High-Specificity Targets in SARS-CoV-2 N Protein for Serological Antibody Detection

Jianhai Yu
Southern Medical University
Zhiran Qin
Southern Medical University
Xiaoen He
Southern Medical University
Xuling Liu
Southern Medical University
Jinxiu Yao
Laboratory of people's Hospital JiangCheng Disrtict Yangjiang City
Xuan Zhou
Guangdong Second Provincial General Hospital
Kun Wen
Southern Medical University Zhujiang Hospital
Nan Yu
Southern Medical University Zhujiang Hospital
Qinghua Wu
Southern Medical University
Weiwei Xiao
Southern Medical University
Li Zhu
Southern Medical University
Chengsong Wan
Southern Medical University
Bao Zhang
Southern Medical University
Wei Zhao (zhaowei@smu.edu.cn)
Southern Medical University

Research article

Keywords: Novel coronavirus (SARS-CoV-2), serological detection, cross-reactivity, epitope

Posted Date: September 8th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-64466/v1

Abstract

Background

The current coronavirus disease (COVID-19) pandemic has created a pressing need to diagnose and screen a large number of close contacts of confirmed and suspected cases. Numerous nucleic acid detection kits are being rapidly developed and approved for viral etiological diagnosis; however, these are limited by the number of false negatives produced in clinical practice. Therefore, there is an urgent need to establish serological detection methods to serve as supplementary diagnostics.

Methods

We (1) performed a conservation and specificity analysis of the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) nucleocapsid (N) protein, which is the main target of serological diagnosis; (2) integrated various B-cell epitope prediction methods to obtain possible dominant epitope regions for the N protein; (3) applied ELISA to analyze differences in the serological antibody levels for different epitopes; and (4) identified N protein epitopes for IgG and IgM with high specificity.

Results

SARS-CoV-2 strains showed low mutation rates for the N protein, and the construction of a phylogeny was a good characterization of its molecular evolutionary lineage in relation to other coronaviruses. SARS-CoV-2 showed the closest genetic relationship with SARS-CoV, which showed multiple consecutive long conserved regions at the amino acid level, but differed substantially from other coronaviruses. Tests targeting the SARS-CoV-2 N protein produced strong positive results in SARS-CoV patients in recovery. Of the five epitope dominant regions, using N18-39 and N183-197 for IgG and IgM detection, respectively, can effectively overcome the limitations of cross-reactivity.

Conclusions

The patients infected with both SARS viruses may exhibit cross-reactivity when using the N protein for antibody detection. However, there are regions of the N protein that can be used for antibody detection and some of these regions showed good specificity even between SARS-CoV-2 and SARS-CoV, and the antibody levels detected were consistent with those detected by the complete N protein. These findings provide a basis for serological diagnosis of SARS-CoV-2 patients, and research ideas for developing vaccines.

Introduction

An acute infectious pneumonia, i.e., coronavirus disease (COVID-19), has erupted across the world. Patients commonly present with pneumonia and chest CT abnormalities as primary symptoms, followed by acute cardiac injury and secondary infection as complications, and even death in serious cases [1]. The pathogen is a novel coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which is the seventh coronavirus in human history that has infected humans [2]. Although the transmission rate of SARS-CoV-2 is yet to be confirmed, it has rampaged across the globe at an alarming rate. On January 30, 2020, the World Health Organization (WHO) declared COVID-19 to be a public health emergency of international concern. As of August 6, 2020, this disease has spread to 216 countries and regions, with nearly 18.61 million cases reported and more than 702,642 deaths [3, 4].

Nucleic acid tests of nasopharyngeal swabs have been recommended for the clinical diagnosis of SARS-CoV-2, and are currently the gold standard for confirming SARS-CoV-2 infection [5–7]. However, the use of nucleic acid tests in epidemic control and clinical diagnosis is limited by their false-negative results. In some cases, nasopharyngeal swabs have given multiple negative results, which has necessitated repeated testing, and some cases could only be confirmed by testing using alveolar lavage fluid. This has undoubtedly increased the difficulties encountered during patient admission, quarantine of close contacts, and discharge. Using antibody testing as a supplementary diagnosis is an effective solution to this problem, especially for suspected patients with consistent negative results from nucleic acid testing. In such patients, the probability of detecting antibodies would be greater than with nucleic acid tests, which would help to reduce the risk of missed detection [8, 9]. Studies have shown that the nucleocapsid (N) protein of coronaviruses has strong antigenicity and plays an important role in inducing the host immune response during SARS-CoV-2 infection. Moreover, it has been widely applied as the main target in the diagnosis of SARS-CoV infection [10–13]. At present, ELISA diagnostic kits for the SARS-CoV-2 N protein or its high-specificity region, have been developed and applied, but related studies and reports are scarce [14]. Several issues have restricted its application in clinical practice, including (1) whether pre-existing N protein antibodies in persons infected with other coronaviruses, especially SARS-CoV, could cross-react with the SARS-CoV-2 N protein, thus affecting the accuracy of diagnostic results; (2) whether the detection level of the N protein high-specificity region is consistent with that of the complete N protein; and (3) the selection of the site for the N-protein high-specificity region. Thus, there is an urgent need to clarify the specificity of the SARS-CoV-2 N protein, as the main target of ELISA antibody testing, and to determine a high-specificity region that can specifically and efficiently determine antibody levels. Given the spread of the epidemic at the global scale, this is of critical importance for the large-scale supplementary diagnosis of suspected patients, close contacts of known cases, and even patients with false negatives.

In this study, the SARS-CoV-2 N protein coding genome was downloaded from a public database to analyze its conservation, while its specificity was analyzed by incorporating other coronaviruses, including Bat-CoV, SARS-CoV, MERS-CoV, HCoV-229E, and HCoV-OC43. B-cell epitope prediction software and online servers were integrated to obtain possible high-specificity epitope regions for the SARS-CoV-2 N protein. Finally, clinical serological testing was performed using ELISA to identify the epitope regions suitable for diagnosing IgM and IgG levels. Our findings provide a basis for elucidating the target of the SARS-CoV-2 N protein involved in the host immune response.

Materials And Methods

Collection of samples and epidemiological information

A total of 22 serum samples were collected from 19 COVID-19 patients, 19 of which were collected during hospitalization. The other three serum samples were from patients who were discharged after qPCR testing of throat swabs turned from positive to negative but reverted to positive when followed up. Seven patients were sampled when they were admitted to the Second People's Hospital of Guangdong Province. Twelve patients were sampled when they were admitted to Yangjiang People's Hospital. Eight serum samples of SARS-CoV infected patients were provided by Zhujiang Hospital, Southern Medical University. These samples were collected during the recovery period of patients infected with SARS-CoV in 2003. The epidemiological data for all patients and samples, including age, gender, time of admission, patient symptoms, disease classification, and sampling time, were obtained and collated by the collection unit, then underwent statistical analysis at our laboratory.

