RESEARCH ARTICLE

The mutation features and geographical distributions of the surface glycoprotein (S gene) in SARS-CoV-2 strains: A comparative analysis of the early and current strains

Rang Liu[1,2] | Xinran Lin[1] | Bing Chen[1] | Zhenhui Hou[1] | Qiuju Zhang[1] |
Shouren Lin[1] | Lan Geng[1] | Zhongyi Sun[1] | Canhui Cao[1,3] ⓘ | Yu Shi[1] | Xi Xia[1]

[1]Center for Reproductive Medicine, Peking University Shenzhen Hospital, Shenzhen Peking University-The Hong Kong University of Science and Technology Medical Center, Shenzhen, Guangdong, China

[2]Shantou University Medical College, Shantou, Guangdong, China

[3]Department of Neurosurgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

Correspondence

Canhui Cao, Yu Shi, and Xi Xia, Center for Reproductive Medicine, Peking University Shenzhen Hospital, No. 1120 Lotus Rd, Futian District, Shenzhen, Guangdong 518036, China.
Email: canhuicao@foxmail.com, shiyu@pkuszh.com and xixia1126@hotmail.com

Funding information

National Natural Science Foundation; Research Team of Female Reproductive Health and Fertility Preservation; Shenzhen Science and Technology Innovation Committee; China Postdoctoral Science Foundation

Abstract

The surface glycoprotein (S protein) of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was used to develop coronavirus disease 2019 (COVID-19) vaccines. However, SARS-CoV-2, especially the S protein, has undergone rapid evolution and mutation, which has remained to be determined. Here, we analyzed and compared the early (12 237) and the current (more than 10 million) SARS-CoV-2 strains to identify the mutation features and geographical distribution of the S gene and S protein. Results showed that in the early strains, most of the loci were with relative low mutation frequency except S: 23403 (4486 strains), while in the current strains, there was a surge in the mutation strains and frequency, with S: 23403 constantly being the highest one, but tremendously increased to approximately 1050 times. Furthermore, D614 (S: 23403) was one of the most highly frequent mutations in the S protein of Omicron as of March 2022, and most of the mutant strains were still from the United States, and the United Kingdom. Further analysis demonstrated that in the receptor-binding domain, most of the loci with low mutation frequency in the early strains, while S: 22995 was nowadays the most prevalent loci with 3 122 491 strains in the current strains. Overall, we compare the mutation features of the S region in SARS-CoV-2 strains between the early and the current stains, providing insight into further studies in concert with emerging SARS-CoV-2 variants for COVID-19 vaccines.

KEYWORDS

ACE2-binding region, COVID-19, genome mutations, SARS-CoV-2, surface glycoprotein

1 | INTRODUCTION

The outbreak of coronavirus disease 2019 (COVID-19), caused by a new coronavirus named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2),[1] a global pandemic, has severely impacted public health systems around the world. Until April 8, 2022, the total confirmed cases have reached more than 494 million, including 6 170 283 deaths globally.[2] The scale of the humanitarian and the economic impact of the pandemic was driving the evaluation of the COVID-19 vaccine through novel platforms to accelerate the process of development. Currently, 24 COVID-19 vaccines are authorized and used.[3] As of April 4, 2022, more than 11 billion doses have been administered.[2]

Evidence has shown that SARS-CoV-2 shared a similar sequence and used the same cell entry receptor as SARS-CoV.[1,4] In coronavirus infection, the surface spike glycoprotein (S protein) on the virion

surface mediated receptor recognition with angiotensin-converting enzyme II (ACE2) and membrane fusion conformation with host cells.[5–7] Notably, both S and N proteins were proposed to be a potential vaccine candidate for immunogenicity and T-cell immune response in Middle East respiratory syndrome coronavirus (MERS-CoV).[8,9] However, only S protein was demonstrated to induce neutralizing antibodies and T-cell immune responses.[9]

Since the first COVID-19 vaccine (messenger RNA [mRNA]-1273) candidate, based on a novel lipid nanoparticle-encapsulated mRNA vaccine that encodes S protein, entered Phase I clinical testing (NCT04283461) on March 16, 2020, there are 114 COVID-19 vaccine candidates reaching clinical trials, with 75 candidates evaluated in the preclinical stage and 48 in the final phase of testing.[10] Most vaccine candidates aimed to induce neutralizing antibodies against the viral S protein,[11] which prevented the uptake by the ACE2 receptor. Compared to the whole virus vaccines, S protein-based vaccines are more well-tolerated and relatively safer.[10] In the very early stage of vaccine development, the adenovirus type-5 (Ad5)-vectored COVID-19 vaccine (NCT04313127) expressing full-length spike protein (S protein) showed tolerance and immunogenic T-cell responses in the Phase I clinical testing.[12] However, given that SARS-CoV-2, especially the S protein, has undergone rapid evolution and mutation, there are several dominant variants worldwide in less than 2.5 years.[13] The future COVID-19 might possibly be treated as "flu" and the efficient vaccine for the whole population is considered one of the most crucial countermeasures, which is still a great concern of our researchers, however.[3]

A spike mutation D614G was revealed in May 2020 and increased in frequency globally, and this mutation was correlated with the increased viral infection.[14] Here we define strains as of April 2020 as the early strains. In this study, we comprehensively compare the early and the current prevalent SARS-CoV-2 strains in the 2019 Novel Coronavirus Resource (2019nCoVR) database to identify the spatiotemporal features of the genome mutations of the S gene in these strains over time and over countries. Moreover, we integrated the ACE2-binding region of S protein to characterize the mutations of the functional region in SARS-CoV-2 strains.

2 | MATERIALS AND METHODS

2.1 | Genome sequences of early SARS-CoV-2 strains in the database

The results of the genome mutations of 12 237 early SARS-CoV-2 strains with complete whole-genome were from the 2019 Novel Coronavirus Resource (2019nCoVR) database (https://bigd.big.ac.cn/ncov)[15] between December 2019 and April 2020. The resources of SARS-CoV-2 genome sequences were integrated from the database, which was supported by the National Genomics Data Center of China National Center for Bioinformation/Beijing Institute of Genomics. The genome sequences were from China National GeneBank DataBase (CNGBdb) (https://db.cngb.org/), GenBank, Genome Warehouse,

GISAID (https://www.gisaid.org/), National Microbiology Data Center (NMDC) (http://www.nmdc.cn/) (Supporting Information: Table 1).

2.2 | Genetic mutations of S proteins in SARS-CoV-2 strains in the database

Genetic mutations of S protein in the early and the current SARS-CoV-2 strains were integrated from the 2019nCoVR data set[15] (Supporting Information: Tables 2 and 3). The reference of the SARS-CoV-2 genome was NC_045512 (NCBI: txid2697049). The mutation types were single-nucleotide polymorphism (SNP), deletion, insertion, and indel. One site could contain more than one type of mutation.

