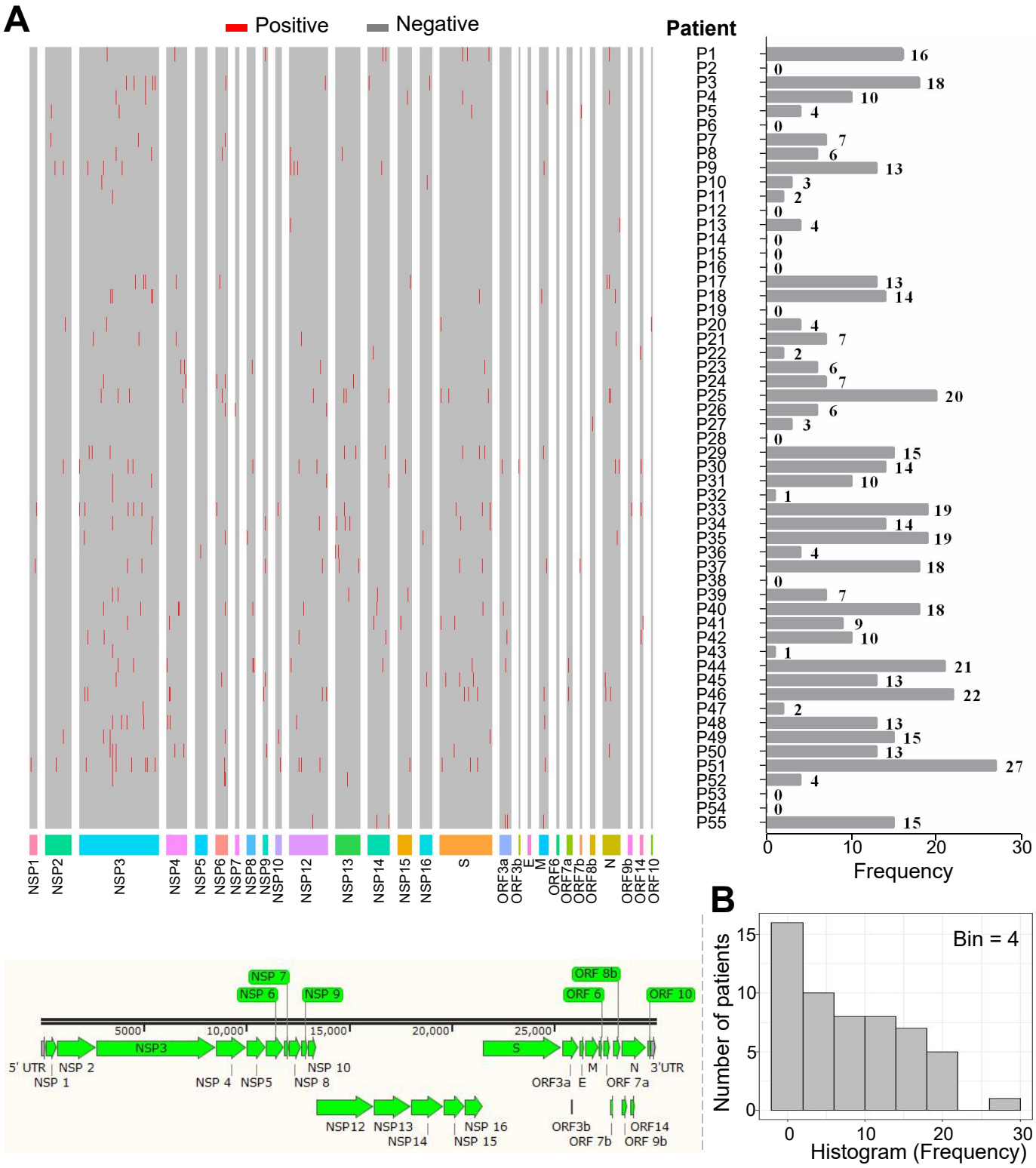


C

Protein	NSP1	NSP2	NSP3	NSP4	NSP5	NSP6	NSP7	NSP8	NSP9	NSP10	NSP12	NSP13	NSP14	NSP15	NSP16	S	ORF3a	ORF3b	E	M	ORF6	ORF7a	ORF7b	ORF8b	N	ORF9b	ORF14	ORF10
Sum_epitopes	8	16	125	30	7	20	1	8	9	3	48	22	26	9	11	69	9	1	1	12	0	3	2	2	20	1	5	1
Length	180	638	1945	500	306	290	83	198	113	139	932	601	527	346	298	1273	275	22	75	222	61	121	43	121	419	97	73	38
Sum_epitopes/ Length	0.04	0.03	0.06	0.06	0.02	0.07	0.01	0.04	0.08	0.02	0.05	0.04	0.05	0.03	0.04	0.05	0.03	0.05	0.01	0.05	0.00	0.02	0.05	0.02	0.05	0.01	0.07	0.03

Fig. S1 The map of SARS-CoV-2 specific IgG binding epitopes revealed from convalescent sera directly. (A) The epitopes (**Table S3**) of each patient (P1-P55) were plotted alongside the known SARS-CoV-2 proteins from N terminal to C terminal. The total number of the identified epitopes for each patient and each protein were defined as frequency and sum_epitopes, respectively. (**B**) The histogram of the epitope frequency (bin = 4). (**C**) The sum_epitopes of each protein, length of proteins and ratio of sum_epitopes and length.

