Ancestral origin, antigenic resemblance and epidemiological insights of novel coronavirus (SARS-CoV-2): Global burden and Bangladesh perspective

Article *in* Infection Genetics and Evolution · July 2020

DOI: 10.1016/j.meegid.2020.104440

CITATIONS
3

READS
204

6 authors, including:

Md Bashir Uddin
Sylhet Agricultural University
50 PUBLICATIONS **309** CITATIONS
SEE PROFILE

Mahmudul Hasan
Sylhet Agricultural University
52 PUBLICATIONS **185** CITATIONS
SEE PROFILE

Ahmed Rashid
Sylhet Agricultural University
31 PUBLICATIONS **62** CITATIONS
SEE PROFILE

Md. Irtija Ahsan
Sylhet Agricultural University
6 PUBLICATIONS **6** CITATIONS
SEE PROFILE

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Infection, Genetics and Evolution 84 (2020) 104440

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research Paper

Ancestral origin, antigenic resemblance and epidemiological insights of novel coronavirus (SARS-CoV-2): Global burden and Bangladesh perspective

Md Bashir Uddin[a,*,1], Mahmudul Hasan[b,1], Ahmed Harun-Al-Rashid[c], Md. Irtija Ahsan[d], Md. Abdus Shukur Imran[b], Syed Sayeem Uddin Ahmed[d,*]

[a] Department of Medicine, Sylhet Agricultural University, Sylhet-3100, Bangladesh
[b] Department of Pharmaceuticals and Industrial Biotechnology, Sylhet Agricultural University, Sylhet-3100, Bangladesh
[c] Department of Aquatic Resource Management, Sylhet Agricultural University, Sylhet-3100, Bangladesh
[d] Department of Epidemiology and Public Health, Sylhet Agricultural University, Sylhet-3100, Bangladesh

ARTICLE INFO

Keywords:
SARS-CoV-2
Evolutionary analysis
Temperature
Descriptive epidemiology
Pandemic
Bangladesh

ABSTRACT

SARS-CoV-2, a new coronavirus strain responsible for COVID-19, has emerged in Wuhan City, China, and continuing its global pandemic nature. The availability of the complete gene sequences of the virus helps to know about the origin and molecular characteristics of this virus. In the present study, we performed bioinformatic analysis of the available gene sequence data of SARS-CoV-2 for the understanding of evolution and molecular characteristics and immunogenic resemblance of the circulating viruses. Phylogenetic analysis was performed for four types of representative viral proteins (spike, membrane, envelope and nucleoprotein) of SARS-CoV-2, HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HKU1, MERS-CoV, HKU4, HKU5 and BufCoV-HKU26. The findings demonstrated that SARS-CoV-2 exhibited convergent evolutionary relation with previously reported SARS-CoV. It was also depicted that SARS-CoV-2 proteins were highly similar and identical to SARS-CoV proteins, though proteins from other coronaviruses showed a lower level of resemblance. The cross-checked conservancy analysis of SARS-CoV-2 antigenic epitopes showed significant conservancy with antigenic epitopes derived from SARS-CoV. Descriptive epidemiological analysis on several epidemiological indices was performed on available epidemiological outbreak information from several open databases on COVID-19 (SARS-CoV-2). Satellite-derived imaging data have been employed to understand the role of temperature in the environmental persistence of the virus. Findings of the descriptive analysis were used to describe the global impact of newly emerged SARS-CoV-2, and the risk of an epidemic in Bangladesh.

1. Introduction

A new strain of coronavirus, 2019 novel coronavirus or SARS-CoV-2, has emerged and infected thousands of humans. It is gaining global importance due to the unprecedented spread and death toll caused by this disease (WHO, 2020c; Zhou et al., 2020; Hui et al., 2020). The first case has been notified in China in late December 2019, and the disease is rapidly spreading to different countries and territories, including Thailand, Japan, South Korea, Iran, and the US – posing pandemic threat (Shirato et al., 2020; Rothe et al., 2020; Benvenuto et al., 2020a; Haider et al., 2020b). Determining the origin, evolutionary process, and antigenic resemblance of SARS-CoV-2 is urgently needed to study its molecular pathogenesis, perform surveillance, identify drug and vaccine targets, and develop a vaccine.

It is important to understand the genetic variability and resemblance of SARS-CoV-2 with other coronaviruses to explore the evolutionary origin of SARS-CoV-2. Human coronavirus strains evolved between 1960 and 2018 (HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HKU1 and MERS-CoV) (Giasuddin et al., 2017; Sharmin et al., 2014) are the important candidates for genetic resemblance and variability analysis. Further, some reports and studies guessed that SARS-CoV-2 originated from the bat. Thus inclusion of bat-originated coronaviruses (HKU4 and HKU5) in resemblance and variability analysis could help in elucidating the evolutionary history of the SARS-CoV-2.

Bangladesh is a densely populated country. The impact of spreading a highly contagious virus-like SARS-CoV-2 would be very high. There is no report of human coronaviruses in Bangladesh before 2020. In the

* Corresponding authors.
E-mail addresses: bashir.vetmed@sau.ac.bd (M.B. Uddin), ahmedssu.eph@sau.ac.bd (S.S.U. Ahmed).
[1] The first two authors contributed equally to this work.

https://doi.org/10.1016/j.meegid.2020.104440
Received 8 May 2020; Received in revised form 19 June 2020; Accepted 21 June 2020
Available online 01 July 2020

database and literature, a single report on Buffalo-originated coronavirus strain collected from Bangladesh (BufCoV-HKU26-M) is available (Lau et al., 2016). The global risk of the 2019 novel coronavirus (COVID-19 [SARS-CoV-2]) has recently been addressed by many scientists (Bogoch et al., 2020; Chinazzi et al., 2020; Global Health Policy, 2020; Wu et al., 2020; Zhu et al., 2020). Outside China, COVID-19 transmission has been found in over 210 countries and territories (Global Health Policy, 2020; WHO, 2020c). The US declared emergency funds because of coronavirus (The Daily Star, 2020). As the outbreak of the 2019 novel coronavirus (COVID-19 [SARS-CoV-2]) is spreading rapidly, the analysis of epidemiological data of COVID-19 is necessary to explore the disease burden and associated factors.

The present study aims to compare the genetic materials of SARS-CoV-2 with different previously reported virulent strains of coronaviruses and to discuss the risk and possible impact of the SARS-CoV-2 epidemic in Bangladesh based on the descriptive epidemiological analysis.

2. Methods

2.1. Retrieving human-, bat- and buffalo-originated coronavirus protein sequences and epidemiological data

Spike (S), membrane (M), envelope (E) and nucleocapsid (N) protein sequences of SARS-CoV-2, HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HKU1, MERS-CoV, HKU4, HKU5 and BufCoV-HKU26 were retrieved from the ViralZone root (https://viralzone.expasy.org/) and UniProt (https://www.UniProt.org/) databases (Supplementary File 1). Additionally, we retrieved epidemiological data for COVID-19 from different websites, such as the World Health Organization (WHO), European Centre for Disease Prevention and Control (ECDC), and other online news portals (ECDC, 2020; Pharmaceutical-technology, 2020; Reliefweb, 2020; WHO, 2020c; Worldometers. COVID-19 Coronavirus outbreak, 2020). Also, we extracted population data for countries and provinces (China) from several websites (City population, 2020; Countrymeters, 2020; Wikipedia, 2020).

2.2. Phylogeny study and pairwise sequence alignment of coronavirus proteins

The retrieved gene sequences were subjected to multiple sequence alignment (MSA). Sequence alignments were performed in ClustalW (Thompson et al., 2003). Phylogenetic relationship (maximum parsimony, MP) was constructed, by using MEGA X (Kumar et al., 2018), to understand the ancestral origin and antigenic resemblance of SARS-CoV-2 with other coronaviruses. Besides, pairwise sequence alignment of SARS-CoV-2 proteins with different viral strains was performed by the EMBOSS Needle online software, which uses the Needleman-Wunsch alignment algorithm to find the optimum alignment (including gaps) of two sequences along their entire length (Chojnacki et al., 2017). Moreover, sequence alignment was also visualized and analyzed by using Jalview software (https://www.jalview.org/).

2.3. Immunogenicity prediction and epitope conservancy analysis

Targeting potential antigens from viral proteins is crucial for constructing peptide-based vaccine molecules that can interact with B lymphocytes (Hasan et al., 2019a). It was reported that peptide flexibility and proper surface accessibility are prerequisites for being a potential B cell epitope. Considering those parameters, the immunogenic peptide sequences from four types of viral proteins were determined by using the B cell epitope prediction tools of the Immune Epitope Database (IEDB) (Vita et al., 2015), which employs the Bepipred linear epitope prediction method (Larsen et al., 2006). The VaxiJen v2.0 server (http://www.ddgpharmfac.net/vaxijen/) was used for screening out the most immunogenic peptides determined from IEDB (Vita et al., 2015). However, epitope conservancy analysis is a crucial step to assess the degree of desired epitope distribution in its homologous protein set. In this study, the conservancy pattern of mostly immunogenic B cell peptide sequences of COVID-19 was compared with other homologous sequences retrieved from the NCBI database by using BLASTp (Hasan et al., 2019b). Moreover, the conservancy study of immunogenic peptides predicted from the SARS-CoV-2 proteins was also compared against other human coronavirus strains (HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HKU1 and MERS-CoV). The epitope conservancy analysis tool (http://tools.iedb.org/conservancy/) of the IEDB was used to continue the conservancy analysis (Hasan et al., 2019c).

