



CORRESPONDENCE

Novel antibody epitopes dominate the antigenicity of spike glycoprotein in SARS-CoV-2 compared to SARS-CoV

Ming Zheng¹ and Lun Song¹

Cellular & Molecular Immunology (2020) 17:536–538; <https://doi.org/10.1038/s41423-020-0385-z>

Recently, the World Health Organization (WHO) has declared the novel coronavirus (2019-nCoV) outbreak a Public Health Emergency of International Concern (PHEIC),¹ which is now formally named as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).² As of 27 February 2020, a total of 82,178 cases of SARS-CoV-2 infection have been confirmed across the world, with 78,630 cases in China (<https://ncov.dxy.cn/ncovh5/view/pneumonia?source=>). The SARS-CoV-2 has been determined as the seventh member of the coronaviruses infected humans.³ Moreover, similar to severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), the SARS-CoV-2 could also cause severe and fatal illness.³ Since the SARS-CoV-2 outbreak, there have been approximately 14,792 clinically severe cases and 2800 dead cases.

Due to the fast spread of SARS-CoV-2 and shortage of specific therapy, many efforts have focused on neutralizing antibody and vaccine development.⁴ Vaccines prevent disease largely depending on inducing neutralizing antibodies against vulnerable epitopes on antigen. Among the structural proteins of coronavirus, the spike glycoprotein contains receptor-binding domain (RBD) to mediate coronavirus entering host cells, which makes spike protein the primary antigenic target of neutralizing antibody and vaccine.⁵

Recently, it has been reported that the genome of SARS-CoV-2 have 79.5% nucleotide sequence identify to that of SARS-CoV.⁶ The genome relatedness indicates the possibilities that pre-clinical drugs against SARS-CoV might be effective to SARS-CoV-2. Also, a recent study was focused on cross-protective epitope between the spike proteins of SARS-CoV-2 and SARS-CoV, and successfully found the cross-protective epitopes in the RBDs of the spike proteins.⁷ Moreover, another study found that the spike RBD of SARS-CoV-2 bound potently to angiotensin-converting enzyme 2 (ACE2), the host cell receptor of SARS-CoV.⁵ However, in spite of the same binding target to ACE2, three of four monoclonal antibodies capable of binding potently to the SRAS-CoV RBD failed to show evident binding to the SARS-CoV-2 RBD.⁵ The limited antibody cross-reactivity suggests the importance to investigate the difference of antibody epitopes between the spike proteins of SARS-CoV and SARS-CoV-2.

In our study, we found the SARS-CoV-2 spike protein had approximately 24.5% amino acid (a.a.) sequence non-conserved to that of SARS-CoV (Supplementary Fig. 1). Because of the divergence of spike proteins, the non-conserved regions of spike proteins might have the main responsibility for the antigenic difference. Thus, to solve the problem, we conducted antibody epitope analysis that focused on the comparison of the conserved

and non-conserved regions of spike glycoproteins between MERS-CoV, SRAS-CoV, and SARS-CoV-2.

The spike proteins of SARS-CoV-2 from Wuhan, Zhejiang, and Guangdong in China and other countries of the United States, France, Australia, and Germany were nearly 100% conserved (Supplementary Fig. 2). Next, alignment and phylogenetic analysis of the amino acid sequences of spike proteins in SARS-CoV-2, MERS-CoV, and SARS-CoVs showed the difference of sequence conservancy (Fig. 1a and Supplementary Fig. 1). As spike proteins of those five SARS-CoVs had approximately 99.5% homologous a.a. sequence (Supplementary Fig. 2), we used SARS-NS1 as a representative SARS-CoV for further analysis.

Currently, bioinformatic approaches of epitope analysis are well-developed and successfully proved to identify both weak and strong epitopes that might be experimentally ignored.⁸ In our study, using antibody epitope bioinformatic tools (Supplementary Materials and Methods), we computed sequence-based antibody epitope scores in spike proteins of MERS-CoV, SARS-CoV, and SARS-CoV-2 (Fig. 1b). The SARS-CoV-2 had significantly lower antibody epitope score compared with MERS-CoV ($p < 0.0001$; Fig. 1c) and significantly higher antibody epitope score compared with SARS-CoV ($p < 0.01$; Fig. 1c), indicating the spike proteins have significantly variable antigenicity. Next, we conducted sequence alignment to acquire the conserved and non-conserved regions of spike proteins (Fig. 1d and Supplementary Fig. 1). Compared with the conserved regions, the non-conserved regions had significantly higher antibody epitope score (Fig. 1e, f), indicating the non-conserved regions of spike proteins are more antigenic.