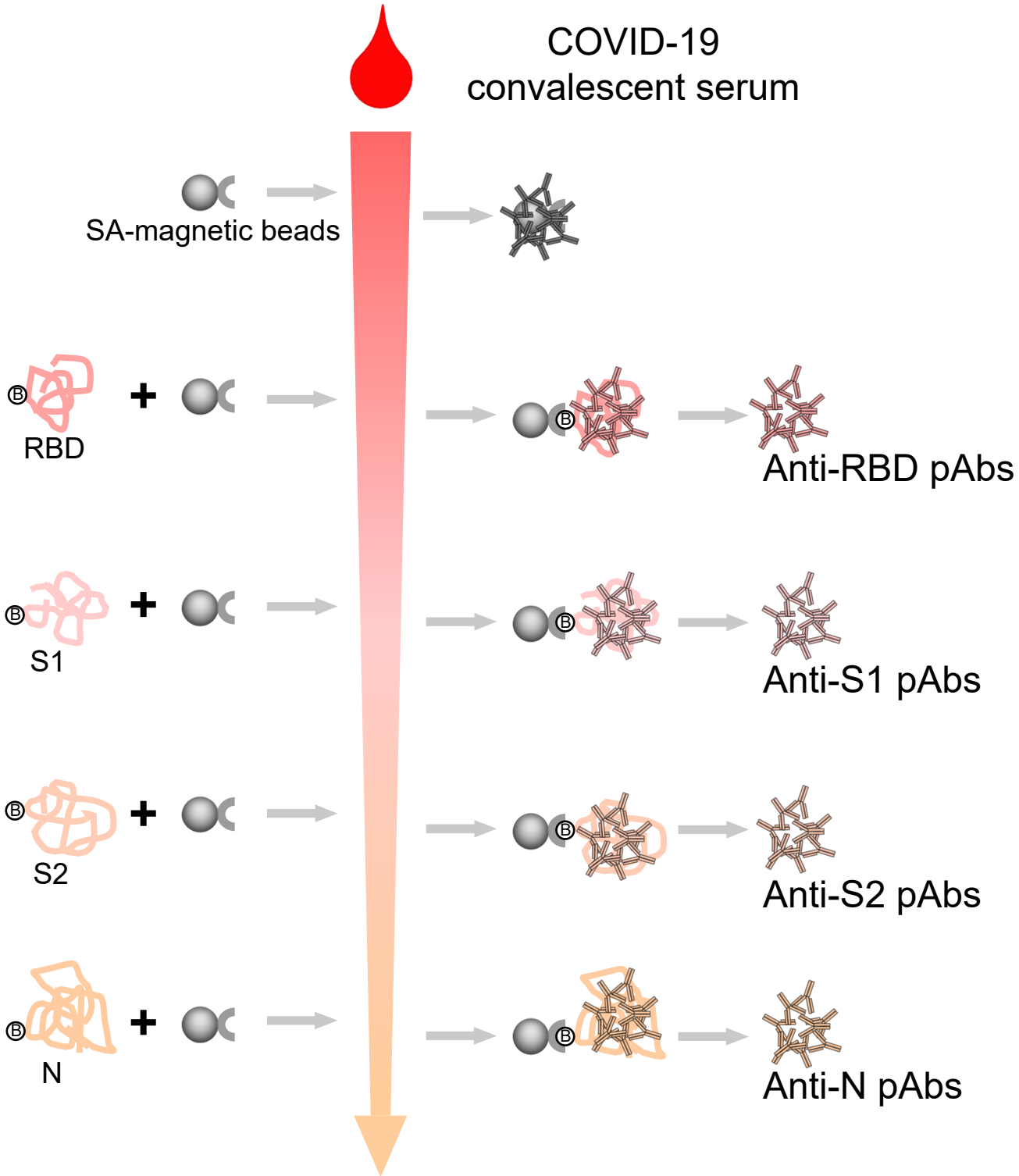


Fig. S2 Protein specific antibodies enrichment from COVID 19 convalescent sera. COVID-19 convalescent sera were used as input. Terminal biotinylated proteins were used as baits to enrich specific antibodies. SA conjugated magnetic beads were adopted to isolate the antibodies.

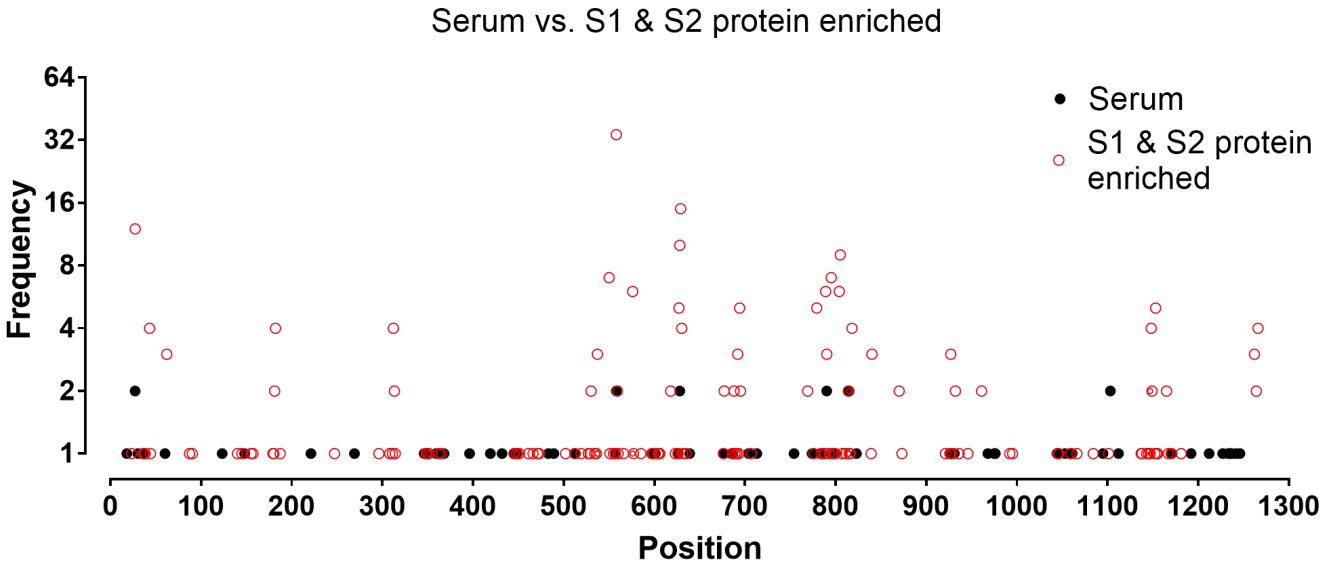


Fig. S3. Protein based antibody enrichment is necessary for AbMap. S protein specific epitopes identified from sera directly (black dot), and S1+S2 protein enriched antibodies (red dot).

