difference for RdRp IgG (**Fig. 3C-E**).
295 ORF-9b has been considered a drug target for the treatment of COVID-19 because it
296 suppresses type I interferon responses[27-29]. To explore the relevance between ORF-9b
297 antibody responses and COVID-19 severity, we compared ORF-9b antibody responses
298 between mild and severe cases, and the results showed that males with severe disease
299 had higher ORF-9b antibody responses than females (**Fig. 3G-H**).

300 To further decipher differences between female and male patients with COVID-19, we
301 employed random forest for machine learning. The results showed creatinine, which is an
302 acute kidney injury marker, to be the most significant factor between males and females
303 (**Fig. 4A**). To explore the relevance between creatinine and sex in COVID-19, we
304 compared the level and dynamic response of creatinine in males and females and
305 observed that the creatinine level in males was significantly higher than that in females
306 (**Fig. 4B-C**). To explore the relevance between creatinine and COVID-19 severity, we
307 compared creatinine levels in mild and severe cases, and similar to ORF-9b antibody
308 responses, male patients with severe COVID-19 had a higher level of creatinine (**Fig.**
309 **4D-E**). Hence, ORF-9b antibodies and creatinine are associated with severe disease in
310 male patients, which suggests different pathogeneses and complications between male
311 and female COVID-19 patients.

312

313 **Discussion**

314 In this study, we built COVID-ONE humoral immune, a COVID-19-specific database,
315 using R Shiny. COVID-ONE humoral immune is based on a comprehensive dataset
316 generated by analysing 2,360 COVID-19 sera using a SARS-CoV-2 protein microarray
317 containing 21 of the 28 known SARS-CoV-2 proteins and 197 peptides completely
318 covering the entire S protein sequence.

319 There are several published studies identifying the clinical characteristics, biomarkers and
320 specific antibody responses of diverse COVID-19 patients (**Table S1**). To strengthen the
321 credibility of our dataset, we compared COVID-19-specific antibody responses with other
322 studies at different levels. At the protein level, we analysed the dynamic responses to the
323 S and N proteins. The results showed that S and N responses peaked at 6 weeks after the
324 onset of symptoms for IgG and 4 weeks for IgM, which is consistent with the results of
325 previous studies[19, 21] (**Fig. S2**). At the peptide level, we compared IgG recognition of
326 immunodominant regions in the SARS-CoV-2 spike protein and found that some high
327 response areas that we identified[12] were consistent with those of Shrock *et al.* [15]:
328 25-36 aa, 553-588 aa, 770-829 aa, 1148-1159 aa and 1256-1273 aa. Another hot spot (aa
329 451-474) was only detected in our study. Regarding antibody diagnosis, Assia *et al.*
330 achieved an AUC of 0.986 for IgG and 0.988 for IgM for the detection of prior
331 SARS-CoV-2 infection when combining N and spike[20]. In our study, the AUC of the N
332 protein was 0.995 for IgG and 0.988 for IgM, and the AUC of the S1 protein was 0.992 for
333 IgG and 0.992 for IgM. We also found that S2-78 (1148-1159 aa) IgG is comparable to
334 S1 IgG for COVID-19 patients, with an AUC of 0.99 for IgG and 0.953 for IgM[11].
335 To our knowledge, COVID-19 humoral immune is the first database for COVID-19-specific
336 immune responses enriched in clinical parameters and has the following features. (i)

337 Universality: COVID-ONE humoral immune contains 783 patients with 16 medical
338 histories, which will be of broad interest for researchers and clinicians from diverse
339 backgrounds. (ii) Accessibility: COVID-ONE-humoral immune provides a one-stop
340 analysis pipeline, by which users can easily obtain meaningful information. (iii) Scalability:
341 COVID-ONE humoral immune is built on the R platform, which is freely accessible, and
342 many modular tools are readily available; thus, we can easily expand and incorporate new
343 analyses for the dataset whenever necessary without changing the overall structure of the
344 database. Nonetheless, there are some limitations for COVID-ONE humoral immunity.
345 For example, it lacks data for convalescent patients, peptide-level humoral responses to
346 proteins other than S protein, and multicentre samples. In the future, we will assay the
347 dynamic responses of SARS-CoV-2-specific antibodies using ~500 serum samples from
348 ~100 COVID-19 convalescent patients. We will also integrate published peptide
349 microarray/phage display-related data[15-17, 30] and attempt to update the database
350 covering the whole SARS-CoV-2 proteome at the peptide or amino acid level. In addition,
351 the SARS-CoV-2 protein microarray has already been promoted by CDI Labs
352 (www.cdi.bio) and ArrayJet (www.arrayjet.co.uk), and we anticipate more diverse data for
353 SARS-CoV-2-specific antibody responses from multicentre samples. We strongly believe
354 that by sharing a large dataset and facilitating data analysis, COVID-19 humoral immune
355 is a valuable resource for COVID-19 research.