2.3 | Heatmaps of S gene mutation in SARS-CoV-2 strains

Heatmaps of S gene mutations in SARS-CoV-2 strains were performed by the 2019nCoVR data set with default settings (https://bigd.big.ac.cn/ncov).[15] The heatmaps of S gene mutations in SARS-CoV-2 strains showed the sites with a mutation frequency of 0.0001, 0.1, and 0.5. The bar color was from 0 to 1. S gene mutations in SARS-CoV-2 strains from Cambodia, Iran, and Poland, were not shown in the heatmaps but were demonstrated in the histogram and line chart (Supporting Information: Figure 1).

2.4 | Amino acid mutations of SARS-CoV-2 strains

Amino acid mutation annotation of S protein of SARS-CoV-2 strains was based on the NCBI reference sequence NC_045512 (GeneID:43740568, QHD43416.1), including coding sequence variation, frameshift variation, inframe variation; missense variation, stop gained variation, and synonymous. All results were performed with default settings.

2.5 | Mutant positions of S protein in the prevalent variants

The composition of the prevalent lineage in every timeframe was integrated from the 2019nCoVR data set with default settings (https://bigd.big.ac.cn/ncov).[15] The heatmaps of different mutant positions of S protein in every prevalent variant were also integrated. Mutations with frequency >0.5 were shown.

2.6 | Variation dynamic curve

The variation dynamic curve was based on the genetic mutation of SARS-CoV-2 strains over time and countries. The information

regarding SARS-CoV-2 strains was integrated with the 2019nCoVR database,[15] including the sample collection date, submitting lab, sample host, and location of the strains. Each curve indicated the genome variation of the mutation site of the S region over time and countries.

2.7 | Three-dimensional (3D) structure of the ACE2 binding region in the S protein

A 3D structure of the receptor-binding domain (RBD), also the ACE2-binding region, was performed in the SWISS-MODEL server,[16–18] which was demonstrated by the rope model in the NGL viewer.[19,20] S protein is composed of three chains (A, B, and C), using PDB_ID: 6VSB structure.[21] The residues from 336 to 516 formed the binding region of the S protein with its human receptor ACE2. Some of the sites with mutant strains were shown in the binding region of S protein.

2.8 | Time and area frequencies

The time frequency of each mutation site of the S gene in the SARS-CoV-2 strains was calculated by the frequency of strains containing the mutation over each time point. The area frequency of each mutation site of the S gene in the SARS-CoV-2 strains was calculated by the frequency of strains containing the mutation over each country. The frequency and the number of strains were indicated in the graph.

2.9 | Statistics of time variance and area variance

Time variance was calculated by the population frequency of each mutant site over time and was evaluated by the dispersion of the site via calculating the variance of population frequency at each time. Area variance was calculated by the population frequency of each mutation site, with country, province, and city as region units, and was evaluated by the variance dispersion of the site via calculating the variance of population frequency in each region (Figure 4A and Supporting Information: Figure 2).

3 | RESULTS

3.1 | Genome mutation heatmaps of S gene in early SARS-CoV-2 strains by countries

We firstly analyzed the genome mutations of the S gene from 12 237 early SARS-CoV-2 strains across 60 countries and we identified 499 loci (frequency > 0.0001, Figure 1A), 32 loci (frequency > 0.1, Figure 1B), and 6 loci (frequency > 0.5, Supporting Information: Figure 1). Notably, most of the mutation loci showed low mutant

frequency by country, and the mutation loci were different among countries except the locus of 23 403. (Figure 1A). Among the strains from different countries, 60% of countries contained only one mutant hot locus (frequency > 0.1) and 18% contained more than three mutant hot loci (frequency > 0.1; Figure 1B).

Although countries like Australia, China, the United Kingdom, and the United States with more than 1000 strains showed more mutant loci, their mutation frequencies were low (Figure 1A). In addition, mutation frequencies in Slovenia and South Africa were high, with two loci (frequency > 0.5), but strains in the two countries were 4 and 7, respectively (Supporting Information: Figure 1).

3.2 | Mutation characteristics of S gene in early SARS-CoV-2 strains

Next, we integrated all the mutant strains in the S gene (NC_ 045512.2) to find out the mutation characteristics of the early SARS-CoV-2 strains. Among the 579 mutant loci, most of them (90%) were with less than five strains and only 2.94% were with more than 20 strains, especially in S: 23403 (4486 strains) and S: 24034 (108 strains) (Figure 2A,B). In addition, 98% of the mutant loci were SNP and 2% were deletion (Figure 2C).

Also, the prevalent type of amino acid mutation was missense (53%) and coding sequence variant (13%) (Figure 2D). Moreover, 410 mutant loci were showing nonsynonymous variations and 189 mutant loci were showing nonsynonymous variations (Figure 2E,F). Most (91%) of nonsynonymous variations were with less than five strains, and only 2% of them were with more than 20 strains, especially in S: 23403 (4486 strains) and S: 21575 (42 strains) (Figure 2G,H).

3.3 | Current lineage and the present mutation characteristics of the S gene

We further integrated the current epidemic lineage and the mutation feature of S protein in SARS-Cov-2 strains as of March 23, 2022. At the end of 2021, the prevalent variant transformed from Delta variant to Omicron, including BA.1 and BA.2 variants (Figure 3A). Among all the amino acid changes (frequency > 0.5), D614G was the only mutant position that was with a frequency > 0.9 in every lineage. In addition, Omicron had the highest number of mutant amino acid positions than other variants. Many amino acid changes showed in the Omicron variant for the first time like S371P, S373P, N764K, N856K, and N969K (Figure 3B).