A

No.	ID	Sequence	Position	Frequency
1	Epitope-S1-1	AYTNSFTR	27-34	12
2	Epitope-S1-2	FRSSVLHS	43-50	4
3	Epitope-S1-3	VTWFHAIH	62-69	3
4	Epitope-S1-4	KQGNFKNL	182-189	4
5	Epitope-S1-5	IYQTSNFR	312-319	4
6	Epitope-S1-6	KCVNFNFN	537-544	3
7	Epitope-S1-7	GVLTESNK	550-557	7
8	Epitope-S1-8	KFLPFQQF	558-565	34
9	Epitope-S1-9	VRDPQTLE	576-583	6
10	Epitope-S1-10	DQLTPTWR	627-634	5
11	Epitope-S1-11	QLTPTWRV	628-635	10
12	Epitope-S1-12	LTPTWRVY	629-636	15
13	Epitope-S1-13	TPTWRVYS	630-637	4
14	Epitope-S2-1	IIAYTMSL	692-699	3
15	Epitope-S2-2	AYTMSLGA	694-701	5
16	Epitope-S2-3	QEVFAQVK	779-786	5
17	Epitope-S2-4	YKTPPIKD	789-796	6
18	Epitope-S2-5	KTPPIKDF	790-797	3
19	Epitope-S2-6	KDFGGFNF	795-802	7
20	Epitope-S2-7	QILPDPSK	804-811	7
21	Epitope-S2-8	ILPDPSKP	805-812	9
22	Epitope-S2-9	IEDLLFNK	818-825	4
23	Epitope-S2-10	CLGDIAAR	840-847	3
24	Epitope-S2-11	FNSAIGKI	927-934	3
25	Epitope-S2-12	FKEELDKY	1148-1155	4
26	Epitope-S2-13	DKYFKNHT	1153-1160	5
27	Epitope-S2-14	EPVLKGVK	1262-1269	3
28	Epitope-S2-15	KGVKLHYT	1266-1273	4
29	Epitope-RBD-1	RDISTEIY	466-473	2

B

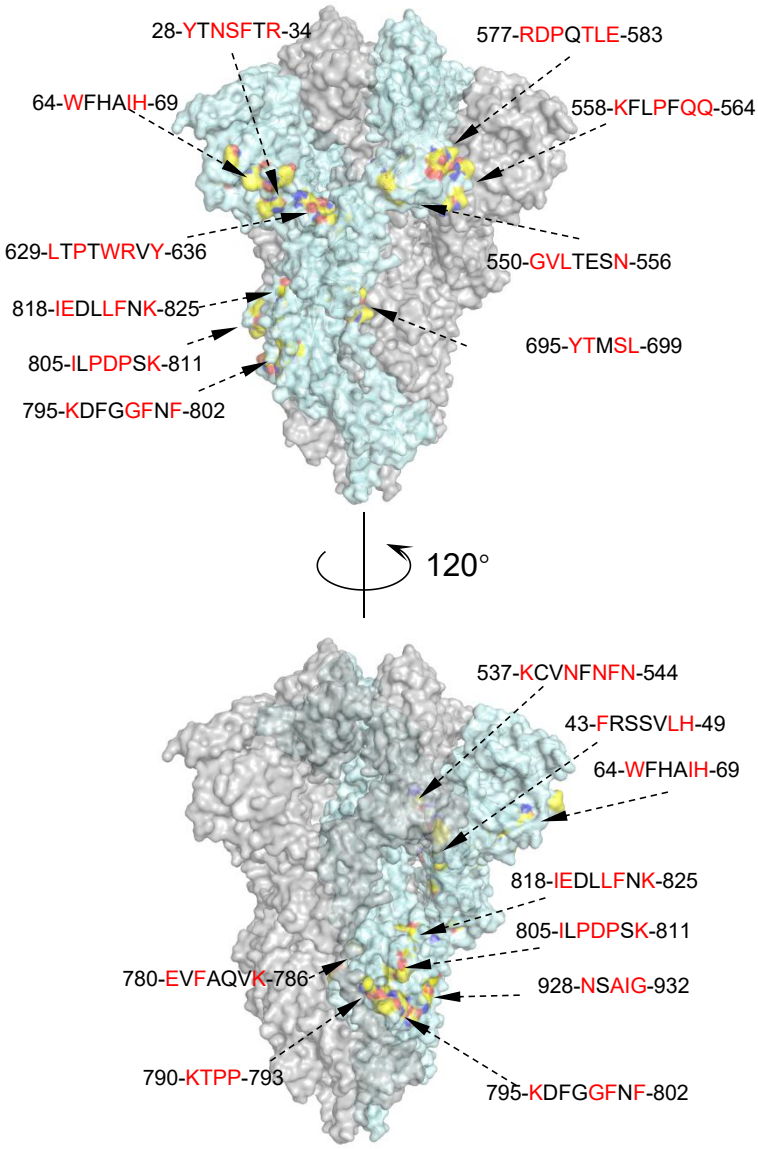


Fig. S4. Detailed information of the significant epitopes on S protein. (**A**) The list of the significant epitopes (frequency >=3), (**B**) The distribution of the significant epitopes (frequency>=3) on the trimer 3D structure of the S protein. Red marked amino acids represent critical residues of epitopes.