2.4. Homology modeling and evaluation of 3D protein structures

Homology modeling of spike glycoprotein (P0DTC2), membrane protein(P0DTC5), an envelope protein (P0DTC4), and nucleoprotein (P0DTC9) of SARS-CoV-2 was performed by using the I-TASSER server (Zhang, 2008). Although multiple threading alignments generated 3D structures in the I-TASSER server, refinement was conducted using ModRefiner (Xu and Zhang, 2011) followed by the FG-MD refinement server to improve the accuracy of the predicted 3D modeled structure (Zhang et al., 2011). ModRefiner allowed for significant improvements in the physical quality of the local structure based on hydrogen bonds, side-chain positioning and backbone topology of the native-state proteins. Again, FG-MD, a molecular dynamics-based algorithm for structure refinement, usually works at the atomic level. The refined protein structure was further validated by RAMPAGE (Wang et al., 2016) and ERAAT analyses (Gundampati et al., 2012). Structures were visualized and analyzed by PyMoL (L DeLano, 2002).

2.5. Illustration of COVID-19 cases and deaths

We illustrated the number of cases and deaths of SARS-CoV-2 through graphs to elucidate the pattern of occurrence. We plotted the number of global cases and Chinese cases by date, the global death toll per day against time. Further, we presented global and country-wise case fatality.

2.6. Calculation of crude mortality rate and case fatality of COVID-19

We calculated the crude mortality rate and case fatality according to the formulas suggested by the CDC (CDC, 2020) as well as Jacob and Ganguli (Jacob, 2008). Here, we calculated the crude mortality rate for countries and Chinese provinces and standardized per 1 crore persons.

2.7. Satellite-based temperature data

It is already known that the SARS-CoV-2 can multiply even at high temperatures, especially temperatures higher than 15 °C (Chan et al., 2011; Kampf et al., 2020); however, SARS-CoV-2 is rapidly inactivated at 20 °C (Kampf et al., 2020). Therefore, temperature plays a great role in its multiplication. For this purpose, recent environmental temperature data from the place of first occurrence as well as Bangladesh were obtained from Landsat-8 satellite data. This satellite provides high spatial resolution (30 m) data at 15-day intervals. Using the brightness temperature of band number 10 (TIR-1) and emissivity data temperature (in °C) of bands 4 and 5 (L8 Data Users Handbook), a large area (a 30-km-wide swath) can be obtained for a time with minor deviation from in situ temperature data (maximum 0.45 degree Celsius SD). Therefore, cloudless or less cloudy images (less than 90%) were obtained from the USGS webpage (www.earthexplorer.usg.gov). A maximum of 2 data points were available for one area in each month. However, neighboring path and row image borders shared some common areas, which provided more frequencies for those overlapped areas. Level-1 Tier-1 images, which are radiometrically and

geometrically corrected, were used in this study.

First, all images fulfilling the cloud-related conditions were downloaded. A total of 90 images covering the land areas of Wuhan, China, Korea, Italy and Bangladesh were downloaded. Then, DN of Band 10 data were converted to emissivity and simultaneously converted to brightness temperature by using "equation number 1" (Cahyono et al., 2017). Then, the emissivity was converted to temperature by using "equation number 2" (Syariz et al., 2015).

The estimated data were obtained by the Landsat 8 Thermal Infrared Sensor (TIRS) of band 10. To obtain the temperature data for coverage of Bangladesh, Landsat-8 images of path 135 rows 44–46, path 136 rows 43–45, path 137 rows 42–45, path 138 rows 42–45, and path 139 rows 41–43 were used. To obtain temperature data for coverage of Korea, Landsat-8 images of path 114 rows 34–36, path 115 rows 33–37, and path 116 rows 33–37 were used. To obtain temperature data for coverage of the city of Wuhan, Landsat-8 images of path 122 rows 39 and path 123 rows 38–39 were used. For Italy, Landsat-8 images of path 186 rows 32, path 187 rows 32–34, path 188 rows 31–35, path 189 rows 31–32 and 34, path 190 rows 30–314, path 191 rows 28–31, path 192 rows 28–30 and 32–33, path 193 rows 27–33, path 194 rows 28–30, path 195 rows 28–29, path 196 rows 28–03, and path 194 rows 28 were used. To obtain temperature data for coverage of the city of Wuhan, Landsat-8 images of path 122 rows 39 and path 123 rows 38–39 were used.

For obtaining the temperature, the digital number (DN) of band 10 for each image was converted to the brightness temperature by using “equation number 1”.

$$L_\lambda = M_L \times Q_{cal} + A_L \quad (1)$$

where L_λ = Top of atmosphere spectral radiance (Watts/($m^2 \times$ sr $\times$ μm)), M_L = Radiance multiplicative scaling factor for band 10, A_L = Radiance additive scaling factor for band 10, Q_{cal} = Level 1 pixel value in D.

Then, by using equation number 2, the real temperature value (˚K) was obtained:

$$T = \frac{K2}{\ln\left(\frac{K1}{L_\lambda} + 1\right)} \quad (2)$$

where T = Top of atmosphere brightness temperature (˚Kelvin), K2 = Thermal conversion constant for band 10, K1 = Thermal conversion constant for band 10.

The effective temperature value (˚K) was converted to the effective temperature value (˚C) using the following equation:

$$T\,(°C) = T\,(°K) - 273 \quad (3)$$

where T (˚C) = Brightness temperature value (˚C), T (˚K) = Brightness temperature value (˚K).This information was automatically obtained from metadata.

3. Results

3.1. Phylogenetic analysis and pairwise sequence alignment of coronavirus proteins

The four phylogenetic trees constructed from four types of representative viral proteins (spike, membrane, envelope and nucleoproteins) of SARS-CoV-2, HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HKU1, MERS-CoV, HKU4, HKU5 and BufCoV-HKU26. Phylogenetic trees demonstrated the ancestral origin and distant evolutionary relationships of the newly emerged novel coronavirus (COVID-19). It was found that SARS-CoV-2 was evolutionarily related to SARS-CoV (**Fig.** 1). In all trees, SARS-CoV-2 aligned with the same