Genome alignment and phylogenetic tree analysis of N protein

The SARS-CoV-2 genome sequence was obtained from the GISAID database (https://www.gisaid.org/), while the genomes of the other coronaviruses, including Bat-CoV, SARS-CoV, MERS-CoV, HCoV-229E, and HcoV-OC43, were obtained from the GenBank database (https://www.ncbi.nlm.nih.gov/). The N protein encoding genomes of different viruses from different epidemic years from different countries and regions were downloaded in the FASTA format. Genome translation and sequence alignment was performed using BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html) and CLUSTALW2

(http://www.ebi.ac.uk/Tools/msa/clustalw2/). The phylogenetic tree was built using MEGA6.0 (http://www.megasoftware.net/), using the maximum likelihood method. Based on the analysis of the best-fit model for each data set, the maximum likelihood method was used to construct the phylogenetic tree using the Tamura-Nei model and Gamma distribution with 1,000 iterations.

B-cell epitope prediction for SARS-CoV-2 N protein

B-cell epitope prediction software and online servers were integrated in order to obtain good dominant epitope regions shared by all three methods. Firstly, ABCpred (http://crdd.osdd.net/raghava/abcpred/ABC_submission.html) and BCEpred (http://crdd.osdd.net/raghava/bcepred/bcepred_submission.html), which have scoring systems and present results in a table format, as well as the authoritative database IEDB (http://tools.iedb.org/bcell/), were used to perform B-cell epitope prediction of the SARS-CoV-2 N protein. SOMPA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html), DNAstar's Protean, and Discovery Studio 2.5 (DS2.5) (http://accelrys.com/products/collaborative-science/biovia-discovery-studio/) were used for stepwise verification of the secondary structure and three-dimensional spatial conformation of the dominant regions. The three-dimensional structure of SARS-CoV-2 N protein originates from the Zhangyang Lab and is available for copyrighted free use (https://zhanglab.ccmb.med.umich.edu/COVID-19/). Finally, possible B-cell epitopes were identified.

ELISA testing

SARS-CoV-2 proteins expressed by prokaryotic vectors or synthetic polypeptide fragments of epitope regions were coated overnight and blocked with 5% skimmed milk powder for 2 h. The proteins or polypeptides were then incubated for 1 h with primary antibody solution consisting of patients' serum samples that were diluted 100-fold and subjected to a two-fold gradient dilution, followed by binding with HRP-labelled mouse anti-human IgG and IgM secondary antibodies, respectively. After TMB staining and reaction termination with 2M sulfuric acid stop solution, OD values were read at 450 nm. (Sample OD-negative OD)/(Negative OD-blank OD) $\geq$ 2.1 was taken as the criterion for determining a positive reaction. The highest dilution for a positive test was regarded as the patient's serum antibody titer.

Colloidal gold strip testing

Two colloidal gold kits for emergency detection of SARS-CoV-2 antibodies, approved by the Food and Drug Administration of China, were used in this study and were purchased from Nanjing Vazyme Medical Technology Co., Ltd and Wondfo Biotech Co., Ltd. The former can detect IgG and IgM at the same time, while the latter can only detect IgG. The test cards were placed horizontally, according to the manufacturer's instructions, and serum was added to the sample well. The sample dilution was added and the samples were observed for 10 to 15 min.

Statistical analysis

ELISA test data were collated and statistically analyzed using SPSS 20.0. An independent samples *t*-test was performed to analyze the serum antibody titer of clinical patients detected using SARS-CoV-2 N protein and nine epitope polypeptides.

Results

Conservation and specificity of SARS-CoV-2 N protein

The amino acid sequences of the complete open reading frame for 116 strains of SARS-CoV-2 from different countries and regions were downloaded. The first sequenced Wuhan-Hu-1 strain (GenBank: MN908947.3) was used as the reference sequence for alignment. The viral structural proteins showed low mutation rates in their amino acid sequences during the transmission process, and were highly conserved. Among them, the S and N proteins, which are closely associated with virus-host interactions, were the most conserved, with mutation rates of only 0.94% and 0.95%, respectively (Table 1).

Table 1
Mutation rates of amino acid sequences for structural proteins during SARS-CoV-2 transmission

Protein	Substitutions/ Length	Rate (%)	Functions in SARS-CoV
S	12/1273	0.94%	Induction of neutralizing antibodies, research target of drugs and vaccines
E	1/75	1.33%	Related to envelope formation
M	4/419	1.35%	Related to envelope formation and membrane transport
N	3/222	0.95%	Strong antigenicity, induction of cellular immune response; main target for the establishment of serological diagnostic methods

The N protein sequences of the other coronaviruses were incorporated into the construction of the phylogenetic tree. Coronaviruses of the same lineage with the capacity for human-to-human transmission were clustered on the same branch, of which SARS-CoV and SARS-CoV-2 showed the closest genetic relationship. Bat-CoV, isolated at different times from different bat species, was clustered with SARS-CoV and SARS-CoV-2 (Fig. 1A). The SARS-CoV-2 N protein differed significantly from the weakly pathogenic HCoV-229E and HCoV-OC43, which are commonly found in the population, as well as MERS-CoV, which can cause severe respiratory syndrome, thus showing strong specificity. The sequence was highly homologous with the N protein sequence of BtRa-CoV. Compared with SARS-CoV, which has several long highly conserved regions, the SARS-CoV-2 N protein showed weak conservation (Fig. 1B).

Feasibility analysis of SARS-CoV-2 N protein as a target for clinical serological antibody detection

The results from the ELISA method, with N protein as the target, were highly consistent with the qualitative results of IgG detection using colloidal gold test strips. The ELISA method showed greater sensitivity and could detect antibodies in serum earlier. The IgM test results varied greatly. Only four of the tests were positive and the regularity of symptom onset distribution was poor. With ELISA, positive results could be detected as early as six days after symptom onset. IgM could be detected in mild patients for approximately 17 d. IgM could still be detected in severe patients (15 and 17) even after they had been infected for more than 20 d (Table 2).

Table 2
Epidemiological information and serological antibody test results of samples of COVID-19 patients

Patient	Age	Sex	City	Date	Fever	Cough	Disease	IgG			IgM	
								Vazyme	Wondfo	ELISA	Vazyme	ELISA
1	41	□	GZ	6	√	×	Mild	-	-	-	+	-
2	24	□	YJ	6	×	√	Mild	-	-	-	-	-
3	27	□	YJ	8	×	√	Mild	+	+	+	+	-
4	58	□	GZ	8	×	√	Mild	+	-	-	+	-
5	35	□	GZ	13	×	√	Mild	+	-	-	+	-
6	49	□	GZ	13	×	√	Mild	+	+	+	+	+
7	57	□	GZ	14	√	×	Mild	+	-	+	+	-
8	42	□	GZ	15	√	√	Mild	+	+	+	+	+
9*	28	□	YJ	16	√	×	Mild	+	+	+	+	-
10	61	□	YJ	17	√	×	Mild	+	+	+	-	-
11	56	□	YJ	17	×	√	Mild	+	+	+	+	-
12*	60	□	YJ	17	×	√	Mild	+	+	+	+	-
13	23	□	YJ	19	√	√	Mild	+	+	-	-	-
14	35	□	YJ	19	×	√	Severe	+	+	+	-	-
15	56	□	YJ	20	√	√	Severe	+	+	+	+	+
16	56	□	GZ	21	√	×	Mild	+	+	+	-	-
17	77	□	YJ	21	×	√	Severe	+	+	+	+	-
18	24	□	YJ	21	×	√	Mild	+	+	+	-	-
19*	69	□	YJ	24	√	√	Mild	+	+	+	-	+

“*” represents patients who tested positive using QPCR and ELISA detection after discharge. □: Male. □: Female. GZ: Guangzhou. YJ: Yangjiang.