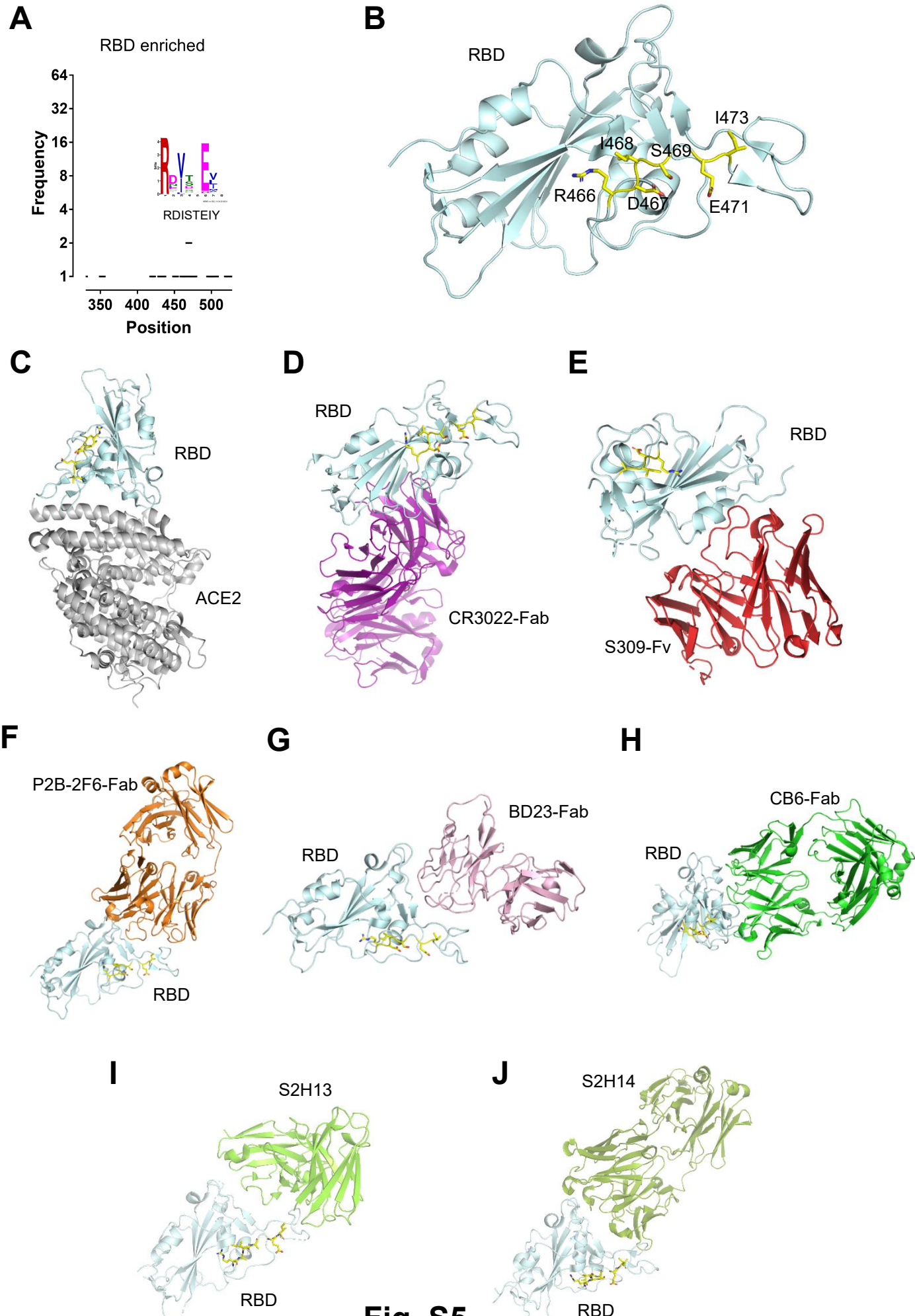
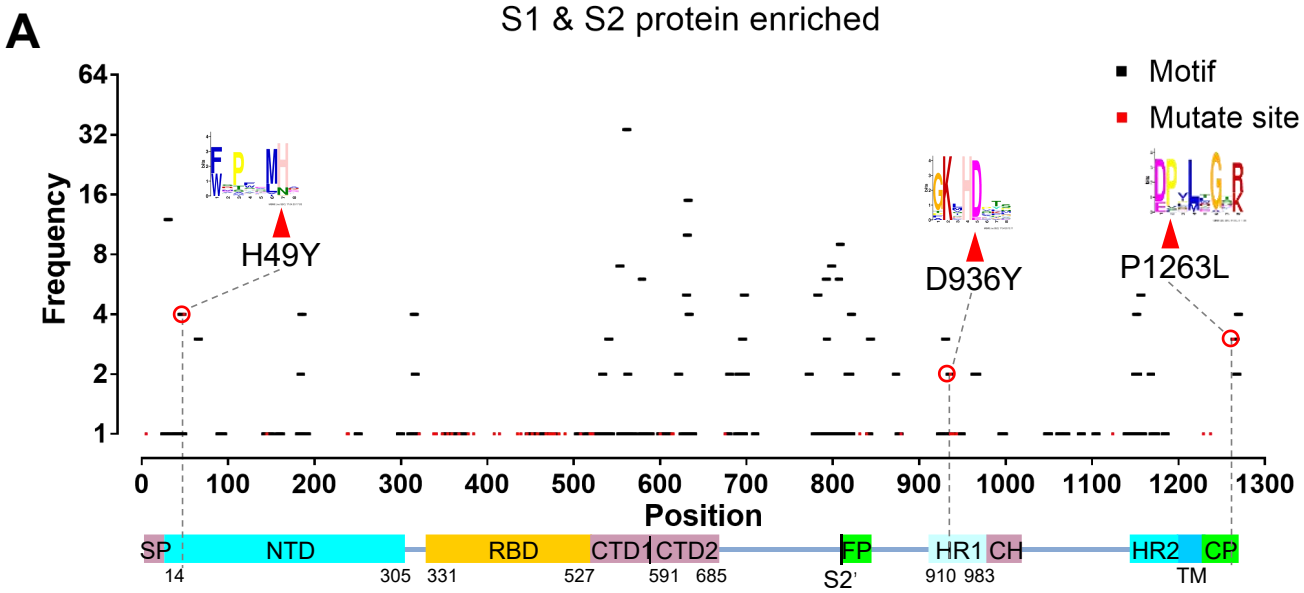