Compared to the early SARS-Cov-2 strains, we found S: 23403 were still the loci with the greatest number of mutant strains, but tremendously increased to 4 713 032 strains, which is approximately 1050 times the number as of April 2020. The other top loci with a high number of strains include S: 23604 (3 970 507 strains), S: 22995 (3 122 491 strains) and S: 21618 (2 432 835 strains) (Figure 3C,D). From March 2020, the mutant frequency of these four loci increased
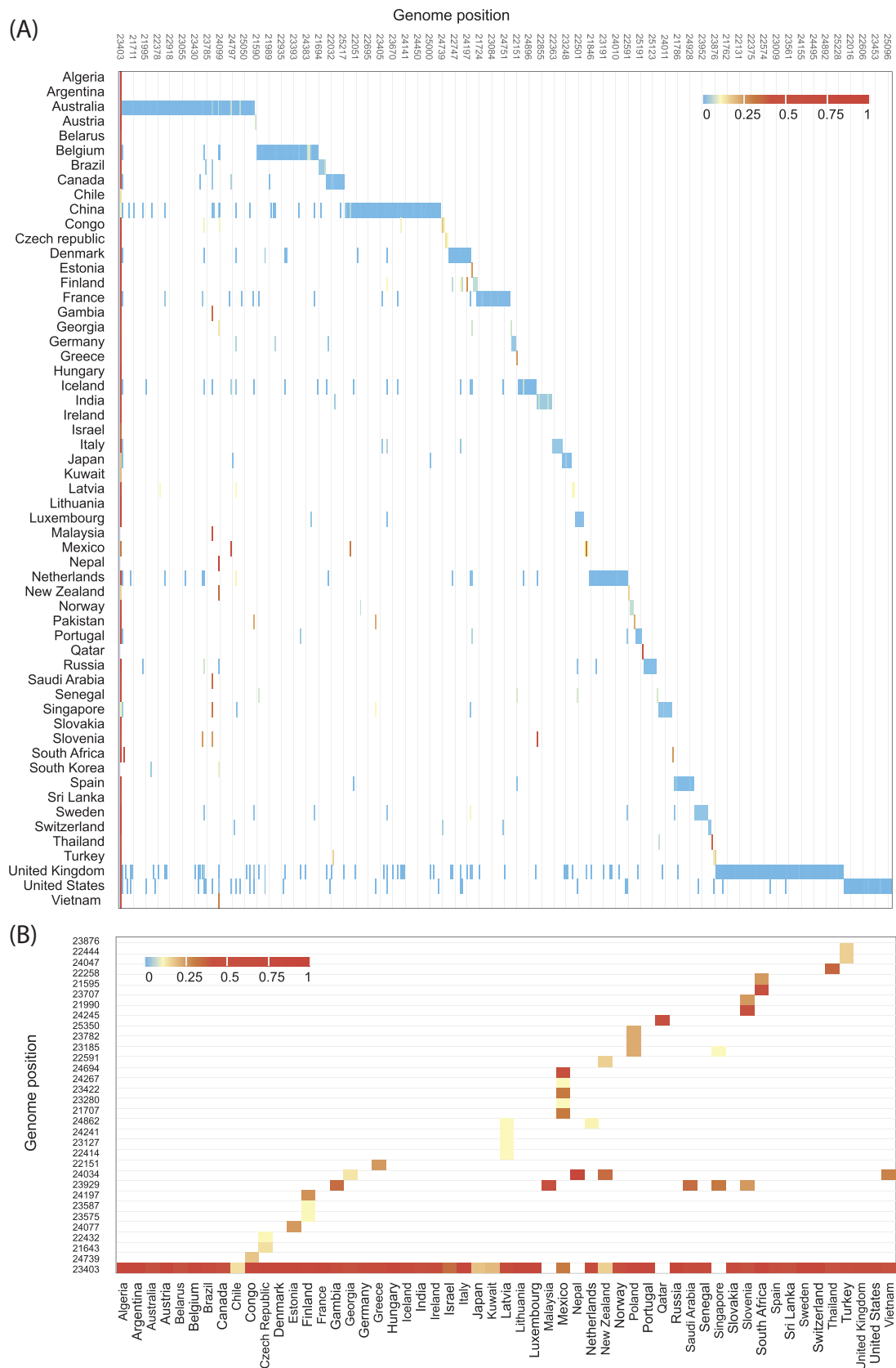
FIGURE 1 Genome mutation heatmaps of S gene in early severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strains by countries. Mutation heatmaps of S gene in different countries, with frequency > 0.0001 (A), with frequency > 0.1 (B). Each vertical line indicates a mutant locus of the S gene; the color bar indicates mutation frequency from 0 to 1; each country is labeled.
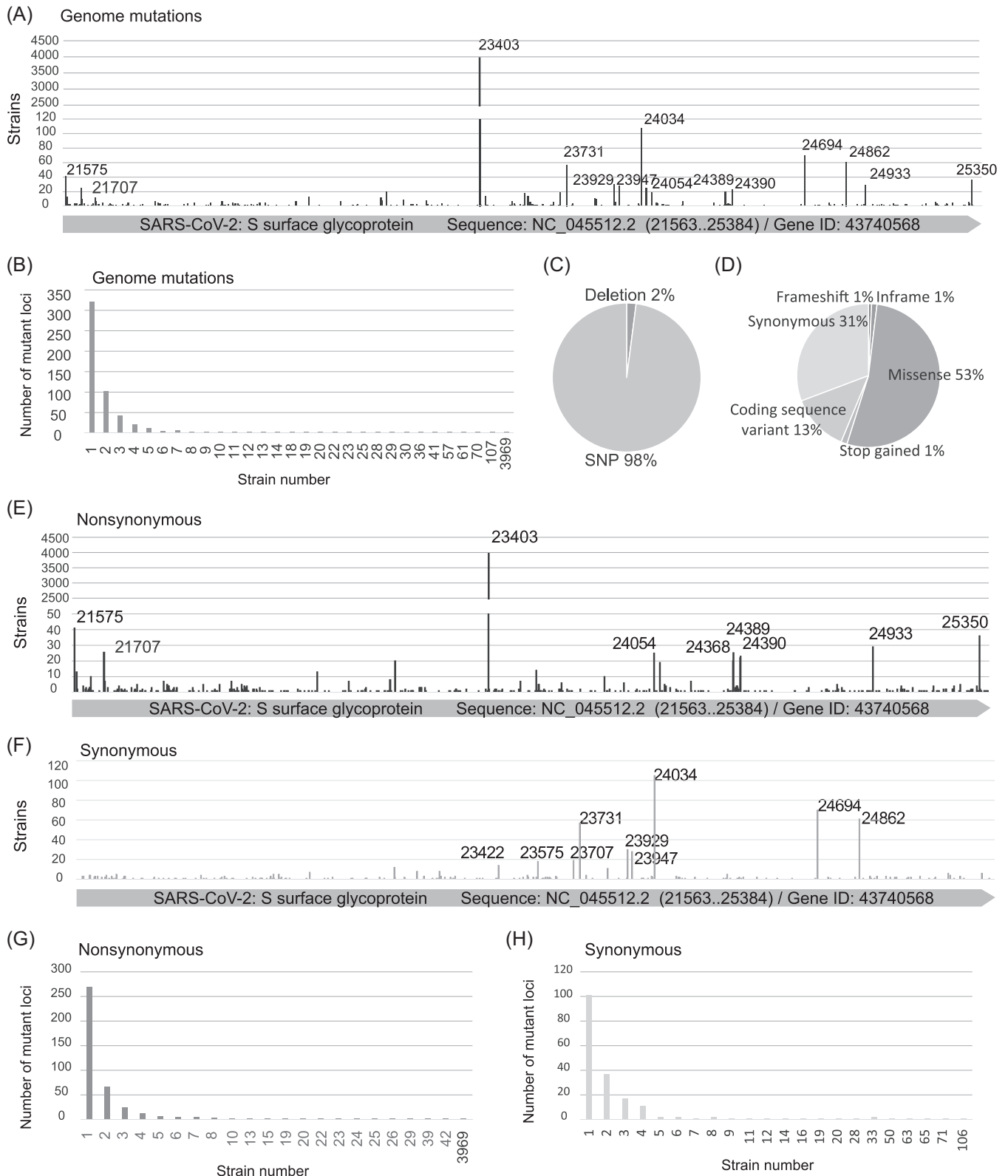
FIGURE 2 Mutation characteristics of S gene in early severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strains. (A) Mutant strains in the S gene of early SARS-CoV-2, mutant loci with a high number of strains are labeled. (B) Histogram of mutant loci, the vertical axis represents the number of mutation loci, and the horizontal axis represents a specific number of strains. Pie chart of genome mutation type (C) and amino acid mutation type (D), mutation types and percentages are shown. Nonsynonymous mutant strains (E) and synonymous mutant strains (F) in the S region of SARS-CoV-2, mutant loci with a high number of strains are labeled, and each column represents the mutant locus. Histogram of nonsynonymous mutant loci (G) and synonymous mutant loci (H), the vertical axis represents the number of mutation loci, and the horizontal axis represents a specific number of strains.
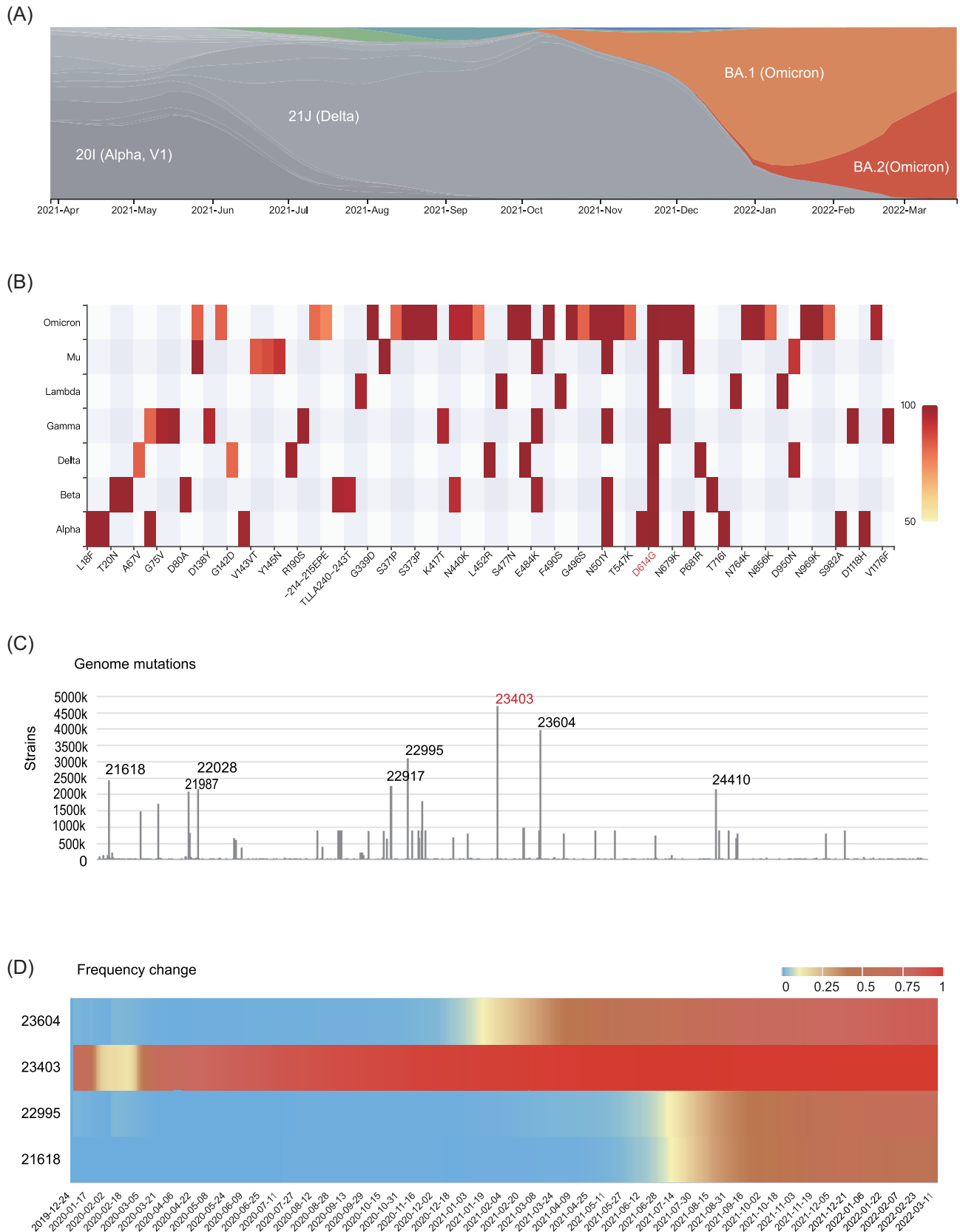
(A)