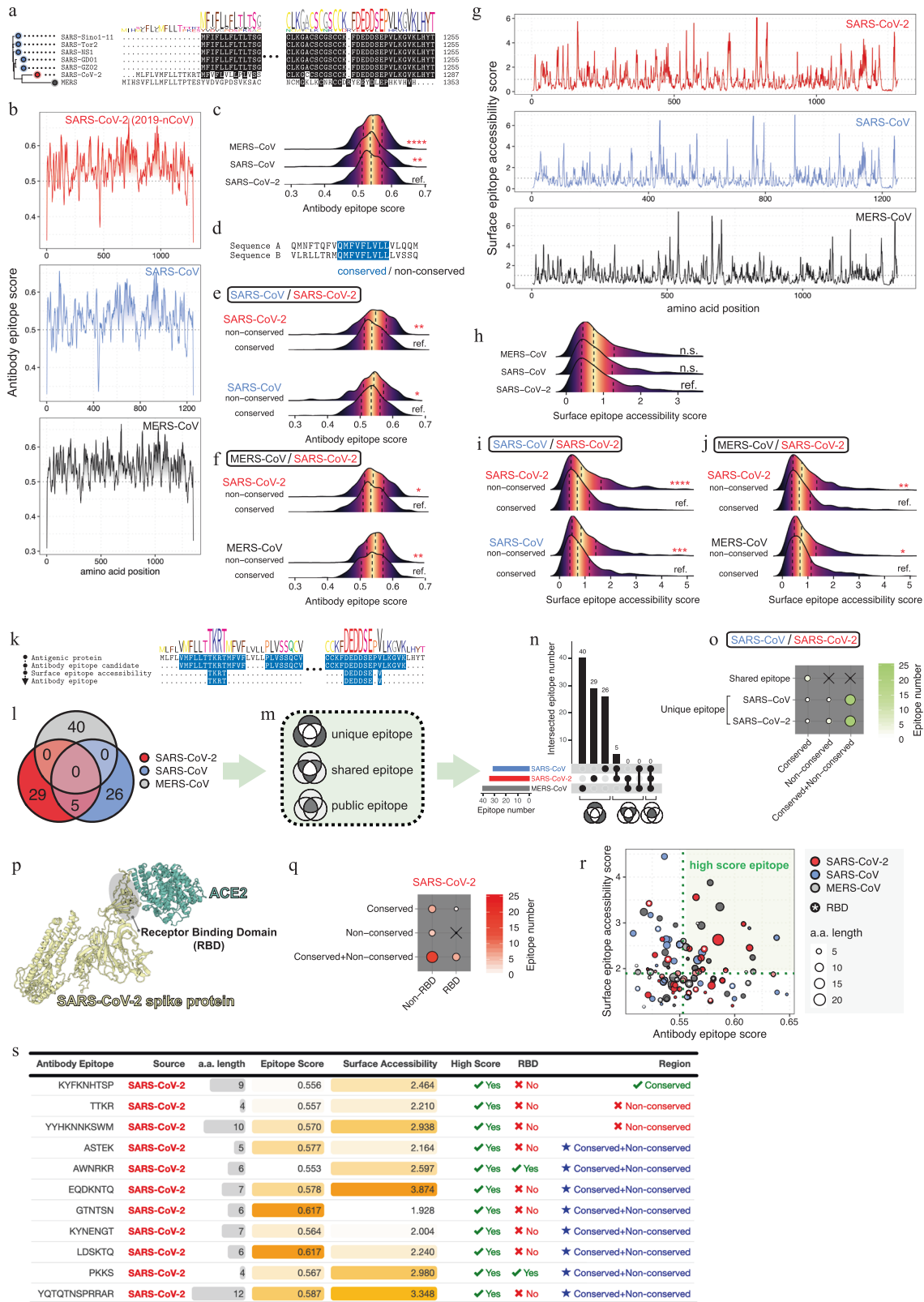
As the surface accessibility of epitope is also important for the interaction of antibody and antigen, we evaluated the surface epitope accessibility of spike proteins (Fig. 1g), no significant difference was observed in the total protein level (Fig. 1h). However, non-conserved regions showed significantly higher surface epitope accessibility score (Fig. 1i, j), indicating the non-conserved regions of spike proteins are more available for antibody recognition.