A
RBD enriched
Frequency
64
32
16
8
4
2
1
RDISTEIY
350
400
450
500
Position
B
RBD
I473
I468
S469
R466
D467
E471
C
RBD
ACE2
D
RBD
CR3022-Fab
E
RBD
S309-Fv
F
P2B-2F6-Fab
RBD
G
BD23-Fab
RBD
H
CB6-Fab
RBD
I
S2H13
RBD
J
S2H14
RBD

Fig. S5

Fig. S5. Functional analysis of several key sites of the significant epitopes of S protein. (**A**) One relatively significant epitope (frequency >= 2) was identified for RBD enriched antibody. (**B**) Match the critical epitope residues to the structure of RBD. (**C**) The critical epitope residues locate adjacent to but not at the binding interfaces of RBD and ACE2. (**D-J**) [1-6] The critical epitope residues locate adjacent to but not at the binding interfaces of RBD strong binding antibodies or highly potent neutralization antibodies, CR3022-Fab (**D**), S309-Fv (**E**), P2B-2F6-Fab (**F**), BD23-Fab (**G**), CB6-Fab (**H**), S2H13 (**I**) and S2H14 (**J**).



Mutation	Location	Mutate rate	Infectivity
H49Y	NTD	0.24%	-
D936Y	HR1	0.01%	60~70% ↓
P1263L	Cytoplasmic Tail	0.01%	98~99% ↓

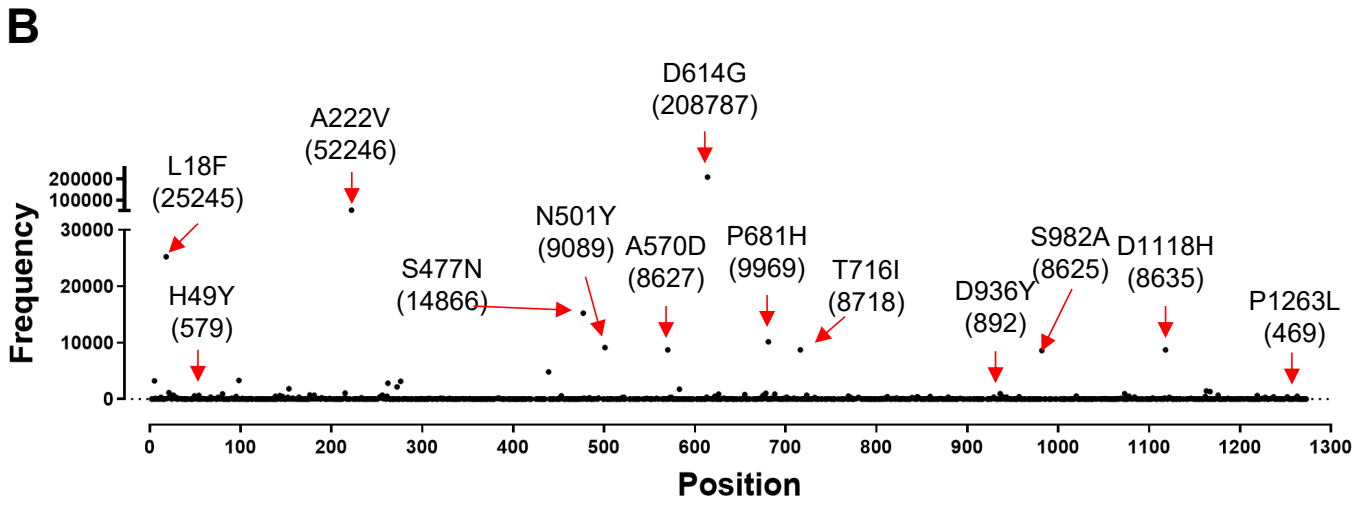


Fig. S6. Match the critical epitope residues with naturelly existing mutants of Spike protein. (**A**) Match the epitopes on S protein with the critical naturally existing mutants [7,8]. (**B**) The Substitution mutants of Spike protein collected by China National Center for Bioinformation were plotted alongside the Spike protein with the mutated frequency. (The data were collected by 2021-1-12)

B

No.	ID	Sequence	Position	Frequency
1	Epitope-N-1	SRNSSRNS	190-197	8
2	Epitope-N-2	FPPTEPKK	363-370	3
3	Epitope-N-3	QALPQRQK	380-387	3
4	Epitope-N-4	KQQTVTLL	388-395	3
5	Epitope-N-5	TLLPAADL	393-400	4
6	Epitope-N-6	KQLQQSMS	405-412	3