(B)

(C)

(D)

FIGURE 3 Current lineage and the present mutation characteristics of the S gene. (A) The composition of the prevalent variants s in every timeframe from April 2021 to March 2022. Different variants were in different colors and labeled. (B) Mutation heatmaps of amino acid sites in different variants, with frequency > 0.5. (C) Mutant strains in the S gene of current severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Mutant loci with a high number of strains are labeled. (D) Frequency change of the top 4 loci with the highest number of strains across time. Colors from blue, yellow, orange to red represent frequency from 0, 0.25, 0.5 to 1.

with time, among which S: 23403 was the most rapid one to reach a frequency > 0.9 (Figure 3D).

3.4 | Geographical distribution of the mutant strains

To further analyze the geographical distribution of the nonsynonymous mutant strains, we integrated the laboratory information and detected data of mutant strains from the 2019nCoVR database. These 4486 strains showed nonsynonymous mutation at locus 23403 (S: 23403_QHD43416.1: p.614-; QHD43416.1: p.614D > G), distributed in 53 countries. Although some countries showed high mutation frequency with a low number of strains, most of the mutant strains were from the United States, the United Kingdom, and Iceland, with medium mutation frequency (Figure 4A,B).

Compared to the early strains, the total number of the current strains showing nonsynonymous mutation at locus 23403 was 4 713 032. Mutant strains in countries like Germany and Japan were now significantly raising. Most of the mutant strains were still from the United States, and the United Kingdom, as well as Denmark and Germany. Mutant frequency was lowest in Niger, which was 0.4615 (Figure 4C,D).