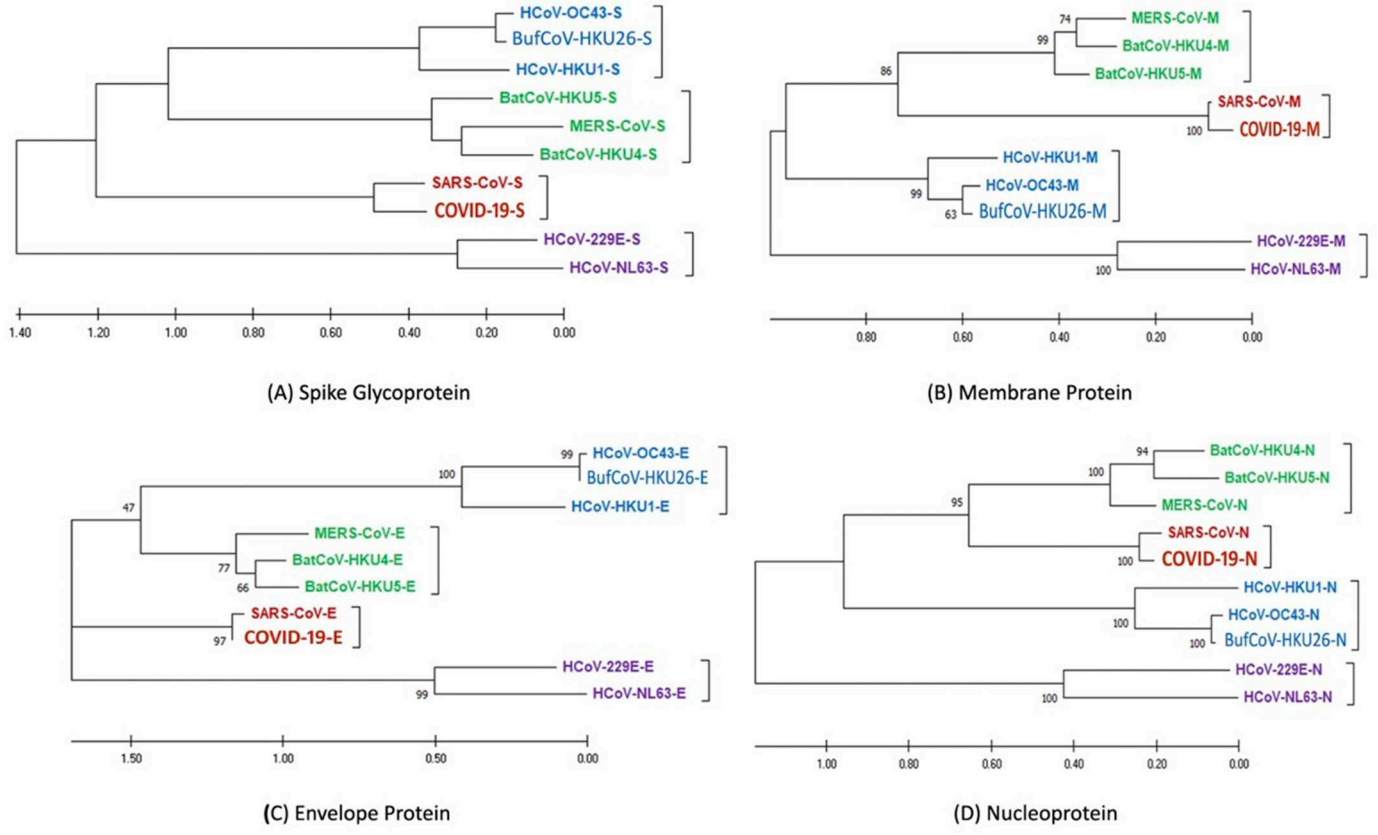

Fig. 1. Phylogeny study of SARS-CoV-2 with the other member of coronavirus family. A phylogentic tree was constructed with (A) Spike glycopreotein, (B) Membrane protein, (C) Envelope protein and (D) Nucleoprotein of COVID-19 with HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HKU1, MERS-CoV, HKU4, HKU5 and BufCoV-HKU26 coronaviruses by using Maximum Likelihood Method of MEGA X.

Fig. 2. (i) Multiple sequence alignment of SARS-CoV-2 proteins with SARS-CoV. Multiple sequence alignment of (A) COVID-19-S and SARS-CoV-S, (B) COVID-19-M and SARS-CoV-M, (C) COVID-19-E and SARS-CoV-E and (D) COVID-19-N and SARS-CoV-N was visualized. Conservation showed based on 11 base scales where yellow color bar indicates the full conservation. Alignment quality was based on BLOSUM 62 substitution matrix score where yellow color indicates good quality. All the colors changes according to the conservation and alignment quality. Black bars showed the consensus sequence. This alignment was visualized by Jalview 2.8 and color scheme used is Clustalx. (ii) Multiple sequence alignment of SARS-CoV-2 proteins with BufCoV-HKU26. Multiple sequence alignment of (A) COVID-19-S and BufCoV-HKU26-S, (B) COVID-19-M and BufCoV-HKU26-M, (C) COVID-19-E and BufCoV-HKU26-E and (D) COVID-19-N and BufCoV-HKU26-N was visualized. Conservation showed based on 11 base scales where yellow color bar indicates the full conservation. Alignment quality was based on BLOSUM 62 substitution matrix score where yellow color indicates good quality. All the colors changes according to the conservation and alignment quality. Black bars showed the consensus sequence. This alignment was visualized by Jalview 2.8 and color scheme used is Clustalx. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

clade of SARS-CoV, where two bat-originated coronaviruses, HKU4 and HKU5, were found to have divergent relationships with COVID-19. However, HKU4 and HKU5 showed ancestral relationships with MERS-CoV. Moreover, BufCoV-HKU26, of Bangladesh origin, always showed a phylogenetic relationship with the animal-originated coronavirus HCoV-OC43. Multiple sequence alignment (MSA) of COVID-19 proteins with SARS-CoV and BufCoV-HKU26 proteins is shown in **Fig. 2** (i) and (ii). The results suggested that major proteins four proteins of SARS-CoV-2 were more aligned with the SARS-CoV proteins, which might substantiate the claim of convergent evolutionary relation of SARS-CoV-2 and SARS-CoV [**Fig. 2** (i)]. On the other hand, the major proteins of BufCoV-HKU26 were also subjected to MSA with SARS-CoV, and it

Table 1
Pairwise sequence alignment of COVID-19 proteins with other viral strains by the EMBOSS Needle.

SARS-CoV-2 Proteins	*Other Coronavirus Proteins*[a]								
	SARS-CoV-S	*HCoV-OC43-S*	*HCoV-229E-S**	*HCoV-NL63-S*	*HCoV-HKU1-S*	*MERS-CoV-S*	*BatCoV-HKU4-S*	*BatCoV-HKU5-S*	*BufCoV-HKU26-S*
Spike Protein	76.4% 87.0%	29.2% 43.3%	26.4% 39.4%	23.35% 35.0%	28.0% 44.2%	29.8% 45.3%	32.1% 46.6%	32.2% 47.0%	29.5% 44.2%
	SARS-CoV-M	*HCoV-OC43-M*	*HCoV-229E-M**	*HCoV-NL63-M*	*HCoV-HKU1-M*	*MERS-CoV-M*	*BatCoV-HKU4-M*	*BatCoV-HKU5-M*	*BufCoV-HKU26-M*
Membrane Protein	90.5% 96.4%	37.2% 55.6%	30.7% 51.3%	29.5% 50.6%	34.4% 54.2%	39.9% 59.2%	41.0% 60.4%	42.7% 59.6%	37.4% 55.7%
	SARS-CoV-E	*HCoV-OC43-E*	*HCoV-229E-E**	*HCoV-NL63-E*	*HCoV-HKU1-E*	*MERS-CoV-E*	*BatCoV-HKU4-E*	*BatCoV-HKU5-E*	*BufCoV-HKU26-E*
Envelope Protein	94.7% 96.1%	22.7% 42.0%	24.7% 44.4%	17.9% 52.6%	26.8% 45.1%	35.4% 46.3%	39.8% 56.6%	29.5% 44.3%	20.9% 44.2%
	SARS-CoV-N	*HCoV-OC43-N*	*HCoV-229E-N**	*HCoV-NL63-N*	*HCoV-HKU1-N*	*MERS-CoV-N*	*BatCoV-HKU4-N*	*BatCoV-HKU5-N*	*BufCoV-HKU26-N*
Nucleoprotein	90.5% 94.3%	32.9% 48.7%	24.6% 35.8%	26.8% 39.4%	29.6% 44.1%	45.6% 59.6%	44.5% 57.4%	43.6% 58.2%	32.7% 48.6%

S = Spike protein; M* = Membrane protein; E* = Envelope protein; N* = Nucleoprotein.*
[a] *First percentage indicates the similarity and second percentage indicates the identity.*

revealed that there were significant dissimilarities between these two strains [**Fig.** 2 (ii)]. Moreover, pairwise sequence alignment by EMBOSS Needle strengthened the phylogenetic relationship of SARS-CoV-2 and SARS-CoV. **Table 1** included the similarity and identical pattern of SARS-CoV-2 proteins with their homologous proteins of other coronaviruses. It had been raveled that SARS-CoV-2 proteins were highly similar and identical with SARS-CoV proteins. Spike, membrane, envelope, and nucleocapsid proteins of SARS-CoV-2 showed that approximately 76.4%, 90.5%, 94.7%, and 90.5% nucleotide similarities with the respective protein molecules of SARS-CoV. Again, in the case of identity pattern, results showed that spike, membrane, envelope, and nucleoproteins of SARS-CoV-2 were 87.0%, 96.4%, 96.1% and 94.3% identical with the respective homologous proteins of SARS-CoV. Besides, minimum similarity and identical patterns were found with other viral strains, including BufCoV-HKU26 of Bangladesh origin (Table 1).

3.2. Immunogenicity prediction and epitope conservancy analysis

Four types of viral proteins (Supplementary File 1) of SARS-CoV-2, HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HKU1, MERS-CoV, HKU4, HKU5 and BufCoV-HKU26 were employed to determine the most antigenic sites by using the B cell epitope prediction tool of IEDB and VaxiJen scoring. The VaxiJen server, which gave a result well above the threshold value (0.40), usually reveals the immunogenic potential to stimulate a protective response in host organisms (Flower et al., 2017). From the analysis, a total of 17 epitopes from S proteins, 1 epitope from M proteins, 1 epitope from E proteins and 5 epitopes from N proteins were found to be mostly immunogenic in SARS-CoV-2, with almost 100% of peptides carrying more than the threshold value of the antigenic score of the VaxiJen server (Table 2). Similarly, other coronavirus proteins also showed different lengths of immunogenic epitope candidates with values exceeding the antigenic threshold (Supplementary Table 1). Moreover, homologous sequences of SARS-CoV-2 S, M, E and N proteins were retrieved from the NCBI database (Supplementary File 2) and were subjected to conservancy analysis with the immunogenic epitopes from SARS-CoV-2 proteins. It was found that antigenic sites are almost conserved in all of the homologous protein sequences deposited in the NCBI database (Table 2). Cross-checked conservancy analysis of COVID-19 antigenic epitopes with SARS-CoV proteins showed that conservancy when cross-checked with other coronaviruses, including BufCoV-HKU26 of Bangladesh origin, was not significant (Table 3).