The mean age of the 19 patients was 46.2 years, and there were more men than women (12/19, 63.2%). More patients presented with cough on admission than fever, with incidences of 73.7% and 47.4%, respectively; even among the three severe cases, only one exhibited a fever (Fig. 2). In terms of the timing of seropositive reactions, IgM appeared earlier than IgG, within the first week after symptom onset. Within the first 6 d of symptom onset, a seropositive reaction could still be detected after an 800-fold dilution of the patient’s serum. The levels of both antibodies increased in the second week, with IgG showing a more significant increase. Thereafter, serum IgG maintained a high antibody titer until the recovery period (after the third week of symptom onset), whereas IgM began to decrease gradually. In the serum of nine patients tested after 19 d, IgM could only be detected in the serum samples collected from two severe cases at 20 and 21 d, whereas serum samples of the remaining patients, subjected to a 100-fold dilution, tested negative (Fig. 2A and B). In addition, there were no significant differences in the serum IgG and IgM levels between the nine mild cases and three severe cases during the recovery period (Fig. 2C). As for the three mild cases with “re-detectable positives”, their serum IgG levels did not differ significantly between the recovery period during hospitalization and after discharge. The main finding was that IgM became positive in all cases after discharge, and the antibody titer was significantly higher than that of serum levels during recovery (Fig. 2D).

The serum colloidal gold test strip results for eight SARS-CoV patients in the recovery stage were all strongly positive for IgG and all negative for IgM (Fig. 2E). The use of the SARS-CoV-2 N protein for ELISA detection was highly consistent with the test strip results, even when the antibody titer was as high as 10^4 (Fig. 2F).

Epitope prediction for SARS-CoV-2 N protein

ABCpred, BCEpred, and IEDB were used to perform the epitope prediction for the N protein of SARS-CoV-2, and the results of the three software were collated and compared. A total of five dominant epitope regions shared by the three were obtained, namely, N18-39, N183-197 l, N249-266, N276-299, and N365-391 (Fig. 3A and Table 3). Verification of secondary structure parameters showed that the five epitope regions had good hydrophilicity, flexibility, surface accessibility, and antigenicity. Simulations of the N protein three-dimensional structure showed that the five regions were located on the surface of the protein. N18-39 and N183-197 were free random-coil loop structures (Fig. 3B), whereas N249-266, N276-299, and N365-391 showed similar random coil + α helix structural characteristics (Fig. 3C and D), and all of them showed good antigenicity.

Table 3
Synthetic polypeptides corresponding to the five epitope regions in SARS-CoV-2 and SARS-CoV N proteins

Polypeptide	Region	AA	Length
SARS-CoV			
P1	N19-40	GGPTDSTDNNQNGGRNGARPKQ	22
P2	N184-198	SSRSSSRSRGNSRNS	15
P3	N277-300	RRGPEQTQGNFGDQDLIRQGTDYK	24
P4	N366-392	PTEPKKDKKKKTDEAQPLPQRQKKQPT	27
SARS-CoV-2			
P5	N18-39	GGPSDSTGSNQNGERSGARSKQ	22
P6	N183-197	SSRSSSRSRNSSRNS	15
P7	N249-266	KSAAEASKKPRQKRTATK	18
P8	N276-299	RRGPEQTQGNFGDQELIRQGTDYK	24
P9	N365-391	PTEPKKDKKKKADETQALPQRQKKQQT	26
The Genbank Accession number MN908947.3 was used for SARS-CoV-2, while NC004718.3 was used for SARS-CoV.			

The N proteins of SARS-CoV-2 and SARS-CoV showed a number of long conserved regions. To analyze the effects of pre-existing antibodies in the SARS-CoV infected population on the detection of SARS-CoV-2 when using the five epitope regions for serological diagnosis, sequence alignment was performed between the five epitope regions of the SARS-CoV-2 and the SARS-CoV N proteins. Our results showed that N18-39 and N365-391 differed significantly between SARS-CoV-2 and SARS-CoV and showed strong specificity. N183-197 and N276-299 showed two and one differential sites, respectively, but still had more than 10 consecutive amino acid residues that were fully conserved. The N249-266 region was fully conserved in both (Fig. 3E). Therefore, a total of nine polypeptides corresponding to the five epitope regions of the SARS-CoV-2 and SARS-CoV N proteins were synthesized and used for verification in subsequent ELISA testing (Fig. 3).

Feasibility analysis of using five epitopes for clinical serological ELISA testing

Based on the positive ELISA results for IgG and IgM in the serum samples of the 19 patients, the IgM antibody testing for N18-39 and N365-391 were negative. The other three polypeptides showed positive reactions for IgG and IgM. (Fig. 4 and Fig. 5).

For the detection of IgG, the five-polypeptide epitope of the SARS-CoV-2 N protein showed specific positive reactions for IgG in serum. However, the N18-39 polypeptide alone did not show any significant difference with the N protein in the detection of IgG antibody levels in the serum samples of the 19 patients. The detection levels of the other four polypeptides were lower than that of the N protein (Fig. 4A). In addition, the detection level of the N18-39 polypeptide in the corresponding region of the SARS-CoV N protein (N19-40) was the lowest among all four corresponding polypeptides. Only 4 of 19 samples tested positive, and the antibody titer was as low as 100-fold (Fig. 4B). At the same time, similar results were obtained for SARS-CoV infected patients, with high antibody levels during the recovery period. Compared with the high IgG antibody titer detected in the N19-40 peptide of SARS-CoV, the N protein of SARS-CoV-2 showed a higher level of antibody titer detection, suggesting that there may be a large number of N proteins targeting different sites of cross-reactive antibodies that can interfere with ELISA test results. Among the four peptides showing differences in the amino acids of the two viruses, the detection level of the N18-39 polypeptide in the serum of eight patients was positive, the difference between the samples was the most stable and the titer was the lowest. Only a 100–800 fold dilution produced positive results (Fig. 4C). This suggests that the N18-39 polypeptide shows good specificity and accuracy when used to detect the IgG antibody levels in the serum of COVID-19 patients.

For the detection of IgM, although a positive IgM reaction could be detected using the N278-299 polypeptide, the antibody level detected was significantly lower than that with the N protein. The detection levels of N183-197 and N249-266 showed good agreement with that of the N protein, with no statistically significant differences (Fig. 5A). There are two differential sites in the amino acid sequence of the SARS-CoV-2 N183-197 region and the corresponding N184-198 region in SARS-CoV, causing N184-198 to produce negative results when detecting the serum IgM of COVID-19 patients. This implies that the SARS-CoV-2 N183-197 polypeptide has good specificity. In contrast, the N249-266 polypeptide sequence is identical between the viruses, and the presence of IgM antibodies can be detected in the sera of patients infected by either virus (Fig. 5B).