356

357 **Data and tool availability**

358 COVID-ONE-humoral immune is freely accessible at www.covid-one.cn. The
359 SARS-CoV-2 proteome microarray data are deposited on Protein Microarray Database
360 under the accession number PMDE244 (<http://www.proteinmicroarray.cn>). If author need
361 the raw data of antibody responses or clinical parameters, please contact the

362 corresponding author (taosc@sjtu.edu.cn).

363

364 **Author's contributions**

365 SCT and XLF developed the conceptual ideas and designed the study. ZWX, LKH, YL,

366 QL, DYL, SJG, HWJ, HNZ, HQ, XL, performed the experiments and data analysis. ZWX,

367 LKH, YL, XL built the database. SCT and ZWX wrote the manuscript with suggestions

368 from other authors.

369

370 **Competing interests**

371 The authors declare no competing interests.

372

373 **Acknowledgments**

374 This work was partially supported by the National Key Research and Development

375 Program of China Grant (No.2016YFA0500600), National Natural Science Foundation of

376 China (No. 31970130, 31600672, 31900112, 21907065 and 32000027).

377

378 **References:**

379 [1] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time.

380 *The Lancet Infectious Diseases* 2020;20:533-4.

381 [2] Knipe DM, Levy O, Fitzgerald KA, Mühlberger E. Ensuring vaccine safety. *Science*

382 2020:eabf0357.

383 [3] Lipsitch M, Dean NE. Understanding COVID-19 vaccine efficacy. *Science* 2020;370:763-5.

384 [4] Chen Y, Zuiani A, Fischinger S, Mullur J, Atyeo C, Travers M, et al. Quick COVID-19 Healers

385 Sustain Anti-SARS-CoV-2 Antibody Production. *Cell* 2020.

386 [5] Peng Y, Mentzer AJ, Liu G, Yao X, Yin Z, Dong D, et al. Broad and strong memory CD4(+) and

387 CD8(+) T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat*

388 *Immunol* 2020;21:1336-45.

389 [6] Cox RJ, Brokstad KA. Not just antibodies: B cells and T cells mediate immunity to COVID-19. *Nat*

390 *Rev Immunol* 2020;20:581-2.

391 [7] Jiang HW, Li Y, Zhang HN, Wang W, Yang X, Qi H, et al. SARS-CoV-2 proteome microarray for

392 global profiling of COVID-19 specific IgG and IgM responses. *Nat Commun* 2020;11:3581.