3.3. Homology Modeling and evaluation of the 3D protein structure of COVID-19

The 3D structures of the SARS-CoV-2 proteins were modeled using the I-TASSER server, where 'SARS-CoV complex with human neutralizing S230 antibody Fab fragment' (PDB ID: 6NB6), 'Membrane protein complex' (PDB ID: 5X5Y), 'Coronavirus Envelope Proteins' (PDB ID: 2 MM4) and 'SARS CoV Nucleocapsid' (PDB ID: 1SSK) acted as the top threading templates of the S, M, E and N proteins, respectively. The SARS-CoV-2 proteins showed 36% to 93% coverage and 40% to above 90% identity to the I-TASSER template proteins (Supplementary Table 3). As I-TASSER usually deduces 5 different models for each operation, refinement by Modrefiner and validation by RAMPAGE and ERRAT were performed to check the accuracy of the modeled structures.

After the refinement process, Ramachandran plot analysis was investigated to fix the structural accuracy from homology modeling. Ramachandran plot could do the placement of amino acids from modeled protein structure following the all attributes of structural biology, from where comparatively better model could be screened out. The present study revealed that the spike protein of SARS-CoV-2 showed 75.2% residues in the favored region, 14.8% residues in the allowed region, and only 10% residues in the outlier region [Fig. 3 (i)A]. On the other hand, membrane protein exhibited 77.7% residues in the favored region, 15% residues in the allowed region, and only 7% residues in the outlier region [Fig. 3 (i)B]. In addition, two other envelopes and nucleoproteins also showed the highest number of residues in the favored regions (74% for E and 68.1% for N) and a lower number of residues (7.3% for E and 11.5% for N) in the outlier regions [Fig. 3 (i)C and (i)D]. Again, for assuring more structural validity of the constructed models, the results of ERRAT had been analyzed. It had been found that values of the highest quality factor from the modeled SARS-CoV-2proteins (S = 71%, M = 88%, E = 89% and E = 73%) were significant [Fig. 3(ii) and Supplementary Table 3]. In addition, the probable immunogenic epitopes of major SARS-CoV-2 proteins were studied in the previous steps, and here, all of the predicted immunogenic epitopes of SARS-CoV-2were presented into the validated 3D structures of the S, M, E and E by using PyMoL in Fig. 4.

3.4. Burden of COVID-19: Occurrence of cases

Over two and half months, from its outbreak on December 31, 2019 to March 20, 2020, COVID-19 has spread to up to 182 countries and territories [**Fig. 5(i)A**] affecting a total of 253,796 persons globally, including in Bangladesh. China was the most affected country [**Fig. 5**(i)

(i)

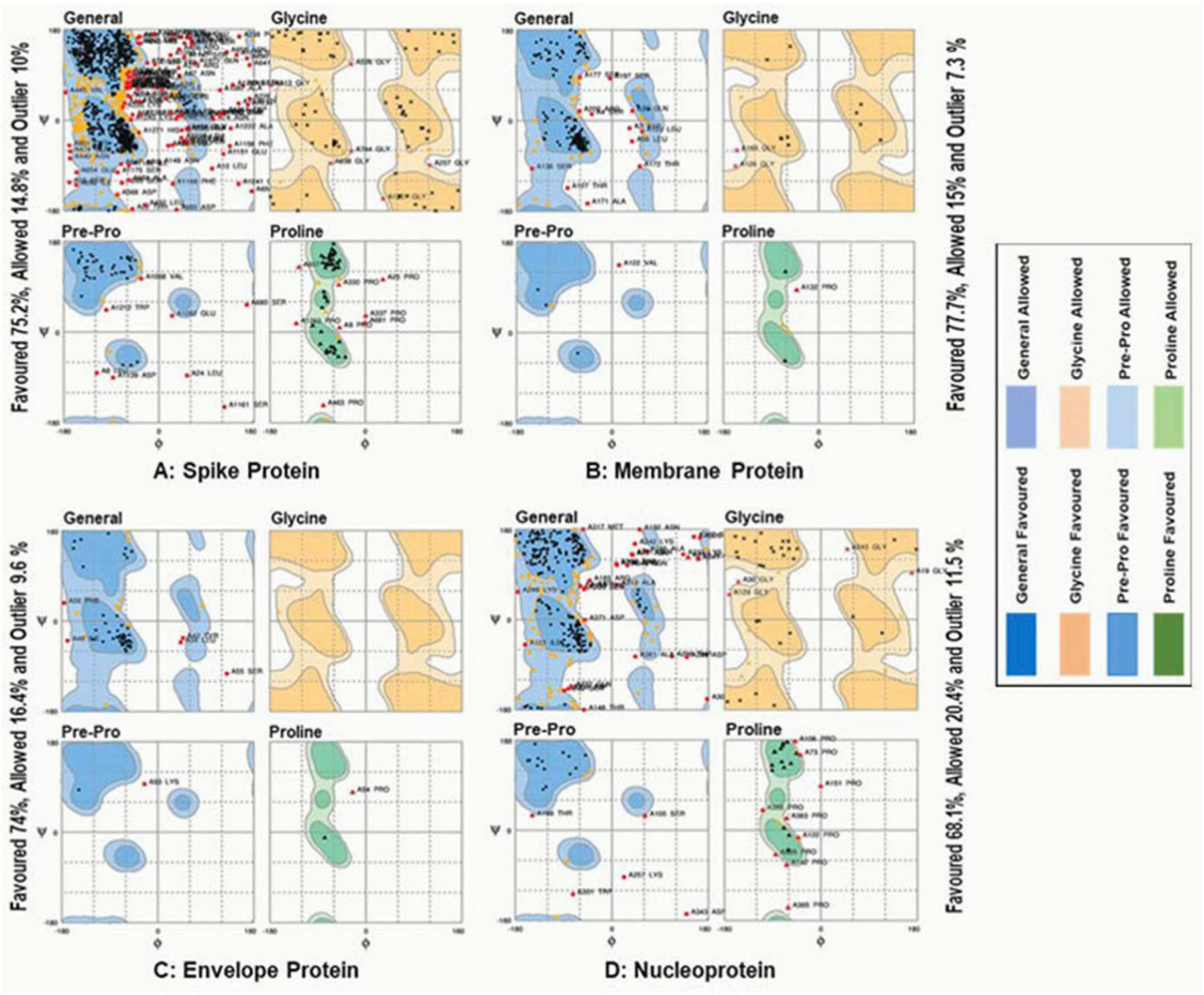

Fig. 3. (i) Ramachandran plot for SARS-CoV-2 proteins. Structure validation of (A) Spike protein, (B) Membrane protein, (C) Envelope protein and (D) Nucleoprotein of COVID-19 proteins predicted by Ramachandran plot analysis of RAMPAGE. (ii) Quality factor analysis of predicted SARS-CoV-2 proteins by ERRAT. This figure shows the overall quality factor of predicted structure of (A) Spike protein, (B) Membrane protein, (C) Envelope protein and (D) Nucleoprotein.

B], holding 32.06% (81,375/253,796) of global cases so far (**Fig. 6A** and **Supplementary File 3**). The second- and third-most affected countries were Italy and Spain, with a total of 41,035 and 19,980 cases, respectively. The cruise ship, Diamond Princess, experienced 706 cases among a total of 3,706 persons. Also, COVID-19 resulted in a substantial burden in Iran, Germany, the United States of America (USA), France, and South Korea infecting 19,644, 17,372, 14,366, 10,995 and 8,652 persons, respectively. However, the number of total cases in other countries ranged from one to five thousand (**Supplementary File 3**). Among 34 provinces in China, including SARs (special administrative regions), up to March 19, 2020, the highest number of cases occurred in Hubei (67,869) (**Fig. 6B**). Other highly affected provinces were Guangdong (1,370), Henan (1,273), Zhejiang (1,232), Hunan (1,018), Anhui (990) and Jiangxi (935) (**Supplementary File 4**). For both global and Chinese cases, there was a steady increasing trend of the occurrence of the disease per day from January 22, 2020 to February 5, 2020 and, in the same way, a steady decreasing trend from February 6, 2020 to February 16, 2020. Interestingly, on February 17, 2020, the number of cases worldwide (19,572 cases) and in China (19,461 cases) soared suddenly; however, the number started declining up until February 25, 2020 and then started rising sharply and reached to peak on March 20, 2020 with 43,957 cases. Fascinatingly, after February 25, 2020, cases outside China increased more than those in China, which was an alarming signal that COVID-19 became a pandemic [**Fig. 5(i) and Supplementary Fig. 1**].

3.5. Burden of COVID-19: Death toll

The first death by COVID-19 occurred on January 11, 2000, and by January 19, 2020, another three persons had died. The global death toll (Fig. 6C) started to increase from January 20, 2020 until February 12, 2020 and then followed an up and down trend until February 22, 2020. The death toll peaked with 158 deaths in a single day on February 23, 2020. Later, it declined up to February 26, 2020 and again rose in a fluctuating trend up to March 07, 2020. Then, the death toll peaked with 1080 deaths in a single day on March 19, 2020 (Fig. 6C), and the cumulative number of deaths up to March 20, 2020 was 10,406 (Supplementary Fig. 2). Among 182 countries and territories, 67 countries had experienced deaths from the COVID-19 outbreak by March 20, 2020.The number of deaths was highest in Italy (3405). In addition,

substantial deaths occurred in China (3254), Iran (1433) and Spain (1002) (51) [Fig. 5(ii) and Supplementary File 3]. In China, 28 of 34 provinces experienced deaths from COVID-19, and the highest death toll occurred in Hubei (3122) Province, followed by Henan (22) and Heilongjiang (13); in other provinces, the death toll was below ten up until March 19, 2020 (Supplementary File 4).