Discussion

The genome sequence data show that SARS-CoV-2 is a member of the betacoronavirus genus and belongs to the subgenus that includes SARS-CoV (*Sarbecovirus*); it shows approximately 79% similarity to SARS-CoV at the nucleotide level. The spike (S) protein, which is a key surface glycoprotein that interacts with the host cell receptor, and N protein, which has strong antigenicity, are capable of inducing cellular immune response, and are the main target of serological diagnosis, also show a high degree of similarity between viruses [15–17]. This is consistent with the results of the conservation and specificity analysis conducted in this study. During the evolution of SARS-CoV-2, the mutation rate of its N protein was relatively low, and the phylogenetic tree shows a good characterization of its molecular evolutionary lineage with other coronaviruses. It has the closest genetic relationship with SARS-CoV, with multiple continuous long conserved regions in the N protein at the amino acid level, suggesting that the two may have cross-reactivity when using the N protein for antibody detection. The SARS-CoV epidemic in 2003 did not spread on a global scale [18]. However, due to limitations at that time, such as inadequate detection methods, deficient diagnostic standards, lack of technical advances, and imperfect prevention and control regulations for public health incidents, there may have been missed diagnoses or misdiagnoses of suspected cases or close contacts. Therefore, in some of the hardest hit regions, such as Guangdong in China, Canada, and Singapore, etc. there may be immunized populations with pre-existing SARS-CoV antibodies. It is still unclear whether these antibodies have disappeared after more than ten years [19, 20]. Nevertheless, it is still necessary to be vigilant to the threat of a SARS-CoV resurgence, which could occur if suitable viral hosts or transmission routes are established. This scenario is not unlike in case of the SARS-CoV-2 epidemic, where the virus may already have existed in certain wild animals, such as bats or pangolins [21–23], but spread to humans via suitable transmission routes, thereby causing a global pandemic.

At present, experimental data on the expression patterns of SARS-CoV-2 antibodies are scarce, and our understanding at this stage is largely based on research on SARS-CoV. For the detection of IgG and IgM antibodies against the SARS-CoV N protein, the average time taken for IgG to produce a positive result was one day lesser than that for IgM, and could be

detected as early as four days after onset [24–26]. Our findings showed that in the serum samples of two patients collected six days after onset, which was the earliest collection time in this study, IgM could be detected, but IgG was negative. However, as the disease progressed, both IgG and IgM could be detected, and the antibody titer of IgG increased until it was higher than that of IgM. During the recovery period, IgG remained at a high level, but IgM gradually disappeared. Kristi et al. found in a small-sample survey that in the first serum sample collected when patients were admitted to the hospital, virus-specific IgM and IgG titers were relatively low, or lower than the detection limit. However, by the fifth day, nearly all patients showed positive or elevated antibody levels [27]. Detecting such early positive reactions for IgG and IgM and plotting the trend of antibody levels are of great significance for clarifying the infection and disease progression of SARS-CoV-2.

Studies have shown that SARS-CoV IgG antibodies can remain at high levels in patients for a long period of time, such that among 257 SARS patients, antibody titers increased steadily for 4 to 6 months after the onset of SARS. The titers of most cases peaked at 6 months. Even in cases where the antibody titer decreased rapidly, ELISA tests yielded positive results even at 48 months, which indicates the possibility that specific antibodies may be present for a long time in SARS patients [28]. The latest research has found that anti-SARS-CoV IgG can last up to 12 years, although IgG titers generally peaked in 2004. Until 2015, the IgG titers of medical workers infected with SARS-CoV remained at a very high level [29]. When SARS-CoV-2 antibody-approved colloidal gold test strips and the ELISA method targeting the SARS-CoV-2 N protein were used to test SARS-CoV patients in the recovery period, strong positive results were revealed. This result suggests that the N protein as a target for antibody detection shows a strong cross-reaction in SARS-CoV infected people with pre-existing immunity, which will affect the accuracy of the results. There is an urgent need to clarify the highly specific regions of the N proteins of the two viruses to solve this problem. In this study, of the five epitope regions obtained by epitope prediction, the use of two high-specificity epitope regions, N18-39 and N183-197, for IgG and IgM detection, respectively, effectively overcame the limitations of cross-reactivity. They showed good specificity even between SARS-CoV-2 and SARS-CoV, and the antibody levels detected were consistent with those detected by the complete N protein. These findings indicate that not all highly conserved epitopes can be used alone in ELISA to detect antibody levels in the serum of patients. It is worth noting that even the most specific N18-39 polypeptide, when testing SARS-CoV patient serum with high IgG antibody levels, also detected low levels of IgG. When testing populations with pre-existing immunity for SARS-CoV, high-sensitivity ELISA detection methods may produce false positives. However, after more than 15 years, the levels of IgG-specific antibodies in people that were infected with SARS-CoV will be significantly reduced; whether they can be detected requires further exploration.

Conclusions

Nucleic acid tests are used to confirm SARS-CoV-2 infection, but these tests can produce false-negative results. Antibody testing can be a useful supplementary method. The N protein of coronaviruses plays important roles in inducing the host immune response and are important targets for antibody testing. However, we do not know whether people previously infected with other coronaviruses, especially SARS, may still have high antibody levels and produce positive results for antibody tests targeting the N protein. In this study, we examined the N proteins of SARS-CoV-2 and other coronaviruses, and tested whether the N protein can be useful for antibody testing. We found that SARS-CoV patients in the recovery stage all produced strong positive test results. However, there are regions of the N protein that can be used for antibody detection and some of these showed good specificity even between SARS-CoV-2 and SARS-CoV, and the antibody levels detected were consistent with those detected by the complete N protein. Our study provides new insights into antibody testing for COVID-19. Given the spread of the epidemic at the global scale, this is of critical importance for the large-scale supplementary diagnosis of suspected patients, close contacts of known cases, and even patients with false negatives.

List Of Abbreviations

Abbreviation	Full name
COVID-19	Current coronavirus disease
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
N	Nucleocapsid
WHO	World health organization
DS2.5	Discovery studio 2.5
S	Spike

Declarations

Ethics approval and consent to participate

As this study involved collecting blood samples from patients, the purpose of the study was explained to all participating COVID-19 patients, and informed consent was obtained in writing.The collection of clinical samples and epidemiological data were reviewed and approved by the ethics review committee of the Southern Medical University and performed in accordance with approved guidelines.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in this published article.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the [National Key R&D Program of China] under Grant [number 2018YFC1602206]; [Yangjiang Science and Technology Program key projects] under Grant [number 2019010]; [Guangzhou Science and Technology Program key projects] under Grant [number 201803040006] and [Guangdong Science and Technology Program key projects] under Grant [number 2018B020207006].

Authors' contributions

JY, ZQ, XH and XL conceived the project, designed the experiments, undertook experiments and wrote the manuscript; JY, XZ, KW, and NY provided the samples, performed the experiments and analysed the data; QW, WX, LZ and CW provided valuable structural insight and helped writing the manuscript. BZ participated in the construction and improvement of the research, and guided the experimental process and article writing. WZ participated in the whole research process and provided financial support.