- 393 [8] Li Y, Lai DY, Zhang HN, Jiang HW, Tian X, Ma ML, et al. Linear epitopes of SARS-CoV-2 spike
394 protein elicit neutralizing antibodies in COVID-19 patients. *Cell Mol Immunol* 2020;17:1095-7.
- 395 [9] Lei Q, Li Y, Hou HY, Wang F, Ouyang ZQ, Zhang Y, et al. Antibody dynamics to SARS-CoV-2 in
396 asymptomatic COVID-19 infections. *Allergy* 2020.
- 397 [10] Lei Q, Yu C-z, Li Y, Hou H-y, Xu Z-w, Yao Z-j, et al. SARS-CoV-2 antibody signatures for
398 predicting the outcome of COVID-19. *medRxiv* 2020.
- 399 [11] Li Y, Lai DY, Lei Q, Xu ZW, Wang F, Hou H, et al. Systematic evaluation of IgG responses to
400 SARS-CoV-2 spike protein-derived peptides for monitoring COVID-19 patients. *Cell Mol Immunol*
401 2021;18:621-31.
- 402 [12] Li Y, Ma M-l, Lei Q, Wang F, Hong W, Lai D-y, et al. Linear epitope landscape of the
403 SARS-CoV-2 Spike protein constructed from 1,051 COVID-19 patients. *Cell Reports* 2021.
- 404 [13] Li Y, Xu Z, Lei Q, Lai DY, Hou H, Jiang HW, et al. Antibody landscape against SARS-CoV-2
405 reveals significant differences between non-structural/accessory and structural proteins. *Cell Rep*
406 2021;36:109391.
- 407 [14] Li Y, Li CQ, Guo SJ, Guo W, Jiang HW, Li HC, et al. Longitudinal serum autoantibody repertoire
408 profiling identifies surgery-associated biomarkers in lung adenocarcinoma. *EBioMedicine*
409 2020;53:102674.
- 410 [15] Shrock E, Fujimura E, Kula T, Timms RT, Lee IH, Leng Y, et al. Viral epitope profiling of
411 COVID-19 patients reveals cross-reactivity and correlates of severity. *Science* 2020;370.
- 412 [16] Wang H, Wu X, Zhang X, Hou X, Liang T, Wang D, et al. SARS-CoV-2 Proteome Microarray for
413 Mapping COVID-19 Antibody Interactions at Amino Acid Resolution. *ACS Cent Sci* 2020;6:2238-49.
- 414 [17] Schwarz T, Heiss K, Mahendran Y, Casilag F, Kurth F, Sander LE, et al. SARS-CoV-2
415 Proteome-Wide Analysis Revealed Significant Epitope Signatures in COVID-19 Patients. *Front*
416 *Immunol* 2021;12:629185.
- 417 [18] Holenya P, Lange PJ, Reimer U, Woltersdorf W, Panterodt T, Glas M, et al. Peptide
418 microarray-based analysis of antibody responses to SARS-CoV-2 identifies unique epitopes with
419 potential for diagnostic test development. *Eur J Immunol* 2021.
- 420 [19] Wu J, Liang B, Chen C, Wang H, Fang Y, Shen S, et al. SARS-CoV-2 infection induces sustained
421 humoral immune responses in convalescent patients following symptomatic COVID-19. *Nat Commun*
422 2021;12:1813.
- 423 [20] de Assis RR, Jain A, Nakajima R, Jasinskas A, Felgner J, Obiero JM, et al. Analysis of
424 SARS-CoV-2 antibodies in COVID-19 convalescent blood using a coronavirus antigen microarray. *Nat*
425 *Commun* 2021;12:6.
- 426 [21] Long QX, Liu BZ, Deng HJ, Wu GC, Deng K, Chen YK, et al. Antibody responses to
427 SARS-CoV-2 in patients with COVID-19. *Nat Med* 2020;26:845-8.
- 428 [22] Hachim A, Kavian N, Cohen CA, Chin AWH, Chu DKW, Mok CKP, et al. ORF8 and ORF3b
429 antibodies are accurate serological markers of early and late SARS-CoV-2 infection. *Nat Immunol*
430 2020;21:1293-301.
- 431 [23] Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated
432 with COVID-19-related death using OpenSAFELY. *Nature* 2020;584:430-6.
- 433 [24] Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. Considering how biological sex
434 impacts immune responses and COVID-19 outcomes. *Nat Rev Immunol* 2020;20:442-7.
- 435 [25] Takahashi T, Ellingson MK, Wong P, Israelow B, Lucas C, Klein J, et al. Sex differences in
436 immune responses that underlie COVID-19 disease outcomes. *Nature* 2020.

437 [26] Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics
438 of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet*
439 2020;395:507-13.

440 [27] Jiang HW, Zhang HN, Meng QF, Xie J, Li Y, Chen H, et al. SARS-CoV-2 Orf9b suppresses type I
441 interferon responses by targeting TOM70. *Cell Mol Immunol* 2020;17:998-1000.

442 [28] Gordon DE, Hiatt J, Bouhaddou M, Rezelj VV, Ulferts S, Braberg H, et al. Comparative
443 host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* 2020;370.

444 [29] Gao X, Zhu K, Qin B, Olieric V, Wang M, Cui S. Crystal structure of SARS-CoV-2 Orf9b in
445 complex with human TOM70 suggests unusual virus-host interactions. *Nat Commun* 2021;12:2843.