3.6. Burden of COVID-19: Crude mortality rate

Upon analysis of mortality data over the period from January 11, 2020 to March 19, 2020, we found the highest mortality in San Marino with 4129 deaths per 10,000,000 persons, followed by Italy (563 deaths per 10,000,000 persons), Spain (214 deaths per 10,000,000 persons), Iran (170 deaths per 10,000,000 persons), Cayman Islands (153 deaths per 10,000,000 persons). However, the mortality in other countries ranged from one to sixty-four deaths per 10,000,000 person (Supplementary Fig. 3 and Supplementary File 5). Chinese record [Fig. 5(ii)B] demonstrates that among the 28 provinces of China, the highest mortality was observed in Hubei Province (528 deaths per 10,000,000 persons) over the period from January 11, 2020 to March 19, 2020. However, in Hainan, Hong Kong, Beijing, Heilongjiang, Henan, Chongqing, Tianjin, Shanghai, and Xinjiang, the mortality ranged from one to seven deaths per 10,000,000 persons. In the rest of the provinces, the mortality rates were below one per 10,000,000 persons (Supplementary File 4).

3.7. Burden of COVID-19: Case fatality

From the analysis of cases and death toll data up to March 20, 2020, we observed that global case fatality was 4.10% (Fig. 6D). Among the 67 countries with morality reports, the highest case fatality was 50.00% in Sudan, while the lowest in Malaysia as 0.19% (Fig. 6D). Case fatality was also high in the Cayman Islands (33.33%), Curacao (33.33%), Gabon (33.33%), and Guyana (20.00%). However, moderate case fatality was observed in Ukraine (11.54%), Algeria (11.11%), Guatemala (11.11%), and San Marino (9.72%). Besides, case fatality was 4.00%, 8.30%, 5.02%, and 7.29% in China, Italy, Spain and Iran, respectively (Fig. 6D and Supplementary File 6). Among 28 provinces of China with death outcomes, the highest case fatality was 4.60% in Hubei Province; however, the lowest was 0.08% in Zhejiang up to March 19, 2020 (Supplementary File 4).

3.8. Recent temperatures in the regions of interest

The satellite-derived temperatures (°C) for Wuhan (China), South Korea, Italy and Bangladesh during February 2020 are shown in Fig. 7. In this figure, a single scale at the rightmost part of the image is used to indicate the temperature for all subimages. However, for the Hubei province of China many parts were cloudy in the Landsat-8 images during the study period. Therefore, only the cloud free Wuhan area of Hubei province is shown in the map. During the study period in the mid-region of Korea, the temperature was very low, which was caused by the presence of heavy and widespread clouds in that region during satellite image acquisition. However, very few clouds covers were found for the Landsat-8 image acquisition for February 2020 for the Italy areas. In almost all areas temperature were lower than 20 °C except a few places where the temperature did not exceeded 25 °C. Therefore, interpretations from the figures for these regions should be guarded to avoid errors.

4. Discussion

The COVID-19, caused by novel coronavirus SARS-CoV-2, became a pandemic (Jung et al., 2020). Coronaviruses are one of the most diverse groups of viruses and have emerged as deadly pathogens for the human race, with SARS-CoV-2 being the latest inclusion to the list. In most cases, human coronaviruses evolve from zoonotic transmission, and bats are reported as the reservoirs for zoonotic viruses (Graham and

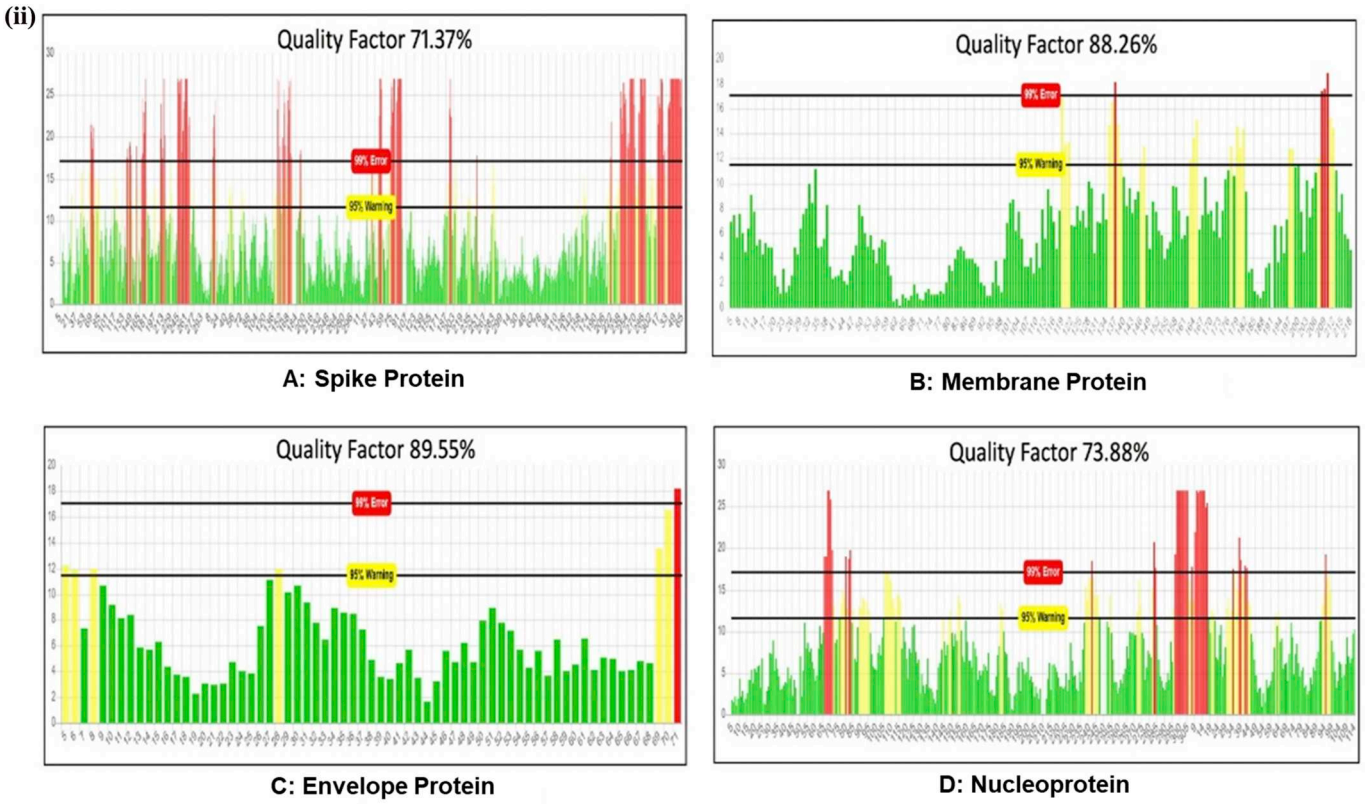

Fig. 3. (*continued*)