Acknowledgements

We would like to thank Editage (www.editage.cn) for English language editing, and Nanjing Vazyme Medical Technology Co., Ltd for providing two colloidal gold kits for emergency detection of SARS-CoV-2 antibodies.

References

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395: 497-506. doi: 10.1016/S0140-6736(20)30183-5.
2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382: 727-733. doi: 10.1056/NEJMoa2001017.
3. Coronavirus disease (COVID-19) outbreak situation. 2020. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019
4. Coronavirus disease (COVID-19) situation dashboard. 2020. Available from: https://covid19.who.int/
5. Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases. Available from: https://www.who.int/publications-detail/laboratory-testing-for-2019-novel-coronavirus-in-suspected-human-cases-20200117(Jan 17, 2020). Accessed February 11, 2020
6. Pan Y, Zhang D, Yang P, Poon LLM, Wang Q. Viral load of SARS-CoV-2 in clinical samples. Lancet 2020;20: 411-412. doi: 1016/S1473-3099(20)30113-4
7. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, Tan W. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA doi: 10.1001/jama.2020.3786.
8. Guo L, Ren L, Yang S, Xiao M, Chang D, Yang F, et al. Profiling early humoral response to diagnose novel Coronavirus disease (COVID-19). Clin Infect Dis. 2020. doi: 10.1093/cid/ciaa310.
9. Dong X, Cao YY, Lu XX, Zhang JJ, Du H, Yan YQ, et al. Eleven faces of Coronavirus disease 2019. Allergy 2020, doi: 10.1111/all.14289.
10. Wang L, Wu YS, Tang YP, Li M. Detection of SARS-associated coronavirus N protein by time-resolved fluoroimmunoassay. Di Yi Jun Yi Da Xue Bao, 2005;25: 429-431 [Chinese].
11. Huang LR, Chiu CM, Yeh SH, Huang WH, Hsueh PR, Yang WZ. Evaluation of antibody responses against SARS coronaviral nucleocapsid or spike proteins by immunoblotting or ELISA. J Med Virol. 2004;73: 338-346. doi: 1002/jmv.20096.
12. Shang B, Wang XY, Yuan JW, Vabret A, Wu XD, Yang RF, et al. Characterization and application of monoclonal antibodies against N protein of SARS-coronavirus. Biochem Biophys Res Commun. 2005;336: 110-117. doi: 1016/j.bbrc.2005.08.032.
13. Wang J, Ji J, Ye J, Zhao X, Wen J, Li W, et al. The structure analysis and antigenicity study of the N protein of SARS-CoV. Genom Proteom Bioinf. 2003;1: 145-154. https://doi.org/10.1016/S1672-0229(03)01018-0.
14. Li Z, Yi Y, Luo X, Xiong N, Liu Y, Li S, et al. Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis. J Med Virol. doi: 10.1002/jm25727.
15. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 2020;395: 565-574. doi: 10.1016/S0140-6736(20)30251-8.
16. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. Nature 2020;579: 265-269. https://doi.org/10.1038/s41586-020-2008-3.
17. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382: 727-733. https://doi.org/10.1056/NEJMoa2001017.
18. Chinese Medical Association. Consensus for the management of severe acute respiratory syndrome. Chin Med J. 2003;116: 1603-1635.
19. Xie L, Liu Y, Fan B, Xiao Y, Tian Q, Chen L, et al. Dynamic changes of serum SARS-coronavirus IgG, pulmonary function and radiography in patients recovering from SARS after hospital discharge. Respir Res. 2005;6: 5. doi: 1186/1465-9921-6-5.

20. Tan YS, Hong CY, Chong PN, Tan ES, Lew YJ, Lin RT. Revisiting upper respiratory tract infection (URTI) in adults in the light of the current severe acute respiratory syndrome (SARS) situation in Singapore. Ann Acad Med. Singapore 2003;32: S27-29.
21. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 2020;579: 270-273.
22. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. Nat Rev Microbiol. 2019;17: 181-192. doi: 1038/s41579-018-0118-9.
23. Lam TTY, Shum MHH, Zhu HC, Tong YG, Ni XB, Liao, YS, et al. Identification of 2019-nCoV related coronaviruses in Malayan pangolins in southern China. BioRXiv 2020. doi: 10.1101/2020.02.13.945485.
24. Hsueh PR, Huang LM, Chen PJ, Kao CL, Yang PC. Chronological evolution of IgM, IgA, IgG and neutralisation antibodies after infection with SARS-associated coronavirus. Clin Microbiol Infect. 2004;10: 1062-1066. doi: 10.1111/j.1469-0691.2004.01009.x.
25. Woo PC, Lau SK, Wong BH, Chan KH, Chu CM, Tsoi HW, et al. Longitudinal profile of immunoglobulin G (IgG), IgM, and IgA antibodies against the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein in patients with pneumonia due to the SARS coronavirus. Clin Diagn Lab Immunol. 2004;11: 665-668. doi: 10.1128/CDLI.11.4.665-668.2004.
26. Lee N, Chan PK, Ip M, Wong E, Ho J, Ho C, et al. Anti-SARS-CoV IgG response in relation to disease severity of severe acute respiratory syndrome. J Clin Virol. 2006;35: 179-184. doi: 10.1016/j.jcv.2005.07.005.
27. Koenig KL, Beÿ CK, McDonald EC. 2019-nCoV: The identify-isolate-inform (3I) tool applied to a novel emerging coronavirus. West J Emerg Med. 2020;21: 184-190. doi: 10.5811/westjem.2020.1.46760.
28. Shi YL, Li LH, Chu XW, Chen Q, Wang YL, Ma QJ, et al. Seroepidemiological study on SARS-CoV IgG antibodies of different populations from several areas. Int J Biomed Sci. 2005;1: 72-75.
29. Guo X, Guo Z, Duan C, Chen Z, Wang G, Lu Y, et al. Long-term persistence of IgG antibodies in SARS-CoV infected healthcare workers. Med 2020. doi: 10.1101/2020.02.12.20021386.