446 [30] Qi H, Ma ML, Jiang HW, Ling JY, Chen LY, Zhang HN, et al. Systematic profiling of
447 SARS-CoV-2-specific IgG epitopes at amino acid resolution. *Cell Mol Immunol* 2021;18:1067-9.

448

449

450 **Figure legends**

451 **Figure 1. Overview of data resources and functional modules of COVID-ONE humoral**
452 **immunity.**

453 **(A)** The patient information of the study cohort showing sex, outcome, severe type, *etc.* **(B)**
454 The framework of COVID-ONE-humoral immune. The one-stop database for
455 COVID-19-specific humoral immune responses and clinical parameters. The COVID-ONE
456 humoral immune dataset includes 220 protein/peptide antibody responses and 96 clinical
457 parameters from 2360 serum samples. Using the Shiny package, COVID-ONE-humoral
458 immune provides single-group or paired-group analysis based on the dataset.

459

460 **Figure 2. SARS-CoV-2-specific antibody responses and their correlations with clinical**
461 **parameters: COVID-19 non-survivors.**

462 **(A)** The antibody IgG response landscape against SARS-CoV-2 proteins (upper part), S1
463 protein peptides (middle part) and S2 protein peptides (lower part). **(B)** Heat map showing
464 correlation analysis of blood parameters. **(C)** Heat map showing correlation analysis of
465 antibody IgG responses against SARS-CoV-2 proteins. **(D)** Scatter plot showing
466 correlation analysis between the S1 IgG response and protein IgG responses/blood

467 parameters.

468

469 **Figure 3. Correlation of the ORF-9b IgG response based on COVID-19 severity in male**
470 **patients.**

471 **(A)** Scatter plot showing uniform manifold approximation and projection (UMAP) for serum
472 samples using 21 protein IgG/IgM responses in sex subgroup analysis. **(B)** Histogram
473 showing different responses in males and females for the IgG response. **(C-E)** Scatter plot
474 showing ORF9b, RdRp and NSP1 IgG dynamic responses using longitudinal samples
475 from male and female patients. **(G-H)** Scatter plot of the dynamic anti-ORF9b IgG
476 response in COVID-19 patients with mild and severe symptoms.

477

478 **Figure 4. Correlation of creatinine response based on COVID-19 severity in male patients.**

479 **(A)** The top 15 sex-specific parameters by random forest analysis ranked by the mean
480 decrease in accuracy and mean decrease in the Gini coefficient **(B)** The boxplot shows
481 the significant difference in creatinine in sex subgroup analysis. The *P*-value was
482 calculated by a two-sided t-test. **(C)** Scatter plot of creatinine levels of male and female
483 COVID-19 patients. **(D-E)** Scatter plot of creatinine levels of COVID-19 patients with mild
484 and severe disease.

485

486 **Figure S1. The antibody IgM response landscape against SARS-CoV-2 proteins (upper**
487 **part), S1 protein peptides (middle part) and S2 protein peptides (lower part).**

488

489 **Figure S2. Dynamic antibody responses to S1 and N proteins.** Scatter plot showing
490 dynamic antibody responses to S1 IgG **(A)**, N protein IgG **(B)**, S1 IgM **(C)**, and N protein
491 IgM **(D)**.