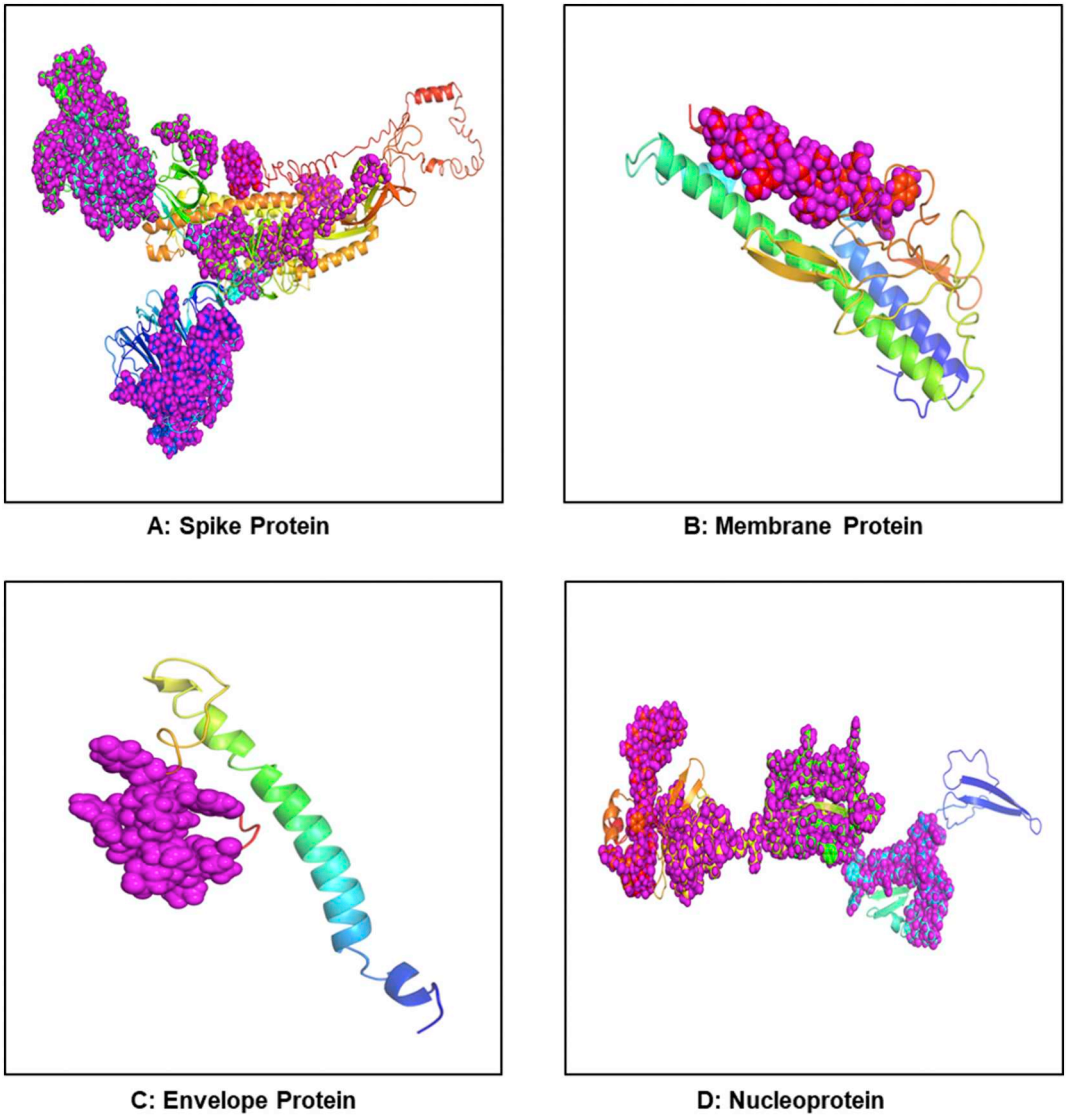

Fig. 4. Representation of predicted 3D model of SARS-CoV-2 proteins. 3D structure of (A) Spike protein, (B) Membrane protein, (C) Envelope protein and (D) Nucleoprotein of COVID-19 proteins with antigenic sites visualized by PyMOL; Ball like structure represents the immunogenic parts.

Baric, 2010). Although the causative agent of COVID-19 has already been declared SARS-CoV-2 (Lai et al., 2020; Phan, 2020), we adopted a protein-based phylogenetic study to reveal the phylogenetic relationships and divergent links of SARS-CoV-2 with other human coronavirus strains. Seven coronaviruses (COVID-19, HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HKU1 and MERS-CoV), two bat-originated coronaviruses (HKU4 and HKU5) and one buffalo-originated coronavirus (BufCoV-HKU26) reported from Bangladesh were subjected to phylogenetic analysis of four types of their viral proteins (S, M, E and N proteins). It was demonstrated that SARS-CoV-2 and SARS-CoV shared a similar pattern of ancestral origin (Fig. 1), and pairwise alignment indicated maximum similarity and identity between these two viruses (Table 1 and Fig. 2). Recently, after the rapid transmission of SARS-CoV-2 and the resulting deaths (Bogoch et al., 2020; Chinazzi et al., 2020; Wu et al., 2020; WHO, 2020c), the WHO announced the genetic similarities between SARS-CoV-2 and SARS-CoV, which were clearly reflected in the present study. Also, HKU4 and HKU5 are bat-originated coronaviruses, and the similar genetic makeup of the other viral strains could indicate a similar host preference (Desjardins et al., 2010), though there were a few unclear reports regarding the possibilities of bats as reservoirs of SARS-CoV-2 (Benvenuto et al., 2020; Giovanetti et al., 2020). However, the divergent phylogenetic relationships and minor alignment patterns of SARS-CoV-2 with the bat-originated viruses HKU4 and HKU5 suggested that there was a low possibility of SARS-CoV-2 being connected with bats (Fig. 1 and Table 1).

In the NCBI database, the gene sequence of a buffalo coronavirus (BufCoV HKU26) from domestic water buffaloes (*Bubalus bubalis*) in Bangladesh is available (Lau et al., 2016). A comparative study of BufCoV-HKU26 with SARS-CoV-2 was performed in this study using phylogenetic analysis and pairwise alignment of the targeted viral proteins. The study revealed that BufCoV-HKU26 showed divergent relationships and low pairwise alignment of S (29.5% identity and 44.2% similarity), M (37.4% identity and 55.7% similarity), E (20.9% identity and 44.2% similarity) and N proteins (32.7% identity and 48.6% similarity) with SARS-CoV-2. However, as BufCoV-HKU26 is an animal-originated coronavirus, its proteins always exhibited a familial relationship with the other animal-originated virus HCoV-OC43 (Fig. 1). BufCoV-HKU26 was detected in Bangladesh, which suggests that the country does provide a favorable environment for the transmission and pathogenesis of buffalo-originated coronaviruses (Rowe and East, 1997; Sooryanarain and Elankumaran, 2015). The genetic architecture of SARS-CoV-2 was highly divergent from that of BufCoV-HKU26 (Figs. 1 and 2 and Table 1). The genetic makeup of the virus directly or indirectly governs viral transmission or pathogenesis.

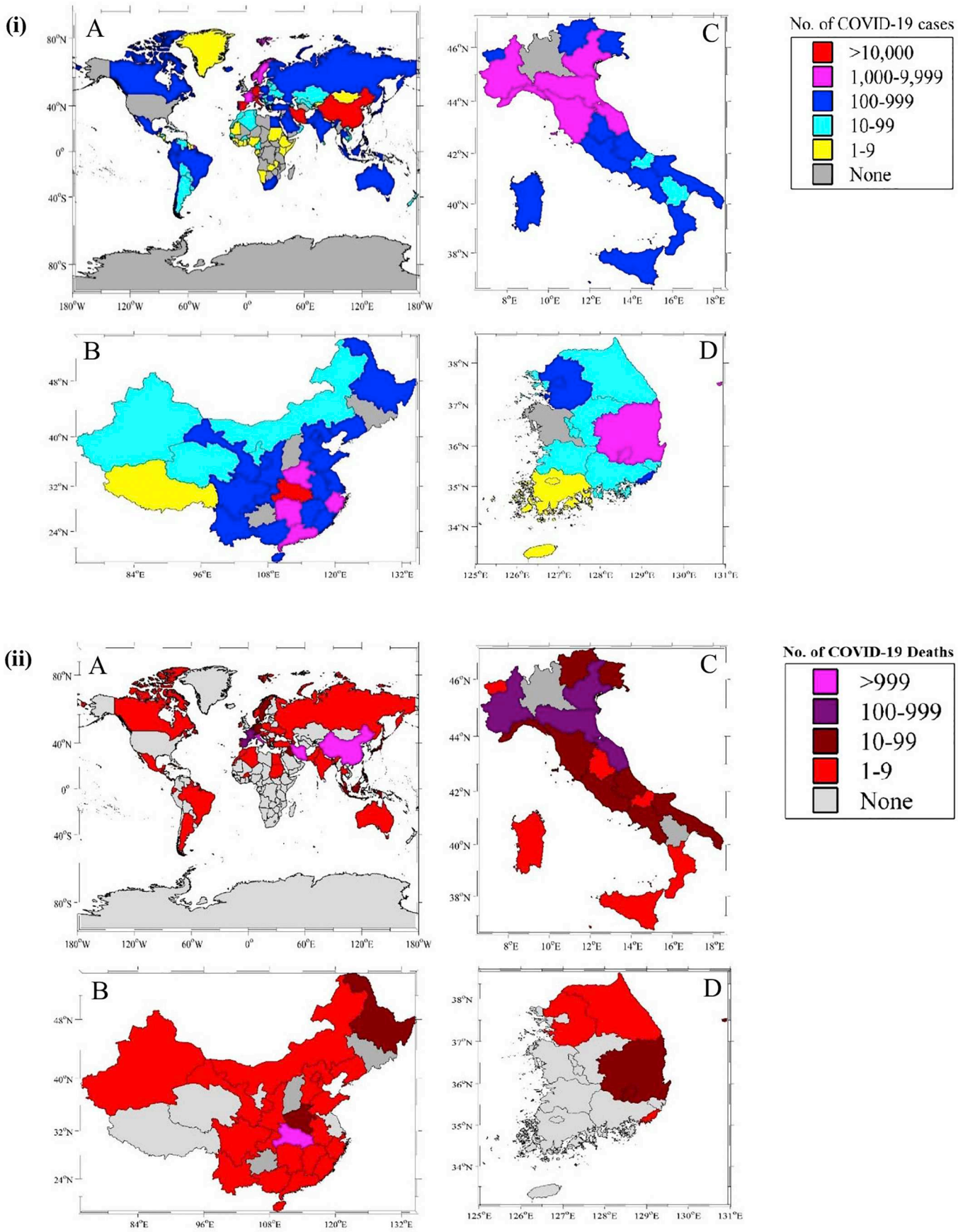

Fig. 5. (i) Geographic distribution of COVID-19 cases as of March 20, 2020. (A) 182 Countries and territories reporting cases and cases reported in (B) China, (C) Italy and (D) Korea. (ii) Geographic distribution of COVID-19 Deaths as of March 20, 2020. (A) 182 Countries and territories death cases and deaths reported in (B) China, (C) Italy and (D) Korea.