Figures

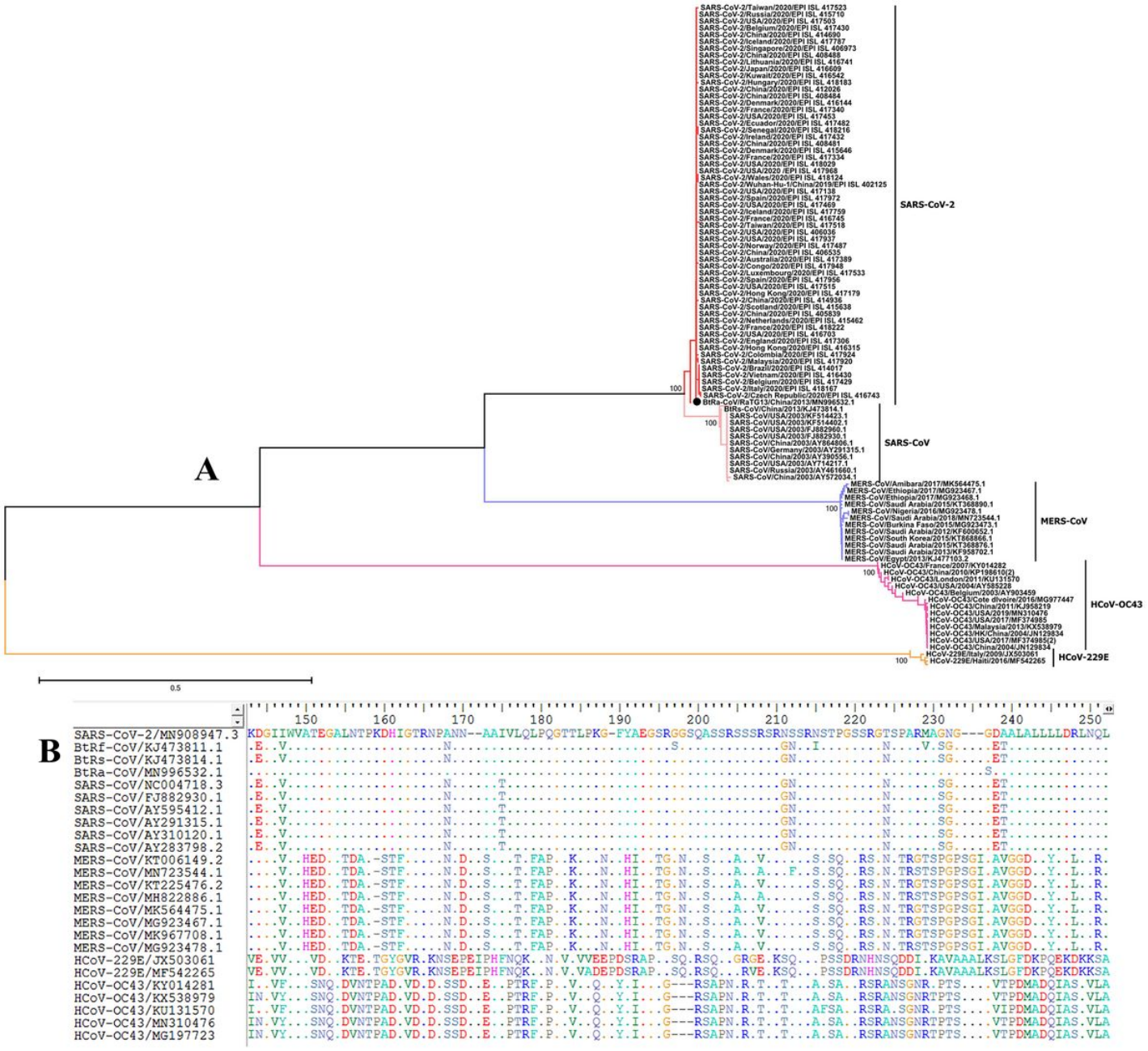

Figure 1

Specificity analysis of the SARS-CoV-2 N protein. A: Phylogenetic tree of the N protein from SARS-CoV-2, Bat-CoV, SARS-CoV, MERS-CoV, HCoV-229E, and HCoV-OC43 from outbreaks in different epidemic years in different countries and regions. B: N protein sequence alignment of SARS-CoV-2, Bat-CoV, SARS-CoV, MERS-CoV, HCoV-229E, and HCoV-OC43