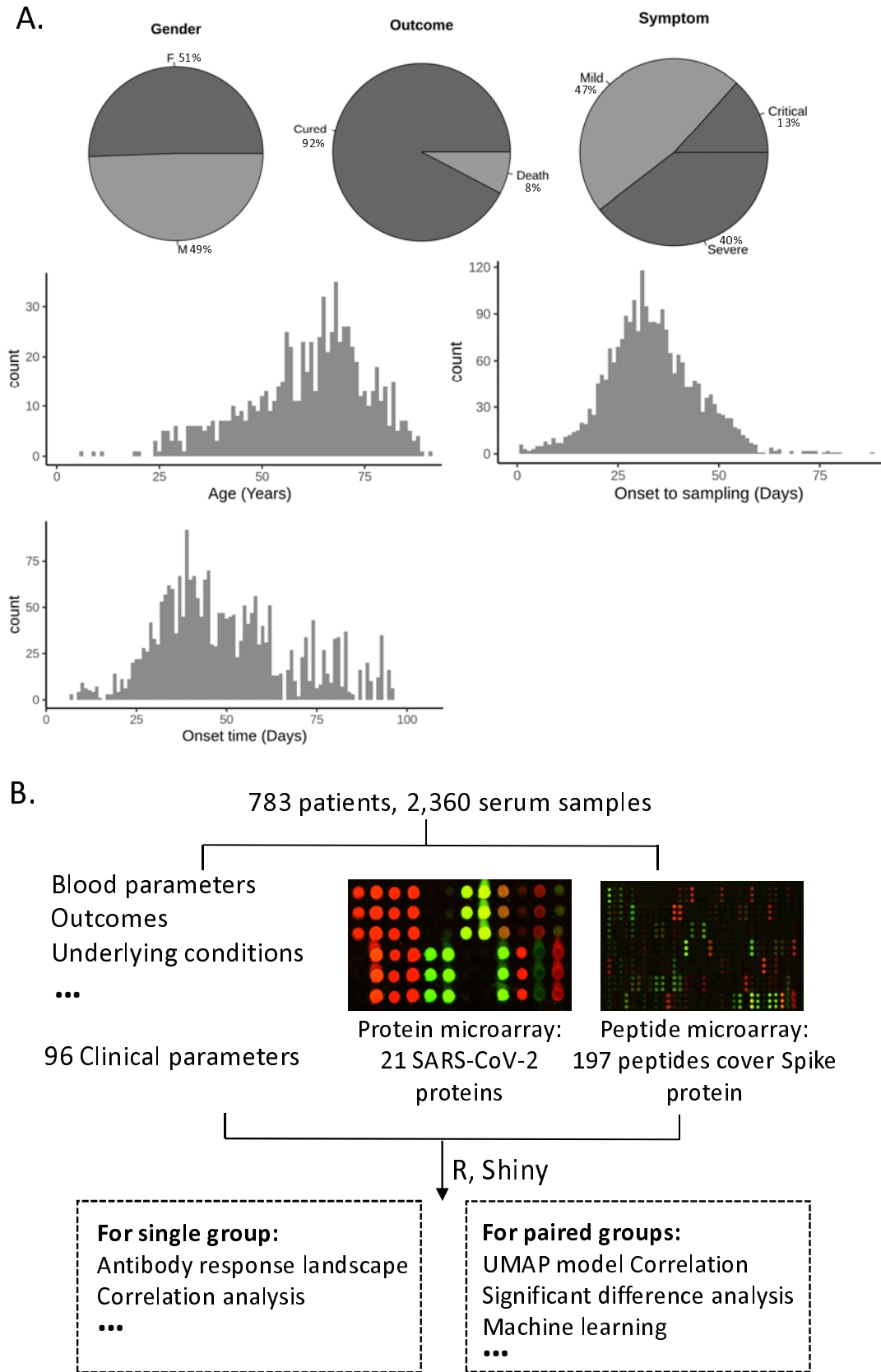
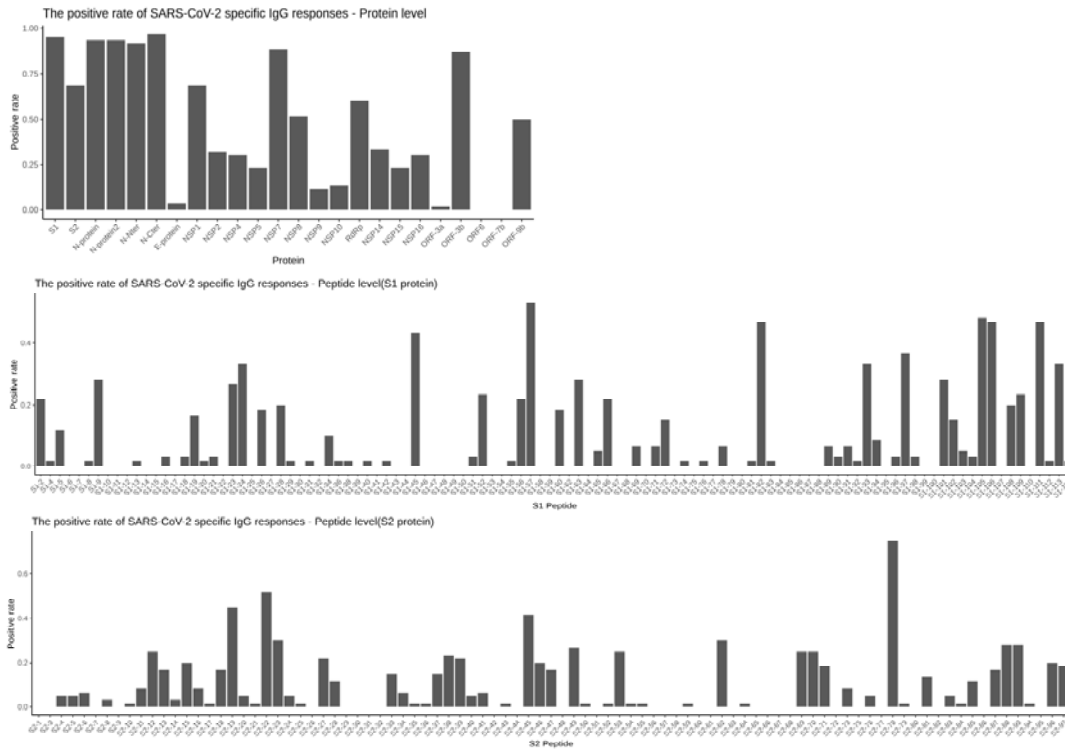