C

```
                  190     197
SARS-CoV-2        RSRNSSRNS
BtCoV-RaTG13      RSRNSSRNS
SARS-CoV          RSRGNSRNS
MERS-CoV          VSRNSSRSS
HCoV-OC43         RTSSRASSA
HCoV-HKU1         RSQSRGPNN
HCoV-NL63         RNNSRDSSR
HCoV-229E         RSQSR--GR
```

```
                  363     370
SARS-CoV-2        FPPTEPKK
BtCoV-RaTG13      FPPTEPKK
SARS-CoV          FPPTEPKK
MERS-CoV          FPKKE-KK
HCoV-OC43         --QQDGMM
HCoV-HKU1         SNQNTDSD
HCoV-NL63         PSSIKEMQ
HCoV-229E         --EMQQHP
```

```
                  380     387
SARS-CoV-2        QALPQRQK
BtCoV-RaTG13      QALPQRQK
SARS-CoV          QPLPQRQK
MERS-CoV          DQMSEPPK
HCoV-OC43         RGHKNGQG
HCoV-HKU1         RGVKQLPE
HCoV-NL63         TVLNASIP
HCoV-229E         NPSQTSPA
```

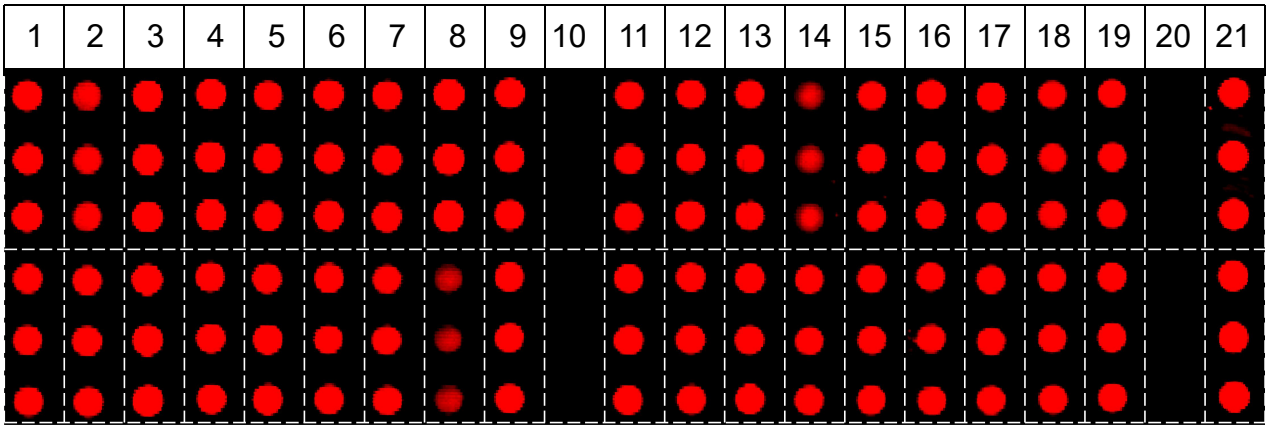
```
                  388     395
SARS-CoV-2        KQQTVTLL
BtCoV-RaTG13      KQQTVTLL
SARS-CoV          KQPTVTLL
MERS-CoV          EQRVQGSI
HCoV-OC43         ENDNISVA
HCoV-HKU1         QFDSLNLS
HCoV-NL63         ESKPLADD
HCoV-229E         TAEPVRDE
```

```
                  393             400
SARS-CoV-2        TLLPA---------ADL
BtCoV-RaTG13      TLLPA---------ADL
SARS-CoV          TLLPA---------ADM
MERS-CoV          GSITQ---------RTR
HCoV-OC43         SVAVPKSRVQQNKSREL
HCoV-HKU1         NLSAG----TQHISNDF
HCoV-NL63         ADDDS-----------
HCoV-229E         RDEVS-----------
```

```
                  405     412
SARS-CoV-2        KQLQQSMS
BtCoV-RaTG13      KQLQQSMS
SARS-CoV          RQLQNSMS
MERS-CoV          VQPGPMID
HCoV-OC43         ISLLKKMD
HCoV-HKU1         HSLLATLD
HCoV-NL63         IEIVNEVL
HCoV-229E         TDIIDEVN
```

Fig. S7. Significant epitopes on N protein. (**A**) The distribution of the sequence-matched epitopes on N protein. (**B**) The list of significant epitopes (frequency >= 3). (**C**) Homology analysis of the significant epitopes among three deadly coronaviruses, i. e., SARS-CoV-2, SARS-CoV and MERS-CoV, four common human coronaviruses, i. e., HCoV-OC43, HKU-1, NL63 and 229E and the highly homologous bat coronavirus BtCoV-RaTG13.

A



B

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Pep1-WT	Pep1-K558A	Pep1-F559A	Pep1-L560A	Pep1-P561A	Pep1-F562A	Pep1-Q563A	Pep1-Q564A	Pep1-F565A	Buffer	Pep2-WT	Pep2-L629A	Pep2-T630A	Pep2-Pep631A	Pep2-T632A	Pep2-W633A	Pep2-R634A	Pep2-V635A	Pep2-Y636A	h-IgM	Cy5-BSA
Pep3-WT	Pep3-D1153A	Pep3-K1154A	Pep3-Y1155A	Pep3-F1156A	Pep3-K1157A	Pep3-N1158A	Pep3-H1159A	Pep3-T1160A	Buffer	Pep4-WT	Pep4-R466A	Pep4-D467A	Pep4-I468A	Pep4-S469A	Pep4-T470A	Pep4-E471A	Pep4-I472A	Pep4-Y4T3A	h-IgG	Cy5-BSA