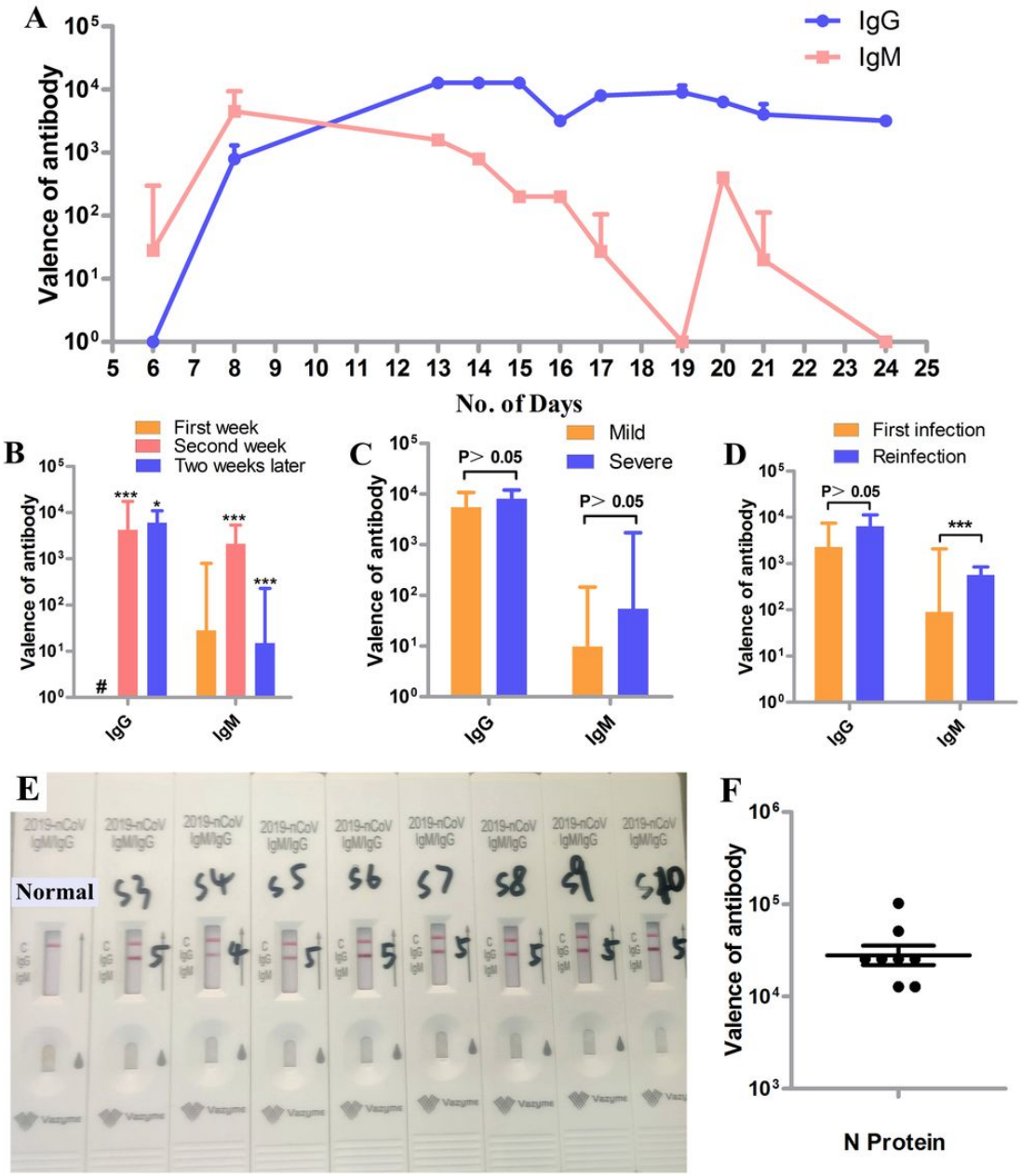

Figure 2

Serum antibody titer of clinical patients for SARS-CoV-2 N protein detected using ELISA. A: During disease progression, patients' serum IgG and IgM antibody titers showed dynamic changes. A detection value of 100 indicates that the ELISA result was negative for serum diluted 100-fold. B: Differences in the serum IgG and IgM antibody titers of COVID-19 patients in the first, second, and two weeks after onset. "#" indicates tested negative for serum diluted 100-fold; "*" and "***" indicate P values of <0.05 and <0.01, respectively, for the statistical analysis between the latter and former periods. C: Differences in serum IgG and IgM antibody titers between mild and severe COVID-19 cases two weeks after onset. D: Comparison of serum IgG and IgM antibody levels in patients with "re-detectable positives" during hospitalization and after discharge, "***" indicates P<0.01. E: SARS-CoV-2 antibody detection kit for testing the serum of SARS-CoV infected patients during the recovery period. F: ELISA method targeting the SARS-CoV-2 N protein to test the serum of SARS-CoV infected patients during recovery.

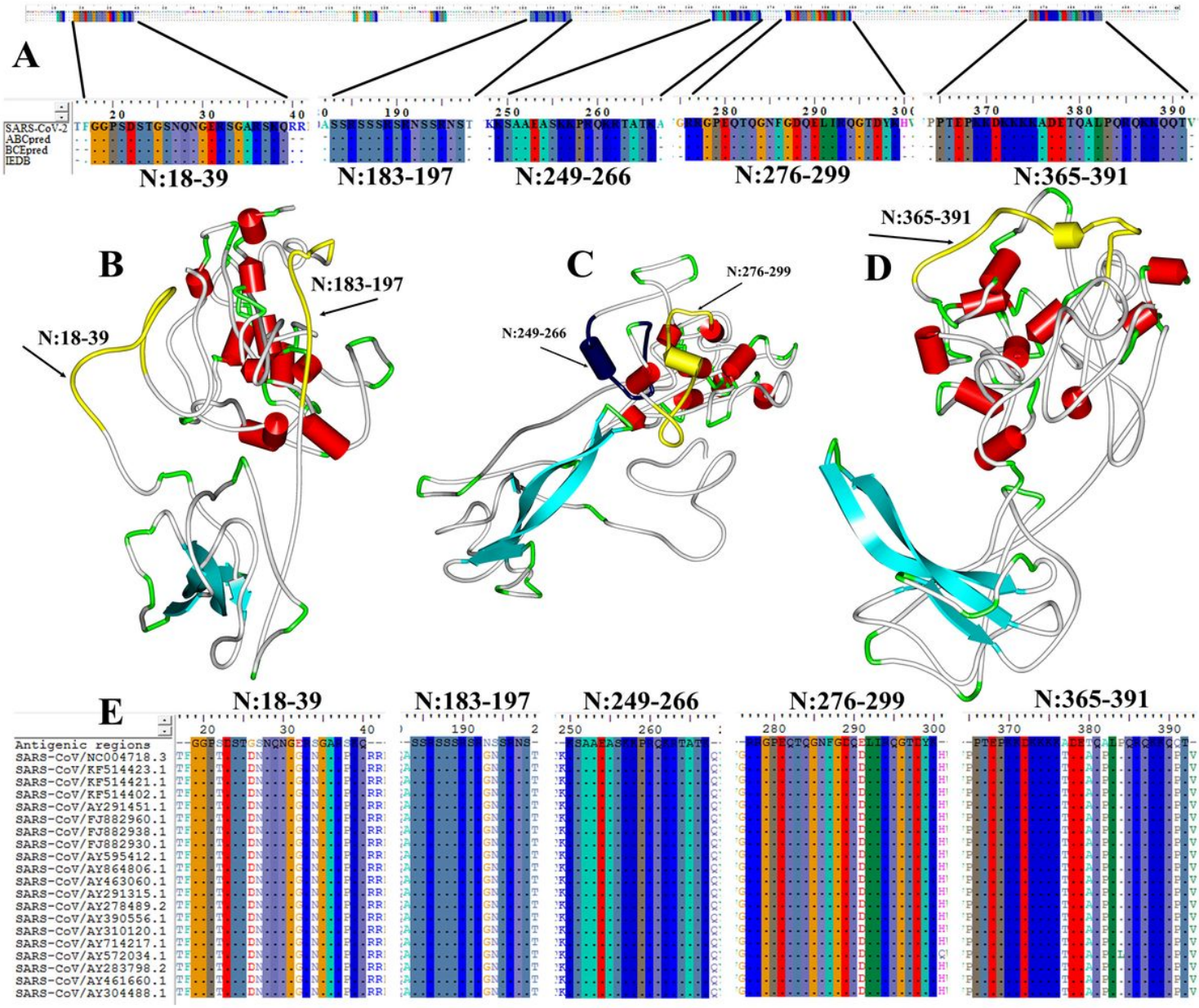

Figure 3

Epitope prediction process for the SARS-CoV-2 N protein. A: Alignment of epitope prediction results from ABCpred, BCEpred, and IEDB for SARS-CoV-2 N protein. B-D: Three-dimensional structure of the five epitope regions. E: Alignment of the SARS-CoV N protein sequences to analyze the specificity of the five epitope regions.