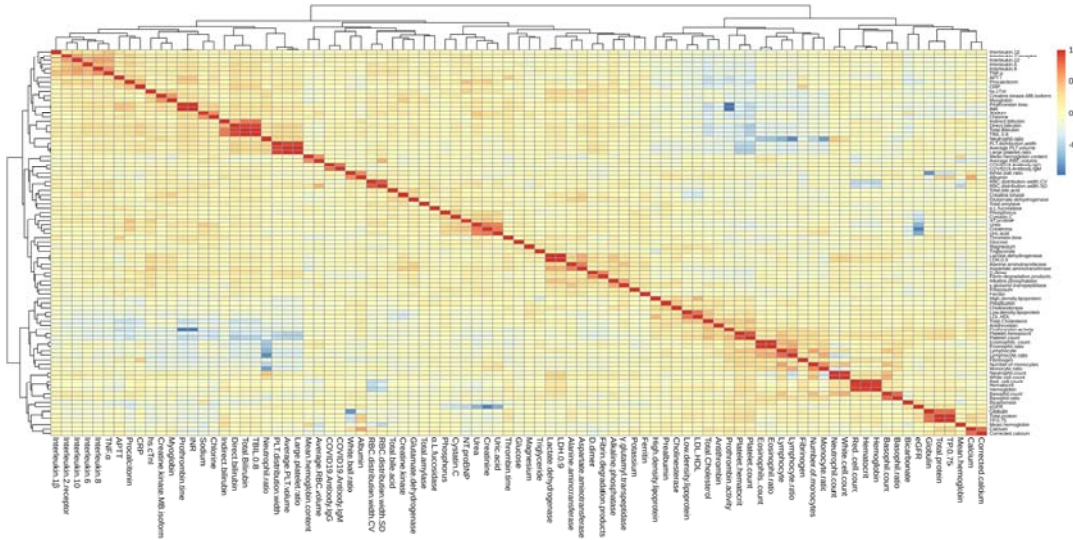