3.5 | Mutation of ACE2-binding region of the S gene in SARS-CoV-2 strains

It was reported that the ACE2-binding region of SARS-CoV-2-mediated receptor recognition with human ACE2 in viral infection.[5,7] We then focused on the ACE2-binding region of S protein from the residues 336 to 516 to display the mutant strains' distribution at line sequence. Although there were 68 genome mutations, only one locus (S: 23010) with 20 strains and 71% of the loci with only one strain (Figure 5A).

Among the 51 nonsynonymous mutant loci, only amino acid mutation at 483 (p.483V > A) contained 20 strains and 75% of the loci contained only one strain (Figure 5B). Besides, 19 loci were synonymous mutations (Figure 5C).

Compared to the ACE2-binding region in early strains, locus 22995 (p.478) was with the highest mutant strain number, 3 122 491 strains, followed by locus 22917 and locus 23063 (Figure 5D). Ninety percent of the mutant loci in the ACE2-binding region were nonsynonymous mutations (Figure 5E).

3.6 | 3D structure of the ACE2-binding region in S protein

We integrated the 3D structure of the ACE2-binding region in the S protein. Amino acid mutation positions p., p.478, p.476, and 414 were close in S protein modeling (Figure 6).

4 | DISCUSSION

COVID-19, a novel respiratory disease caused by SARS-CoV-2, has become a global pandemic. Although Food and Drug Administration approved Veklury (remdesivir) as the first and the only antiviral drug for treating COVID-19 on October 22, 2020,[22] the early randomized, double-blinded, controlled clinical trial showed no difference in time to clinical improvement between remdesivir and placebo.[23] Results from SARS-CoV vaccines, an inactivated virus vaccine or a spike-based DNA vaccine, indicated encouraging performance with safe and neutralizing antibody titers[24,25]

As previous findings have suggested that vaccines targeting S protein could induce immune responses and protective efficiency for SARS-CoV and MERS,[26–28] an Ad5-nCo vaccine (a recombinant Ad5-vectored COVID-19 vaccine) targeting full-length S protein in Phase I clinical testing (NCT04313127) also testified the immunogenicity for SARS-CoV-2.[12] However, it was reported that vaccines expressing full-length S protein could not only induce nonneutralizing antibodies in the host[29,30] but also facilitate the virus entry into host cells via the FcγR-dependent pathway.[31] Moreover, SARS-CoV-2 has been evolving exponentially and undergoing mutations since its outburst,[13] making it necessary to determine the mutation landscape of the S gene and S protein. In general, we integrated and comparatively analyze the early and current mutant heatmap and dynamic variation curve of the S gene and the amino acid mutations of S protein, demonstrating that, although up to the present, the epidemic displayed tremendous evolution compared to the early time, with the nonsynonymous mutant locus S: 23403 having the highest number of strains among all mutations in the S protein. Meanwhile, the United States and the United Kingdom are the two top-ranking countries for the number of strains.