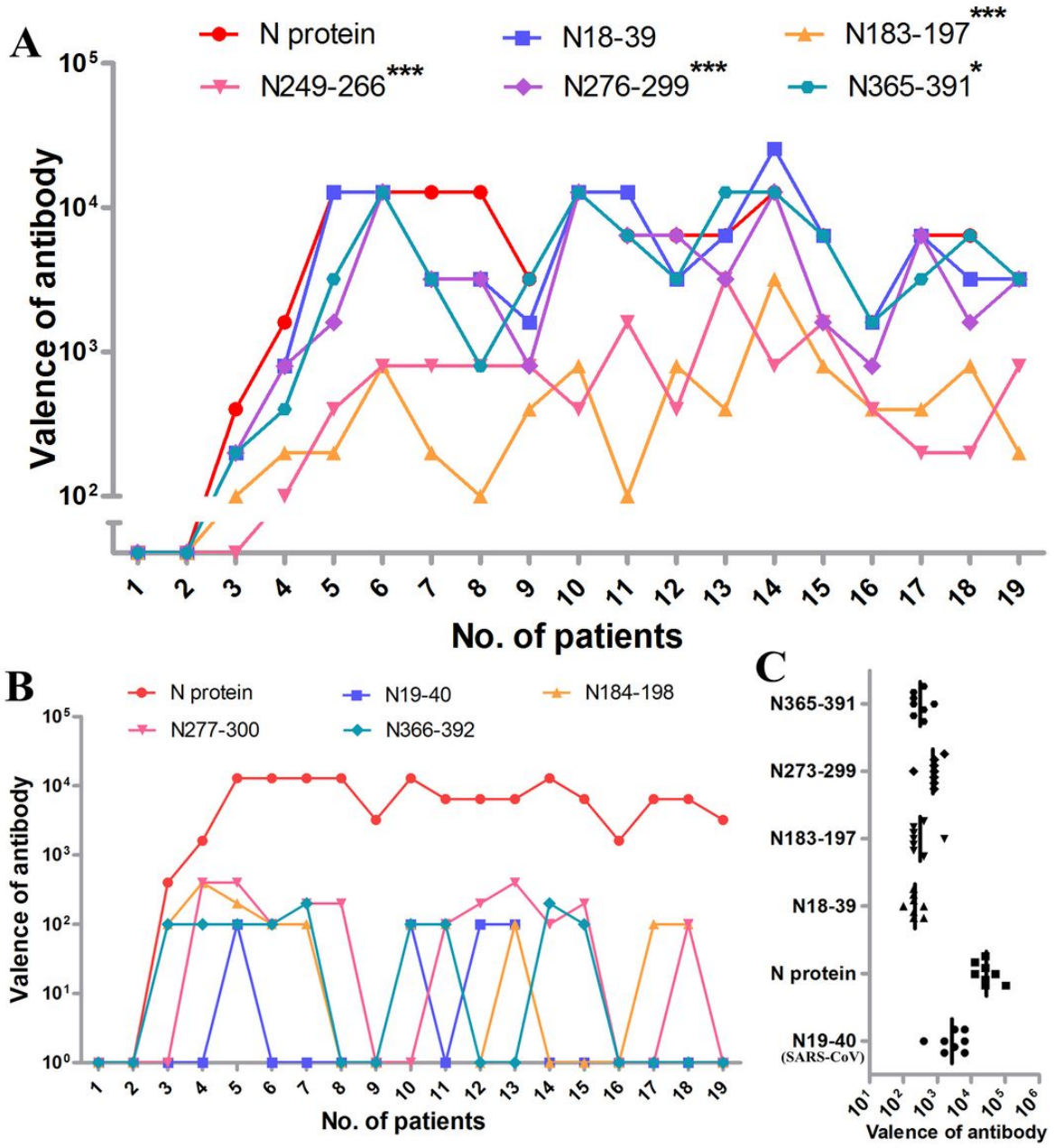

Figure 4

Use of nine polypeptides in ELISA to detect serum IgG levels in patients. A-B: Detection value of 100 indicates that the ELISA result was negative for serum diluted 100-fold. A: Comparison of detection levels between the five epitope polypeptides and N protein of SARS-CoV-2. "*" and "***" indicate P<0.05 and P<0.01, respectively, in the statistical analysis between the polypeptide and N protein detection results. B: Comparison of detection levels between the four polypeptides in corresponding regions and the N protein of SARS-CoV-2. C: Verification of serum cross-reactivity in the recovery stage of SARS-CoV infected patients.

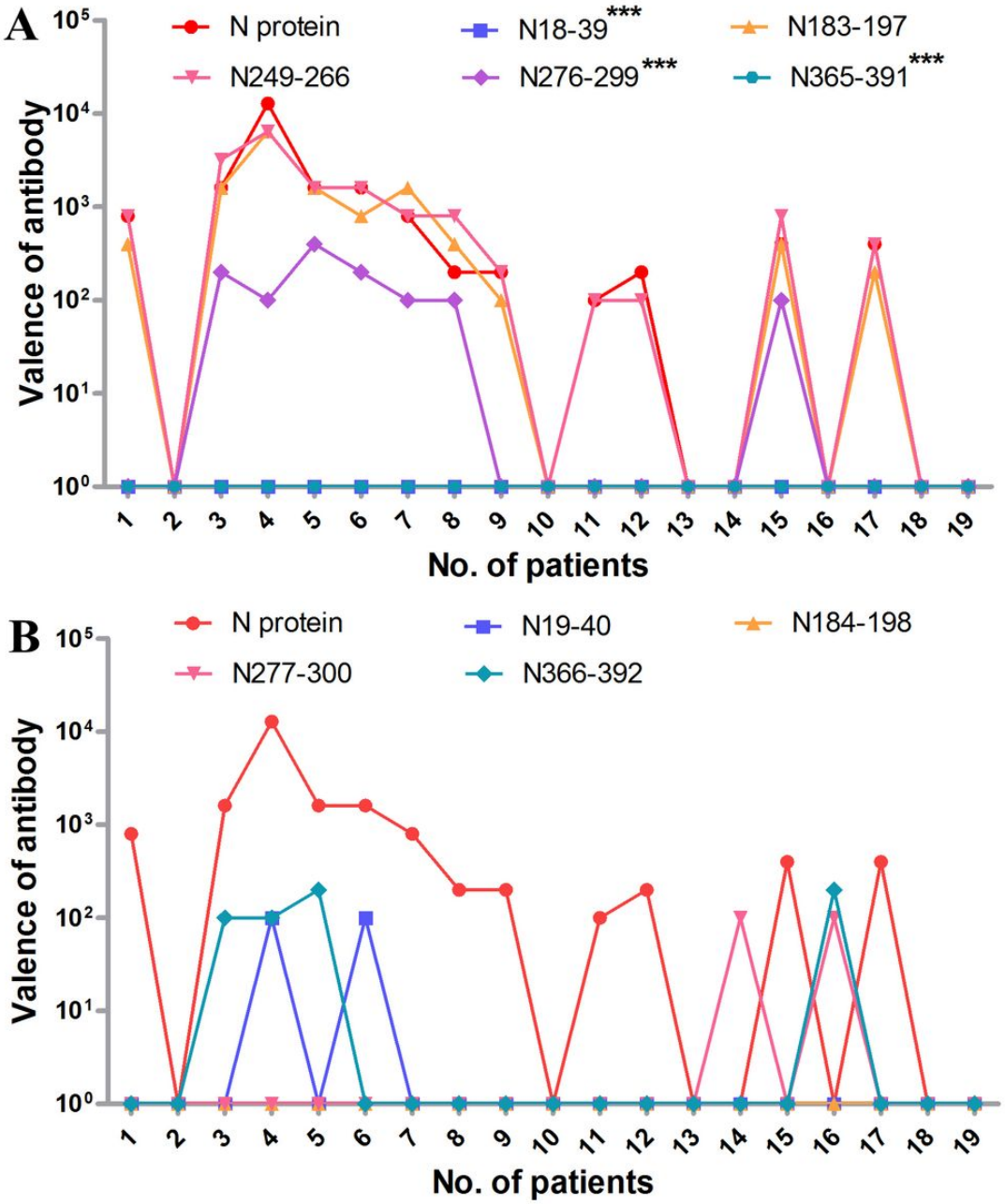

Figure 5

Use of nine polypeptides in ELISA to detect serum IgM levels in patients. A-B: Detection value of 100 indicates that the ELISA result was negative for serum diluted 100-fold. A: Comparison of detection levels between the five epitope polypeptides and N protein of SARS-CoV-2. "***" indicates P<0.01 in the statistical analysis between the polypeptide and N protein detection results. B: Comparison of detection levels between the four polypeptides in corresponding regions and the N protein of SARS-CoV-2.