Furthermore, we identified the antibody epitopes considering both the antibody epitope and surface epitope accessibility scores (Fig. 1k and Supplementary Materials and Methods). The antibody epitopes of spike proteins were compared between MERS-CoV, SARS-CoV, and SARS-CoV-2 (Fig. 1l), and the unique, shared, and public epitopes were identified (Fig. 1m). No public epitope could be found. Although five epitopes were shared between SARS-CoV and SARS-CoV-2, there were apparent dominances of unique epitopes in SRAS-CoV (83.9%) and SRAS-CoV-2 (85.3%) (Fig. 1n).

¹Institute of Military Cognition and Brain Sciences, Academy of Military Medical Sciences, 27 Taiping Road, Beijing 100850, China
Correspondence: Ming Zheng (mmzheng@fmmu.edu.cn) or Lun Song (lunsong0752@163.com)

Received: 19 February 2020 Accepted: 19 February 2020

Published online: 4 March 2020



Moreover, among these unique epitopes, 92.7% of them were derived from the non-conserved regions and the combinations of the conserved and non-conserved regions (Fig. 1o), indicating the divergence of spike proteins could lead to major changes in the antibody epitopes.

Next, according to the cryo-electron microscopy structure of the SARS-CoV spike protein complexed with human ACE2 protein (PDB accession: 6ACJ),⁹ we used Swiss-model bioinformatic tool¹⁰ to model the three-dimensional complex structure of the SARS-CoV-2 spike protein binding to its host cell receptor

Fig. 1 Antibody epitope analysis of spike proteins in MERS-CoV, SARS-CoV, and SARS-CoV-2. **a** Alignment and phylogenetic analysis of the amino acid sequences of spike proteins in SARS-CoV-2, MERS-CoV, and five representative SARS-CoVs (total sequence alignment in Supplementary Fig. 2). **b** Antibody epitope scores in spike proteins of MERS-CoV, SARS-CoV, and SARS-CoV-2; the grey dashed lines indicate the default threshold of antibody epitope scores. **c** Density plot of the distributions of antibody epitope scores in spike proteins of MERS-CoV, SARS-CoV, and SARS-CoV-2, with colors showing tail distribution probability, and the grey dashed lines representing the first, second, and third quartiles, respectively. The results were considered statistically significant when $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$ using Kruskal–Wallis test. **d** Illustration of conserved and non-conserved sequences. **e** Density plot of the distributions of antibody epitope scores in conserved and non-conserved sequences in the comparison between spike proteins from SARS-CoV and SARS-CoV-2. The results were considered statistically significant when $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$ using Wilcoxon's signed-rank test. **f** The distributions of antibody epitope scores in conserved and non-conserved sequences in the comparison between spike proteins from MERS-CoV and SARS-CoV-2 using Wilcoxon's signed-rank test. **g** Surface epitope accessibility scores in spike proteins of MERS-CoV, SARS-CoV, and SARS-CoV-2, the grey dashed lines indicate the default threshold of surface epitope accessibility scores. **h** The distributions of surface epitope accessibility scores in spike proteins of MERS-CoV, SARS-CoV, and SARS-CoV-2 using Kruskal–Wallis test. **i** The distributions of surface epitope accessibility scores in conserved and non-conserved sequences in the comparison between spike proteins from SARS-CoV and SARS-CoV-2 using Wilcoxon's signed-rank test. **j** The distributions of surface epitope accessibility scores in conserved and non-conserved sequences in the comparison between spike proteins from MERS-CoV and SARS-CoV-2 using Wilcoxon's signed-rank test. **k** Illustration of antibody epitope identification. **l** Venn plot of antibody epitopes in spike proteins of MERS-CoV, SARS-CoV, and SARS-CoV-2. **m** Illustration of analyzing intersected epitopes of unique, shared and public groups. **n** Plot of the number of intersected epitopes in spike proteins of MERS-CoV, SARS-CoV, and SARS-CoV-2. **o** Number of shared and unique epitopes from conserved, non-conserved, and the combination of conserved and non-conserved regions in the spike proteins of SARS-CoV and SARS-CoV-2. The color and size of cycles represent the number of epitopes, and the "x" represents no epitope could be found. **p** Structural model of the SARS-CoV-2 spike protein (yellow) in complex with its human cell receptor ACE2 (cyan). The model was superimposed using the structure of the SARS-CoV spike protein and ACE2 protein complex (PDB accession: 6ACJ) as template. The gray shadow represents the receptor-binding domain (RBD). **q** Number of epitopes located in RBD and non-RBD from conserved, non-conserved, and the combination of conserved and non-conserved regions in the SARS-CoV-2 spike protein. The color and size of cycles represent the number of epitopes, and the "x" represents no epitope could be found. **r** Plot of antibody epitope scores and surface epitope accessibility scores in the epitopes, each dot represent an epitope, the green dashed lines indicate the medians, respectively, the dot color represents the source of epitopes, the dot size represents the amino acid length of epitopes, the "*" represents the epitopes located in RBD, and the light green area shows the high-score epitopes. **s** Detail information of the high-score antibody epitopes from the SARS-CoV-2 spike protein