C

ID	Sequence	ID	Sequence
Pep1-WT	**K F L P F Q Q F**	Pep2-WT	**L T P T W R V Y**
Pep1-K558A	**A** F L P F Q Q F	Pep2-L629A	**A** T P T W R V Y
Pep1-F559A	K **A** L P F Q Q F	Pep2-T630A	L **A** P T W R V Y
Pep1-L560A	K F **A** P F Q Q F	Pep2-P631A	L T **A** T W R V Y
Pep1-P561A	K F L **A** F Q Q F	Pep2-T632A	L T P **A** W R V Y
Pep1-F562A	K F L P **A** Q Q F	Pep2-W633A	L T P T **A** R V Y
Pep1-Q563A	K F L P F **A** Q F	Pep2-R634A	L T P T W **A** V Y
Pep1-Q564A	K F L P F Q **A** F	Pep2-V635A	L T P T W R **A** Y
Pep1-F565A	K F L P F Q Q **A**	Pep2-Y636A	L T P T W R V **A**
Pep3-WT	**D K Y F K N H T**	Pep4-WT	**R D I S T E I Y**
Pep3-D1153A	**A** K Y F K N H T	Pep4-R466A	**A** D I S T E I Y
Pep3-K1154A	D **A** Y F K N H T	Pep4-D467A	R **A** I S T E I Y
Pep3-Y1155A	D K **A** F K N H T	Pep4-I468A	R D **A** S T E I Y
Pep3-F1156A	D K Y **A** K N H T	Pep4-S469A	R D I **A** T E I Y
Pep3-K1157A	D K Y F **A** N H T	Pep4-T470A	R D I S **A** E I Y
Pep3-N1158A	D K Y F K **A** H T	Pep4-E471A	R D I S T **A** I Y
Pep3-H1159A	D K Y F K N **A** T	Pep4-I472A	R D I S T E **A** Y
Pep3-T1160A	D K Y F K N H **A**	Pep4-Y4T3A	R D I S T E I **A**

Fig. S8. The layout of the peptide microarray of 4 significant epitopes. (**A**) Quality control of the peptide microarray. All the peptides were chemically synthesized and conjugated to BSA. The BSA conjugated peptides were printed and immobilized on the microarray. The microarray was stained with an anti-BSA antibody and followed with a Cy5 conjugated secondary antibody. (**B**) The layout of the peptide microarray. (**C**) The sequences of the peptides that were included on the peptide microarray.

A

Patient No.	epitope-KFLPFQQF		epitope-LTPTWRVY		epitope-RDISTEIY	
	AbMap	Microarray	AbMap	Microarray	AbMap	Microarray
P1	√	√	N.D.	√	N.D.	√
P42	√	√	√	N.D.	√	√
P44	√	√	√	√	N.D.	√
P45	√	√	√	√	N.D.	√
P28	√	√	√	√	N.D.	√