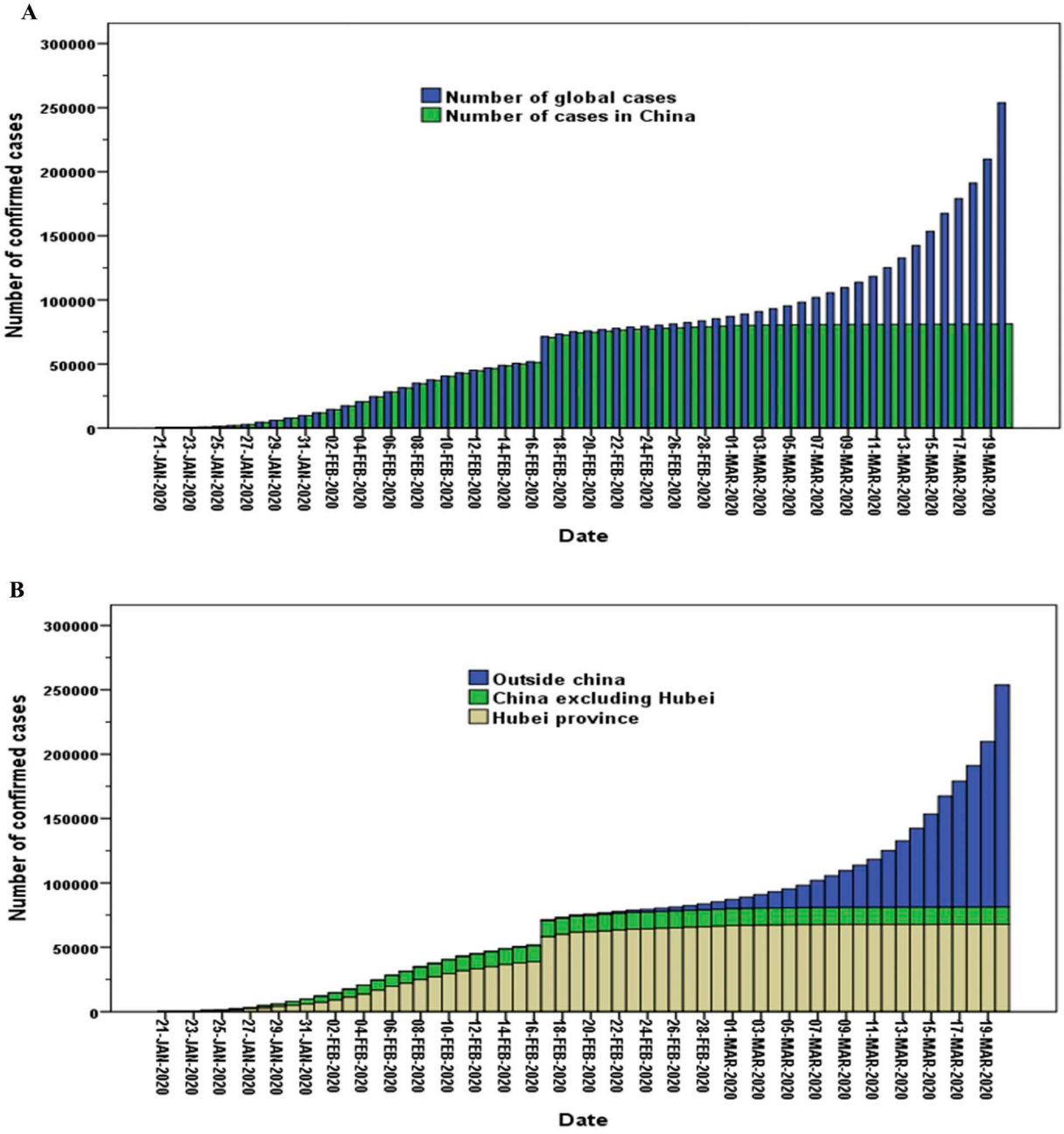

Fig. 6. (A) Cumulative confirmed cases of COVID-19 in globe and China* from January 21, 2020 to March 20, 2020**. (B) Cumulative confirmed cases of COVID-19 in Hubei province, China* excluding Hubei and outside china from January 21, 2020 to March 20, 2020**. (C) Cumulative global death toll of COVID-19 from January 11, 2020 to March 20, 2020*. (D) Global and country-wise case fatality of COVID-19 up to March 20, 2020*. *Confirmed cases in China include cases confirmed in Hong Kong, Macau, and Taiwan; ** Data from WHO (WHO, 2020c).

COVID-19 has already spread to all continents except Antarctica. Movement of people from one place to another allows the rapid spread of any infectious disease from the place of origin to all over the world. Geographically, Bangladesh is located in a temperate zone. We performed a comparative study of the temperatures between Bangladesh, Wuhan, South Korea, and Italy. The temperature could be a factor reducing the risks of COVID-19 to Bangladesh, although the country has a high-risk index (Haider et al., 2020a). The high temperature and high humidity can effectively reduce the environmental persistence and transmission of the COVID-19 (Wang et al., 2020a), and temperature variation and humidity may be important factors affecting the COVID-19 mortality (Ma et al., 2020). However, experts cannot yet predict the trajectory of the virus. COVID-19 (SARS-CoV-2) might have different persistence times on inanimate surfaces, but these times should be similar to those of SARS-associated viruses (3-4 days) (Kramer et al., 2006). Human coronaviruses such as Middle East Respiratory Syndrome (MERS) coronavirus, Severe Acute Respiratory Syndrome (SARS) coronavirus, and endemic Human Coronaviruses (HCoV) can persist on inanimate surfaces but not for long periods (Chan et al., 2011; Kampf et al., 2020). However, at lower temperatures, such as 4°C or less, this virus persisted for as long as 28 days, and the lowest level of inactivation occurred at 20% relative humidity. Inactivation was more rapid at 20°C than at 4°C at all humidity levels. Likewise, coronaviruses were inactivated more rapidly at 40°C than at 20°C (Kampf et al., 2020).

Recently, the authors established that temperature significantly changed COVID-19 transmission in 429 cities, in which COVID-19 appeared to spread fastest at 8.72°C (Wang et al., 2020b). However, based on this brief discussion, it can be presumed that this virus has a lower

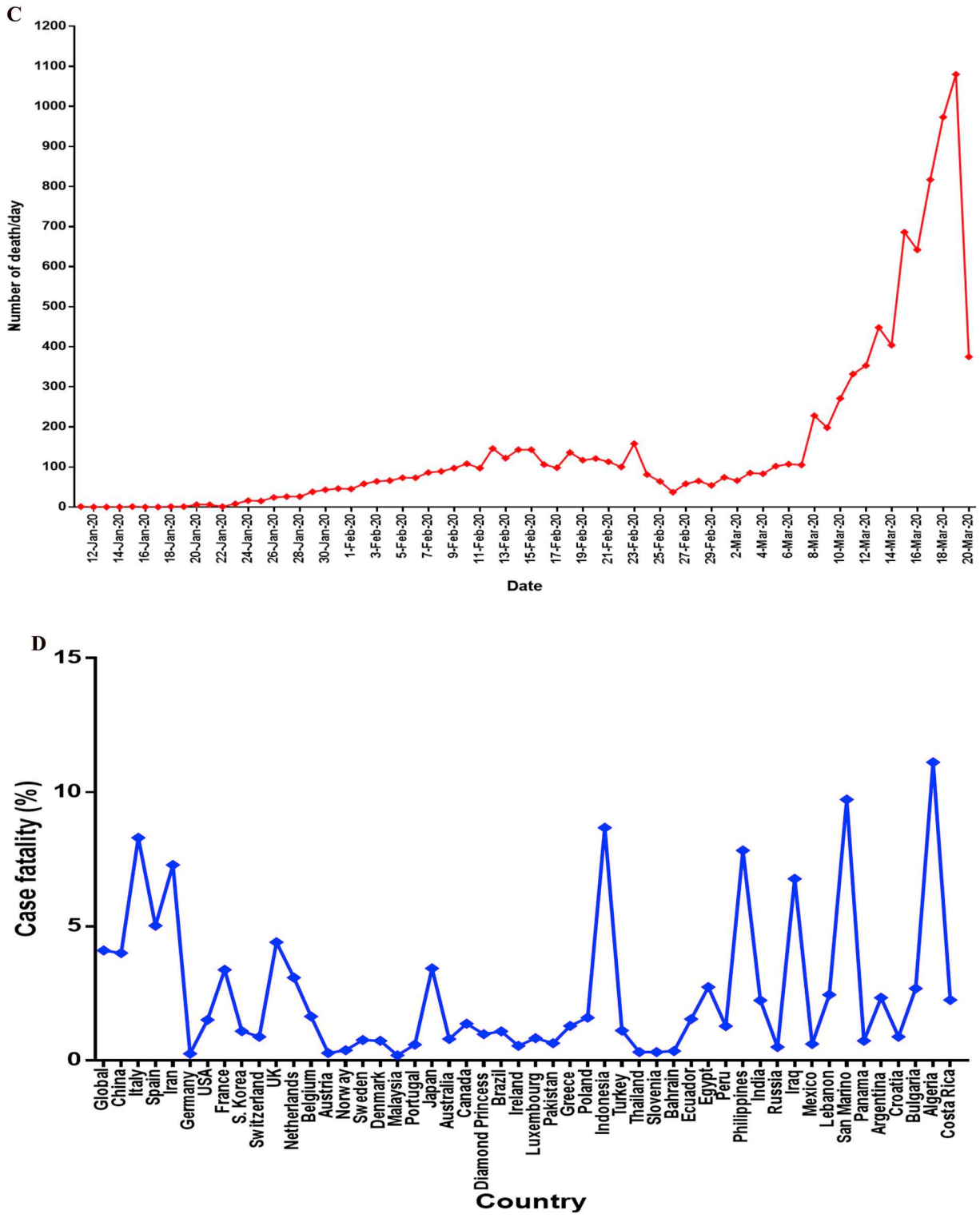

Fig. 6. (*continued*)