Since May 2020, the spike mutation of D614G was revealed globally, which was considered correlated with the increased viral infection[14] and higher RNA loads in the nasopharyngeal tract of COVID-19 patients. It was demonstrated that the spike D614G substitution increased infectivity of SARS-CoV-2 in human lung epithelial cell lines and in a primary airway tissue model. The stability of the G614 virus was reported to be increased. However, G614 SARS-CoV-2 was more susceptible to neutralization by sera from hamsters infected with D614 virus, which eased the concern about D614G mutation causing resistance to the previous COVID-19 vaccine based on the D614 sequence.[32] The current mutation heatmap indicated that D614G was also a mutant amino acid position with high frequency in Omicron. Since Omicron transmitted at a significantly higher rate than the previous variants and was reported as escapable from almost all approved antibody treatments, the countermeasures including vaccines received unprecedented concerns. However, a booster was believed to be effective to protect from Omicron infection for a three-dose vaccine could increase the proportion of broad-spectrum antibodies by B memory cells.[33] In general, the emergence of SARS-CoV-2 variants may affect COVID-19 vaccine development and antibody treatment;[34] thus, the
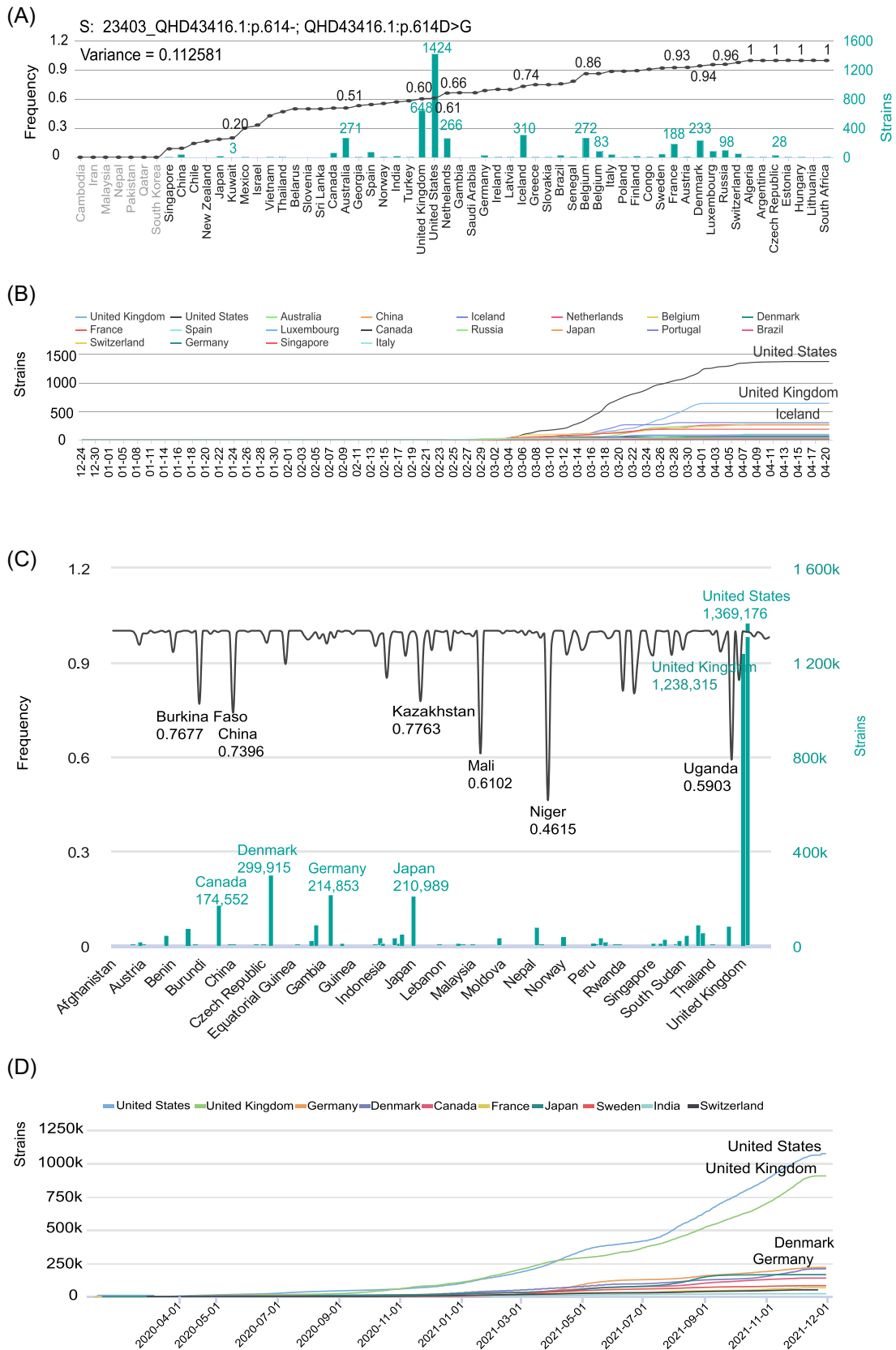
FIGURE 4 Geographical distribution of the mutant strains. Geographical distributions of the mutant strains in S: 23403 were shown as of April 2020 in (A) and as of March 2022 in (C). The histogram represents the number of strains; the line chart represents the mutation frequency; the horizontal axis represents the countries. Variation dynamic curves of countries were shown as of Apr 2020 in (B) and as of Dec 2021 (D). The vertical axis represents the number of strains. The horizontal axis represents the detected time.
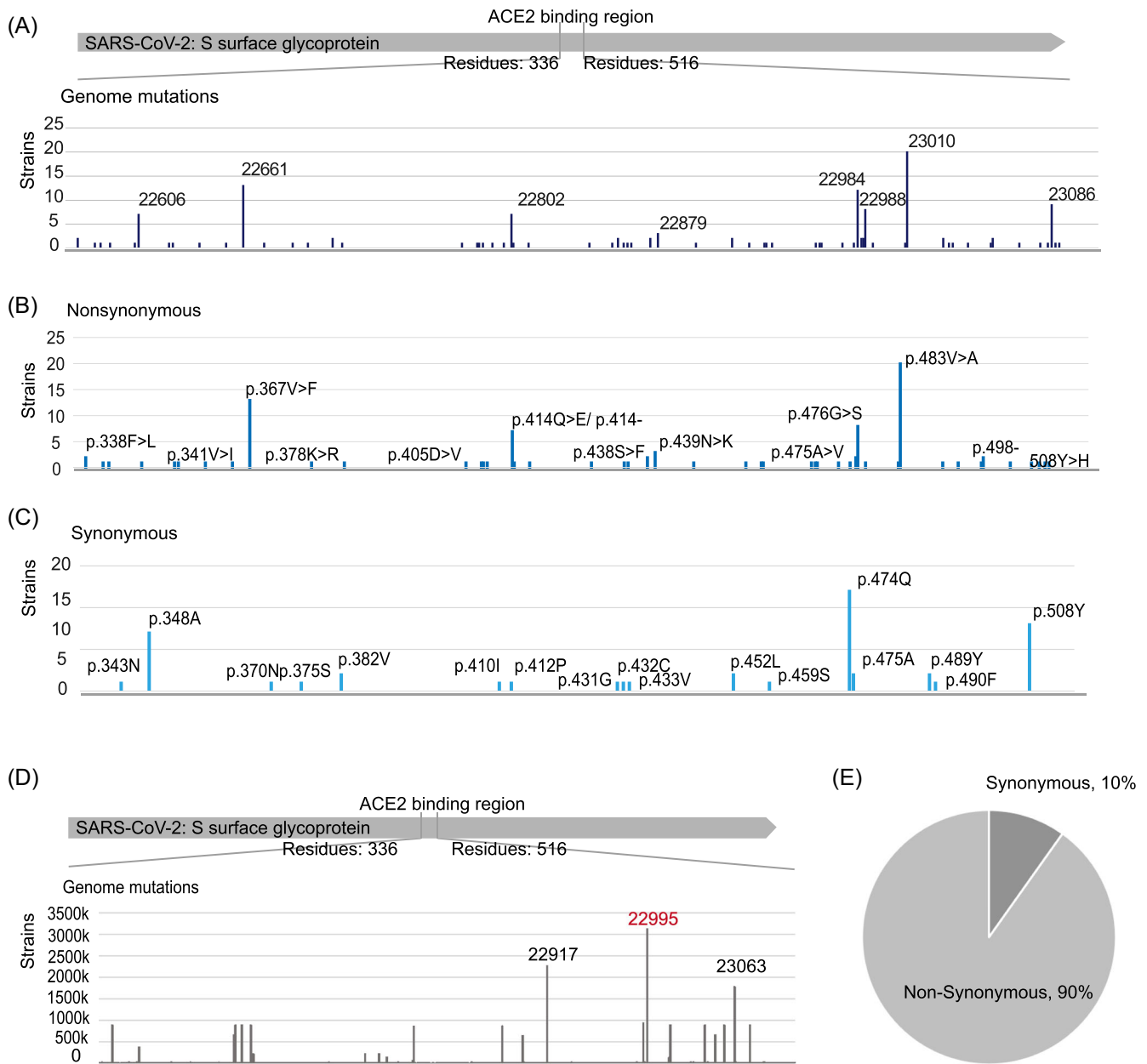
FIGURE 5 Mutation of the angiotensin-converting enzyme II (ACE2)-binding region of the S gene in angiotensin-converting enzyme II (SARS-CoV-2) strains. (A) Mutant strains in the ACE2-binding region of the S gene in early strains and mutant loci with a high number of strains are labeled. Nonsynonymous mutant strains (B) and synonymous mutant strains (C) in the ACE2-binding region of the S region, mutant loci with a high number of strains are labeled, and each column represents the mutant locus. (D) Mutant strains in the ACE2-binding region of the S gene in current strains. (E) The composition of mutation types in the ACE2-binding region is shown in the pie chart.

immunogenicity of novel COVID-19 vaccines for diverse S protein variants needs to be further analyzed.