ACE2 (Fig. 1p). We discovered that the SARS-CoV-2 spike RBD was in the interaction interface with ACE2 (Fig. 1p). In the RBD of SARS-CoV-2 spike protein, we found seven epitopes and only one of them was from the conserved region homologous to SARS-CoV, yet the rest are novel epitopes using the combinations of conserved and non-conserved regions (Fig. 1q). Furthermore, we identified the high-score epitopes with both high epitope and surface accessibility score (Fig. 1r and Supplementary Fig. 3). Finally, we found 11 high-score epitopes for SARS-CoV-2 and only 1 of them was from the conserved region, but located outside RBD; nevertheless, we identified two novel high-score epitopes located in RBD (Fig. 1s), which might be used to block the spike-ACE2 interaction to inhibit the SARS-CoV-2 infection.

In summary, our study showed that, although SARS-CoV-2 spike protein displayed high (75.5%) homology toward that of SARS-CoV, the novel epitopes contributed to 85.3% of all the antibody epitopes, 85.7% of the RBD antibody epitopes, and 90.9% of the high-score antibody epitopes in SARS-CoV-2, implying remarkable alterations in the antigenicity. Notably, these results might explain why the most of the antibodies against SARS-CoV spike protein were invalidated for SARS-CoV-2 in the previous study⁵ and indicate the necessity to develop new antibodies and vaccines specific for SARS-CoV-2. Importantly, we discovered novel and high-score antibody epitopes for SARS-CoV-2 spike protein and analyzed their RBD locations, which should be potent and specific targets for developing antibody drugs and vaccines of SARS-CoV-2 in the future. Taken together, our study found that the antigenicity of SARS-CoV-2 spike protein is remarkably dominated and altered by novel antibody epitopes, which provides promising leads for the research and development of vaccine for SARS-CoV-2.

ACKNOWLEDGEMENTS

This project is supported by the National Natural Science Foundation of China (numbers 31570758, 91743115, and 31270797) to L.S.

AUTHOR CONTRIBUTIONS

M.Z. conceived the project, developed the method, conducted data analysis, and wrote the manuscript. L.S. supervised the study and wrote the manuscript with M.Z.

ADDITIONAL INFORMATION

The online version of this article (<https://doi.org/10.1038/s41423-020-0385-z>) contains supplementary material.

Competing interests: The authors declare no competing interests.

REFERENCES

1. World Health Organization (WHO), *Statement on the second meeting of the International Health Regulations (2005). Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV)*. [https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)).
2. Gorbalenya, A. E. Severe acute respiratory syndrome-related coronavirus – The species and its viruses, a statement of the Coronavirus Study Group. *bioRxiv* **382**, 727–733 (2020).
3. Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
4. Lu, H. Drug treatment options for the 2019-new coronavirus (2019-nCoV). *Biosci. Trends* **9**, 382–385 (2020).
5. Tian, X. et al. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. *Emerg. Microbes Infect.* **9**, 382–385 (2020).
6. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* (2020).
7. Qiu, T. et al. Identification of potential cross-protective epitope between 2019-nCoV and SARS virus. *J. Genet. Genomics* (2020).
8. Jespersen, M. C., Peters, B., Nielsen, M. & Marcattili, P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* **45**, W24–W29 (2017).
9. Song, W., Gui, M., Wang, X. & Xiang, Y. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathog.* **14**, e1007236 (2018).
10. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).