Figure 1. Overview of data resources and functional modules of COVID-ONE-humoral immune.

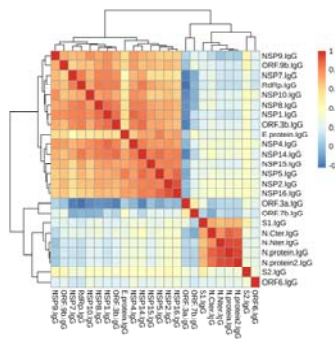
A.



B.



C.



D.

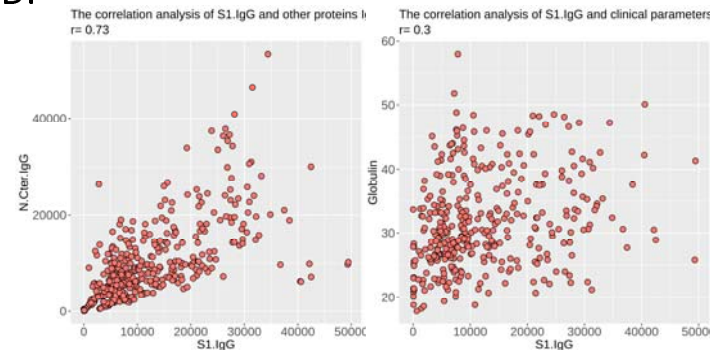


Figure 2. SARS-CoV-2 specific antibody responses and its correlations with clinic parameters: COVID-19 non-survivors.

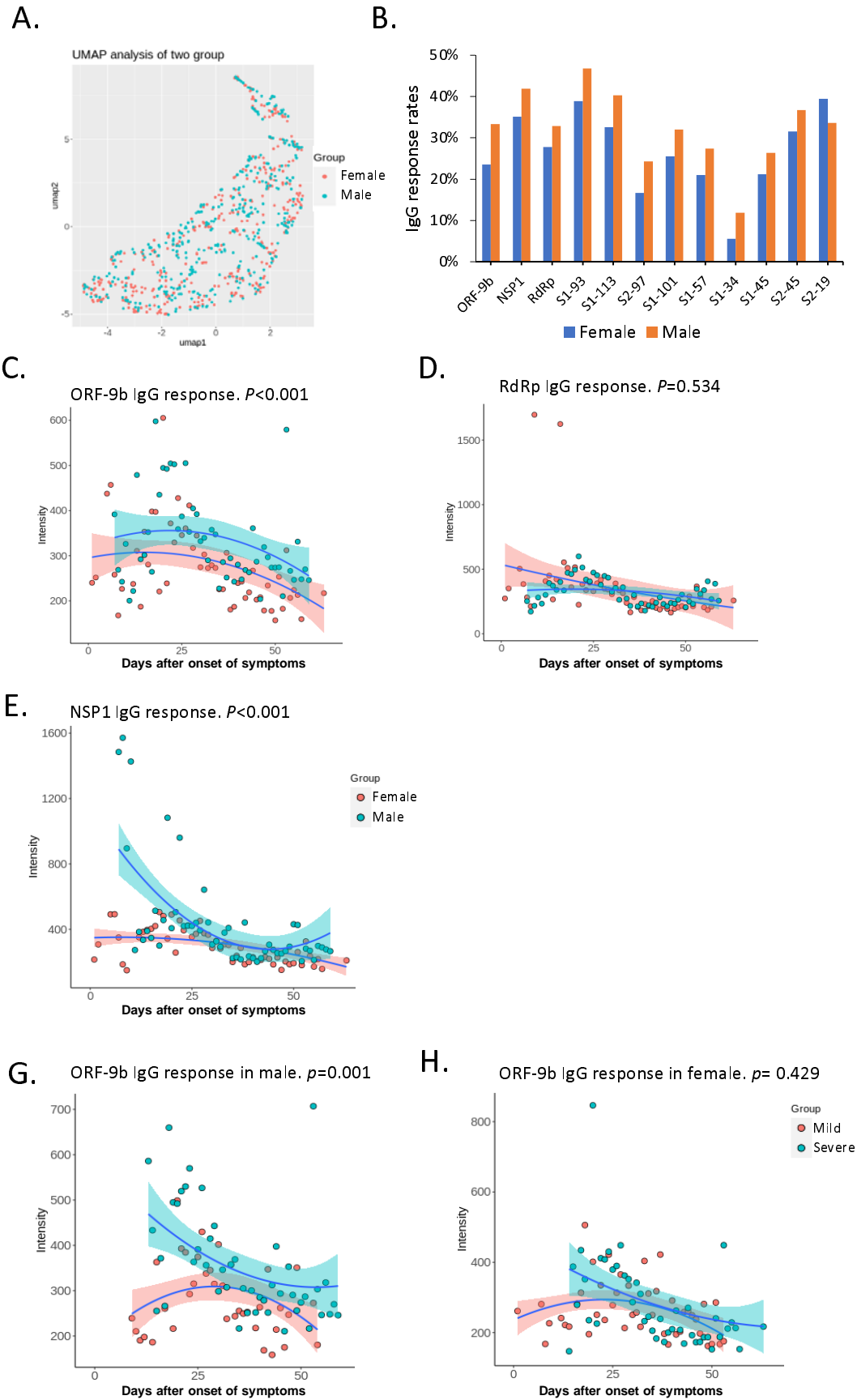


Figure 3. Correlation of ORF-9b IgG response with severe male in COVID-19 patients.