Accumulating evidence demonstrated that subunit vaccines targeting RBD of the S protein showed a higher safety profile among current vaccines despite the low immunogenicity in the host.[35,36] Unlike the full-length S protein, RBD comprises the critical neutralizing regions but lacks the nonneutralizing domains, which restricted RBD-based vaccines from producing neutralizing immune responses.[36,37] Moreover, RBD-specific vaccines of MERS-CoV were

reported to produce neutralizing antibodies against multiple strains with a single mutation in epitopes, due to there being several conformational neutralizing epitopes.[37,38] In our work, we found that mutant strains in the RBD region were less than those outside the RBD region of the early S protein, especially at the locus S: 23403 (p.614-/p.614D > G), which was previously identified in European countries, such as the Netherlands, Switzerland, and France.[14] However, compared to the early time, the current difference of mutated strains between the RBD region and the outside region
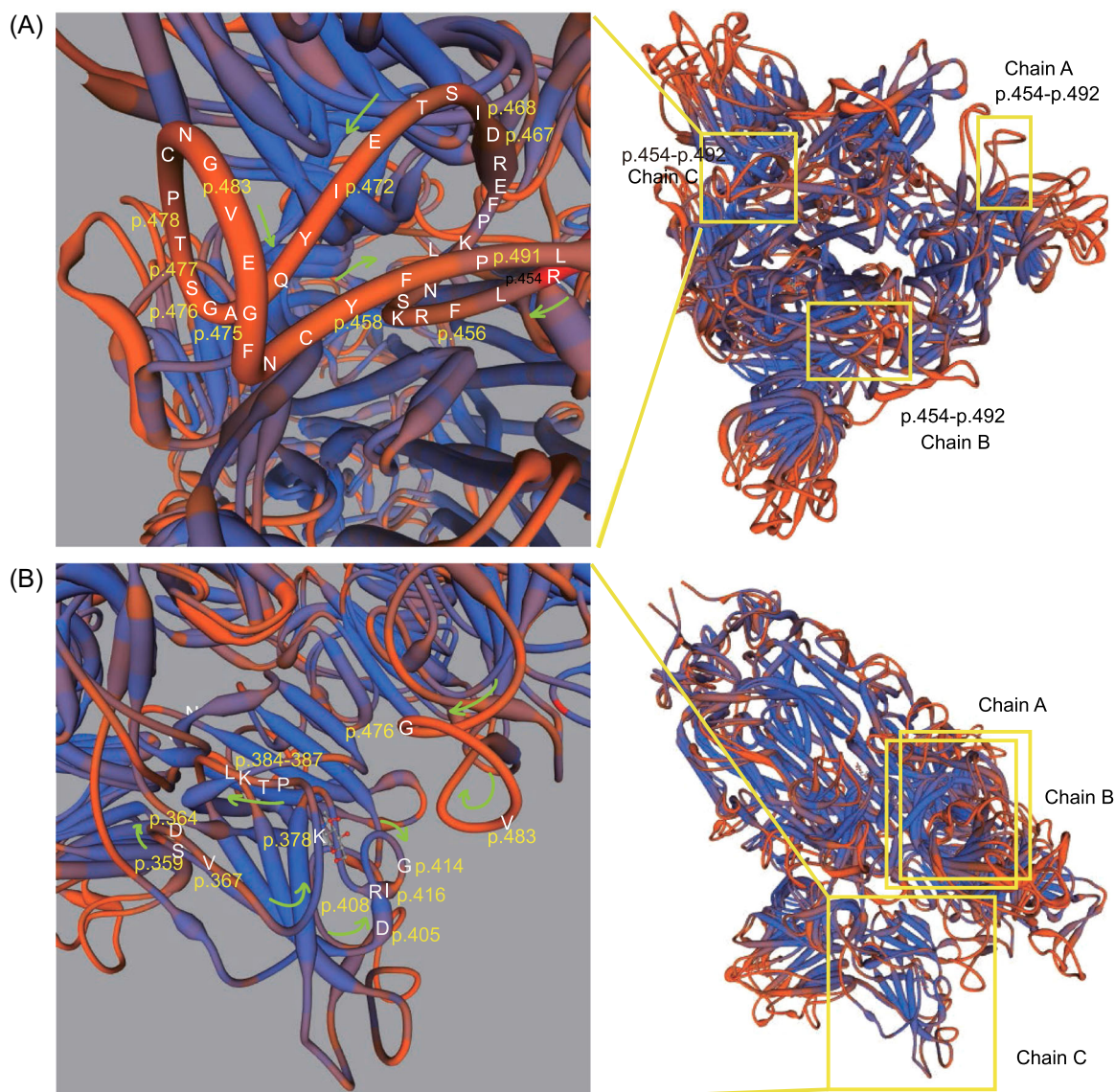
FIGURE 6 Three-dimensional structure of the angiotensin-converting enzyme II (ACE2)-binding region in S protein. 3D structure of ACE2-binding region of S region. Mutant amino acid sites were shown in the protein model.

shown to be reduced since S: 22995 (p.478) in the RBD region mutated with more than 3 million strains and 90% of the mutations were nonsynonymous.

Still, our analyses only focused on the mutant features of S protein in SARS-CoV-2 strains, lacking functional analyses of diverse variants. Besides, further experimental assays in T and B cells are necessary to identify the potential epitopes for inducing the neutralizing response against SARS-CoV-2.

5 | CONCLUSIONS

Overall, we comprehensively analyze the early and the current prevalent SARS-CoV-2 strains to identify the spatiotemporal features of the genome mutations of the S gene and S protein over time and over countries, demonstrating the mutation landscape toward the vaccine development against SARS-CoV-2. More generally, the surge of mutant strains and frequency highlights the urgency of further studies on effective vaccine development, especially S protein-based vaccines, in concert with emerging SARS-CoV-2 variants of the COVID-19 epidemic.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the 2019 Novel Coronavirus Resource (2019nCoVR) database (https://bigd.big.ac.cn/ncov).[15]