B

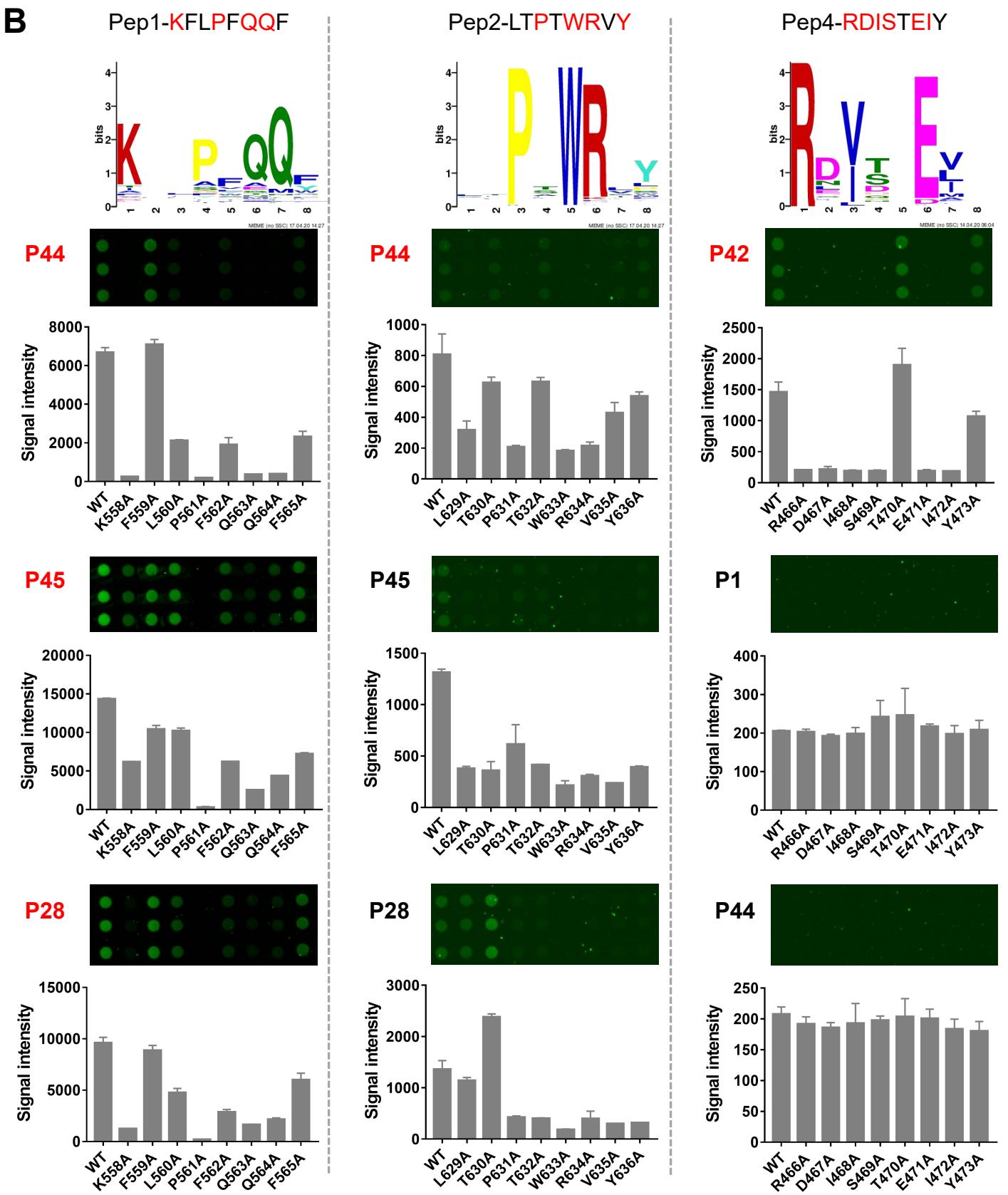


Fig. S9

Fig. S9. Validation of critical residues in epitopes by a peptide microarray. (**A**) Selected samples were subjected for peptide microarray validation. The summary of the validation results. (**B**) The microarray results of selected samples against three representative epitopes. Sample with red label indicates the corresponding epitope was identified from this sample by AbMap. While other samples were included as negative controls for corresponding epitopes.

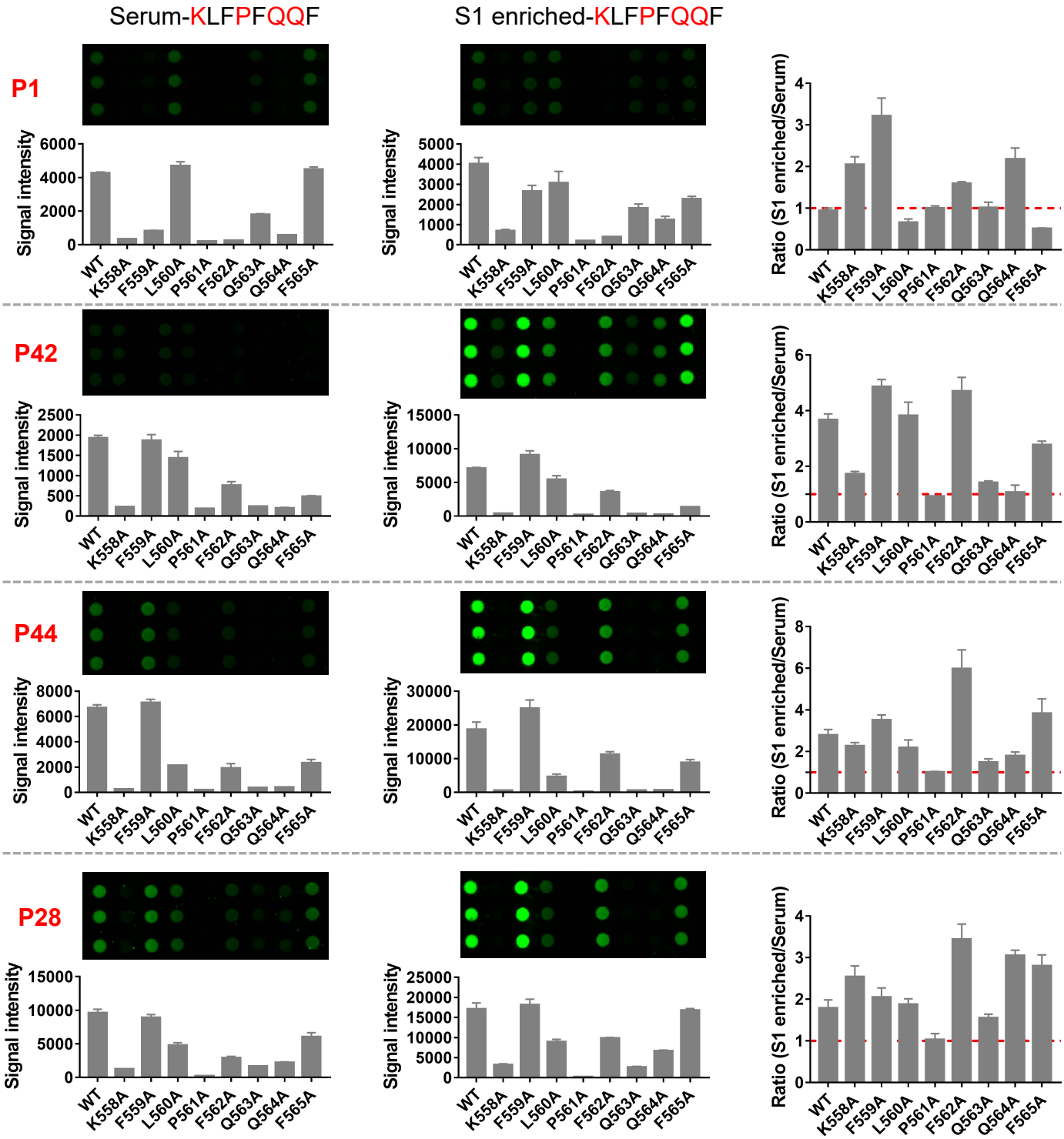


Fig. S10. Validation of epitope identification procedures by a peptide microarray for more samples: the comparison of sera and protein enriched antibodies. Peptide 1 (Pep 1), which corresponds to Epitope-S1-8 (KLFPFQQF) was selected as an example. Sample with red label indicates the corresponding epitope was identified from this sample by AbMap. The ratio of signal_intensity (S1 enriched)/ signal_intensity (serum) for each residue was plotted (right).

Table S1. Serum samples tested in this study.

COVID-19 patients		n = 55
Gender	Male	27
	Female	28
Age	Mean ± s.d.	41.5 ± 14.8
	< 60	48
	≥ 60	7
Severity	Mild	55
	Severe and critical	0
Outcome	Recovered	55
	Death	0
Days after symptom onset for sampling		27.5 ± 7.7
Control group		n = 25
Healthy controls		5
Lung cancer		20
Gender	Male	15
	Female	10
Age	Mean ± s.d.	51.7 ± 12.7
	< 60	18
	≥ 60	7