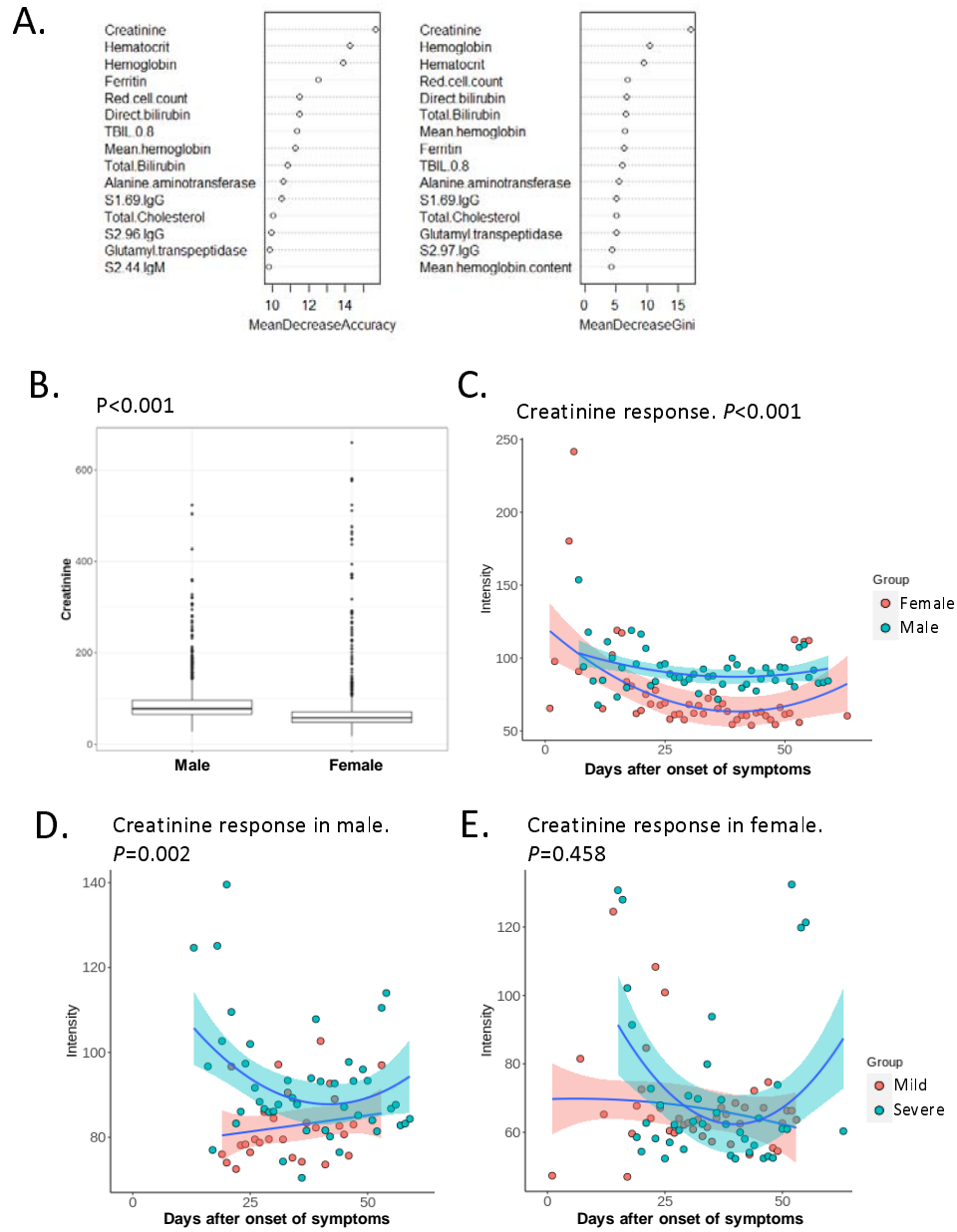


Figure 4. Correlation of creatinine response with severe male in COVID-19 patients.