ORCID

Canhui Cao http://orcid.org/0000-0002-7328-0264

REFERENCES

1. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579:270-273. doi:10.1038/s41586-020-2012-7
2. Coronavirus disease (COVID-19) situation dashboard. World Health Organization. Published online ahead of print April 8, 2022. Retrieved from: https://covid19.who.int/
3. Altmann DM, Boyton RJ. COVID-19 vaccination: the road ahead. *Science*. 2022;375(6585):1127-1132. doi:10.1126/science.abn1755
4. Lu R, Zhao X, Li J, 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
5. Lan J, Ge J, Yu J, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. 2020;581:215-220. doi:10.1038/s41586-020-2180-5
6. Simmons G, Zmora P, Gierer S, Heurich A, Pohlmann S. Proteolytic activation of the SARS-coronavirus spike protein: cutting enzymes at the cutting edge of antiviral research. *Antiviral Res*. 2013;100:605-614. doi:10.1016/j.antiviral.2013.09.028
7. Gallagher TM, Buchmeier MJ. Coronavirus spike proteins in viral entry and pathogenesis. *Virology*. 2001;279:371-374. doi:10.1006/viro.2000.0757
8. Agnihothram S, Gopal R, Yount BL, Jr., et al. Evaluation of serologic and antigenic relationships between Middle Eastern respiratory syndrome coronavirus and other coronaviruses to develop vaccine platforms for the rapid response to emerging coronaviruses. J Infect Dis. 2014;209:995-1006. doi:10.1093/infdis/jit609
9. Shi J, Zhang J, Li S, et al. Epitope-based vaccine target screening against highly pathogenic MERS-CoV: an in silico approach applied to emerging infectious diseases. *PLoS One*. 2015;10:e0144475. doi:10.1371/journal.pone.0144475
10. Gasmi A, Srinath S, Dadar M, et al. A global survey in the developmental landscape of possible vaccination strategies for COVID-19. *Clin Immunol*. 2022;237:108958.
11. Thanh Le T, Andreadakis Z, Kumar A, et al. The COVID-19 vaccine development landscape. *Nat Rev Drug Discov*. 2020;19:305-306. doi:10.1038/d41573-020-00073-5
12. Folegatti PM, Bittaye M, Flaxman A, et al. Safety and immunogenicity of a candidate Middle East respiratory syndrome coronavirus viral-vectored vaccine: a dose-escalation, open-label, non-randomised, uncontrolled, phase 1 trial. *Lancet Infect Dis*. 2020;20:P816-P826. doi:10.1016/S1473-3099(20)30160-2
13. Zhang Y, Zhang H, Zhang W. SARS-CoV-2 variants, immune escape, and countermeasures. *Front Med*. 2022;16:1-12. doi:10.1007/s11684-021-0906-x
14. Canhui C, Huang L, Liu K, et al. Amino acid variation analysis of surface spike glycoprotein at 614 in SARS-CoV-2 strains. *Genes Dis*. 2020;7(4):567-577. doi:10.1016/j.gendis.2020.05.006
15. Zhao WM, Song SH, Chen ML, et al. The 2019 novel coronavirus resource. *Yi Chuan*. 2020;42:212-221. doi:10.16288/j.yczz.20-030
16. Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46:W296-W303. doi:10.1093/nar/gky427
17. Bienert S, Waterhouse A, de Beer TA, et al. The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res*. 2017;45:D313-D319. doi:10.1093/nar/gkw1132
18. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis*. 2009;30(suppl 1):S162-S173. doi:10.1002/elps.200900140
19. Rose AS, Bradley AR, Valasatava Y, Duarte JM, Prlic A, Rose PW. NGL viewer: web-based molecular graphics for large complexes. *Bioinformatics*. 2018;34:3755-3758. doi:10.1093/bioinformatics/bty419
20. Rose AS, Hildebrand PW. NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res*. 2015;43:W576-W579. doi:10.1093/nar/gkv402
21. Wrapp D, Wang N, Corbett KS, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 2020;367:1260-1263. doi:10.1126/science.abb2507
22. FDA approves first treatment for COVID-19. US Food & Drug Administration. Published online ahead of print October 22, 2022. Retrieved from: https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19
23. Moreno S, Alcázar B, Dueñas C, González Del Castillo J, Olalla J, Antela A. Use of antivirals in SARS-CoV-2 infection. Critical review of the role of remdesivir. *DDDT*. 2022;16:827-841. doi:10.2147/DDDT.S356951
24. Martin JE, Louder MK, Holman LA, et al. A SARS DNA vaccine induces neutralizing antibody and cellular immune responses in healthy adults in a Phase I clinical trial. *Vaccine*. 2008;26:6338-6343. doi:10.1016/j.vaccine.2008.09.026
25. Lin JT, Zhang JS, Su N, et al. Safety and immunogenicity from a phase I trial of inactivated severe acute respiratory syndrome coronavirus. vaccine. *Antivir Ther*. 2007;12:1107-1113.
26. Yong CY, Ong HK, Yeap SK, Ho KL, Tan WS. Recent advances in the vaccine development against Middle East respiratory syndrome-coronavirus. *Front Microbiol*. 2019;10:1781. doi:10.3389/fmicb.2019.01781
27. Duan J, Yan X, Guo X, et al. A human SARS-CoV neutralizing antibody against epitope on S2 protein. *Biochem Biophys Res Commun*. 2005;333:186-193. doi:10.1016/j.bbrc.2005.05.089
28. Ura T, Okuda K, Shimada M. Developments in viral vector-based vaccines. *Vaccines*. 2014;2:624-641. doi:10.3390/vaccines2030624
29. Weingartl H, Czub M, Czub S, et al. Immunization with modified vaccinia virus Ankara-based recombinant vaccine against severe acute respiratory syndrome is associated with enhanced hepatitis in ferrets. *J Virol*. 2004;78:12672-12676. doi:10.1128/JVI.78.22.12672-12676.2004

30. Czub M, Weingartl H, Czub S, He R, Cao J. Evaluation of modified vaccinia virus Ankara based recombinant SARS vaccine in ferrets. *Vaccine*. 2005;23:2273-2279. doi:10.1016/j.vaccine.2005.01.033

31. Jaume M, Yip MS, Kam YW, et al. SARS CoV subunit vaccine: antibody-mediated neutralisation and enhancement. *Hong Kong Med J*. 2012;18(Suppl 2):31-36.

32. Plante JA, Liu Y, Liu J, et al. Spike mutation D614G alters SARS-CoV-2 fitness [published correction appears in *Nature*. 2021;595(7865):E1]. *Nature*. 2021;592(7852):116-121.

33. Hu Y, Sun Q. A booster with SARS-CoV-2 vaccines: protection against Omicron infection. *Signal Transduct Target Ther*. 2022;7(1): 115.

34. Koyama T, Weeraratne D, Snowdon JL, Parida L. Emergence of drift variants that may affect COVID-19 vaccine development and antibody treatment. *Pathogens*. 2020;9(5):324. doi:10.3390/pathogens9050324

35. Du L, Tai W, Zhou Y, Jiang S. Vaccines for the prevention against the threat of MERS-CoV. *Expert Rev Vaccines*. 2016;15:1123-1134. doi:10.1586/14760584.2016.1167603

36. Du L, Jiang S. Middle East respiratory syndrome: current status and future prospects for vaccine development. *Expert Opin Biol Ther*. 2015;15:1647-1651. doi:10.1517/14712598.2015.1092518

37. Wang L, Shi W, Joyce MG, et al. Evaluation of candidate vaccine approaches for MERS-CoV. *Nat Commun*. 2015;6:7712. doi:10.1038/ncomms8712

38. Ahmed SF, Quadeer AA, McKay MR. Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies. *Viruses*. 2020;12(3):254. doi:10.3390/v12030254

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.