possibility of transmission and severity in countries or seasons with relatively higher temperatures, respectively. Thus, in affected countries, COVID-19 might disappear with rising temperatures when spring comes. In Fig. 7, the temperatures of Wuhan (China), South Korea, Italy and Bangladesh during February 2020 are compared. It can be observed that the temperature ranged from 13 to 22°C in many places of Wuhan, Korea as well as in Italy. However, across Bangladesh, the temperatures were above 25°C except for very small areas with temperatures of approximately 20°C, which were caused by small patches of cloud cover. Thus, the temperature is much higher in Bangladesh than in Wuhan, Korea, and Italy. In March, the temperature has increased considerably, giving hope that the transmission of the virus will be limited. Only 20 cases (including one death) of COVID-19 have been confirmed in Bangladesh as of March 20, 2020. The largest international airport in Bangladesh is situated in Dhaka and is the main entrance for people from all over the world. As Dhaka is an industrial and commercial zone and all the official works of Bangladesh are centralized here, most of the visitors or people returning to the country stay here for an extended

Table 2
Predicted antigenic sites (B cell epitope) of COVID-19 with VaxiJen score and conservancy analysis with homologous protein sets.

Protein	Peptide	Length (aa)	Vaxijen Score	Percent of protein sequence matches at identity ≤ 100%	Minimum identity	Maximum identity
Spike protein	SQCVNLTTRTQLPPAYTNSFTRGVY	25	0.6860	83.33% (5/6)	96.00%	100.00%
	FSNVTWFHAIHVSGTNGTKRFDN	23	0.6767	100.00% (6/6)	100.00%	100.00%
	DPFLGVYYHKNNKSWME	17	0.5821	100.00% (6/6)	100.00%	100.00%
	SQCVNLTTRTQLPPAYTNSFTRGVY	25	0.6860	83.33% (5/6)	96.00%	100.00%
	MDLEGKQGNFKNL	13	1.2592	100.00% (6/6)	100.00%	100.00%
	KHTPINLVRDLPQGFS	16	0.6403	83.33% (5/6)	93.75%	100.00%
	KSFTVEKGIYQTSNFRVQP	19	0.5729	100.00% (6/6)	100.00%	100.00%
	FPNITNLCPFGEVFNATRFASVYAWNRKRISNCVA	35	0.4466	100.00% (6/6)	100.00%	100.00%
	YNSASFSTFKCYGVSPTKLNDLCFT	25	1.4031	100.00% (6/6)	100.00%	100.00%
	GDEVRQIAPGQTGKIADYNYKLP	23	1.1017	100.00% (6/6)	100.00%	100.00%
	NLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTN	62	0.3951	100.00% (6/6)	100.00%	100.00%
	SNKKFLPF	8	1.3952	100.00% (6/6)	100.00%	100.00%
	NCTEVPVAIHADQLTPT	17	0.3987	100.00% (6/6)	100.00%	100.00%
	VNNSYECDIPI	11	0.6124	100.00% (6/6)	100.00%	100.00%
	YTMSLGAENSVAYSNN	16	0.6434	100.00% (6/6)	100.00%	100.00%
	GQSKRVDFC	9	1.7790	100.00% (6/6)	100.00%	100.00%
	SCCKFDEDDSEPVLKG	16	0.4347	100.00% (6/6)	100.00%	100.00%
Nucleo-protein	HGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLS	47	0.5773	100.00% (6/6)	100.00%	100.00%
	TTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGD	52	0.5206	33.33% (2/6)	15.38%	100.00%
	RLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKA	42	0.5627	83.33% (5/6)	16.67%	100.00%
	RRGPEQTQGNFGDQELIRQGTDYK	24	0.6277	83.33% (5/6)	20.83%	100.00%
	DAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLLPAADLDD	45	0.4968	83.33% (5/6)	15.56%	100.00%
Membrane protein	YRIGNYKLNTDHSSSSDNIA	20	0.2216	0.00% (0/2)	19.05%	95.24%
Envelop protein	YVYSRVKNLNSSRVP	15	0.4492	100.00% (2/2)	100.00%	100.00%

Table 3
Conservancy analysis of Antigenic Sites from COVID-19 proteins with other coronaviruses corresponding proteins.

Antigenic Sites of COVID-19	*Conservancy with other coronaviruses corresponding proteins (%)*								
	SARS-CoV	*Buffalo BufCoV-HKU26*	*BatCoV-HKU5*	*BatCoV-HKU4*	*MERS-CoV*	*HCoV-HKU1*	*HCoV-NL63*	*HCoV-OC43*	*HCoV-229E*
S proteins									
SQCVNLTTRTQLPPAYTNSFTRGVY	28.00	24.00	28.00	24.00	28.00	24.00	28.00	28.00	28.00
FSNVTWFHAIHVSGTNGTKRFDN	39.13	26.09	26.09	26.09	30.43	26.09	26.09	26.09	26.09
DPFLGVYYHKNNKSWME	29.41	29.41	23.53	29.41	29.41	29.41	29.41	29.41	29.41
SQCVNLTTRTQLPPAYTNSFTRGVY	28.00	24.00	28.00	24.00	28.00	24.00	28.00	28.00	28.00
MDLEGKQGNFKNL	53.85	30.77	30.77	30.77	38.46	30.77	30.77	30.77	38.46
KHTPINLVRDLPQGFS	56.25	31.25	31.25	31.25	25.00	31.25	37.50	31.25	37.50
KSFTVEKGIYQTSNFRVQP	78.95	31.58	42.11	31.58	47.37	42.11	26.32	31.58	36.84
FPNITNLCPFGEVFNATRFASVYAWNRKRISNCVA	88.57	25.71	22.86	20.00	22.86	25.71	25.71	25.71	22.86
YNSASFSTFKCYGVSPTKLNDLCFT	84.00	28.00	32.00	32.00	28.00	32.00	28.00	24.00	28.00
GDEVRQIAPGQTGKIADYNYKLP	91.30	26.09	34.78	39.13	30.43	26.09	26.09	26.09	21.74
NLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTN	33.87	17.74	17.74	17.74	17.74	17.74	17.74	17.74	17.74
SNKKFLPF	62.50	37.50	50.00	50.00	50.00	50.00	37.50	37.50	50.00
NCTEVPVAIHADQLTPT	76.47	29.41	35.29	35.29	41.18	29.41	29.41	29.41	29.41
VNNSYECDIPI	81.82	45.45	36.36	45.45	27.27	45.45	36.36	45.45	36.36
YTMSLGAENSVAYSNN	81.25	31.25	31.25	31.25	31.25	31.25	43.75	37.50	31.25
GQSKRVDFC	100.00	55.56	66.67	66.67	66.67	55.56	55.56	55.56	66.67
SCCKFDEDDSEPVLKG	100.00	31.25	31.25	31.25	31.25	31.25	31.25	31.25	31.25
M proteins									
HGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLS	89.36	17.78	48.94	46.81	48.94	31.91	19.15	21.28	19.15
TTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARMAGNGGD	88.46	29.17	44.23	42.31	50.00	26.92	25.00	28.85	26.92
RLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKA	95.24	26.19	33.33	40.48	50.00	28.57	21.43	26.19	21.43
RRGPEQTQGNFGDQELIRQGTDYK	95.83	30.77	41.67	45.83	50.00	29.17	37.50	29.17	29.17
E proteins									
YRIGNYKLNTDHSSSSDNIA	85.00	30.00	20.00	25.00	20.00	30.00	15.00	30.00	20.00
N proteins									
YVYSRVKNLNSSRVP	80.00	20.00	20.00	20.00	20.00	40.00	20.00	20.00	26.67

period. Due to industrialization, the temperature of Dhaka is much higher than that of any other parts of Bangladesh, which is also evident from Fig. 7 (D) (indicated by the black circle in the figure). Many of the scientists speculated and revealed that high temperature and humidity have a role in restricting the spread of COVID-19, and the spread of disease would be suppressed in hot and humid weather (Sajadi et al., 2020; Zhang and Zeng, 2020), are in the line of our hypothesis. Coronaviruses that cause common colds do tend to subside in warmer

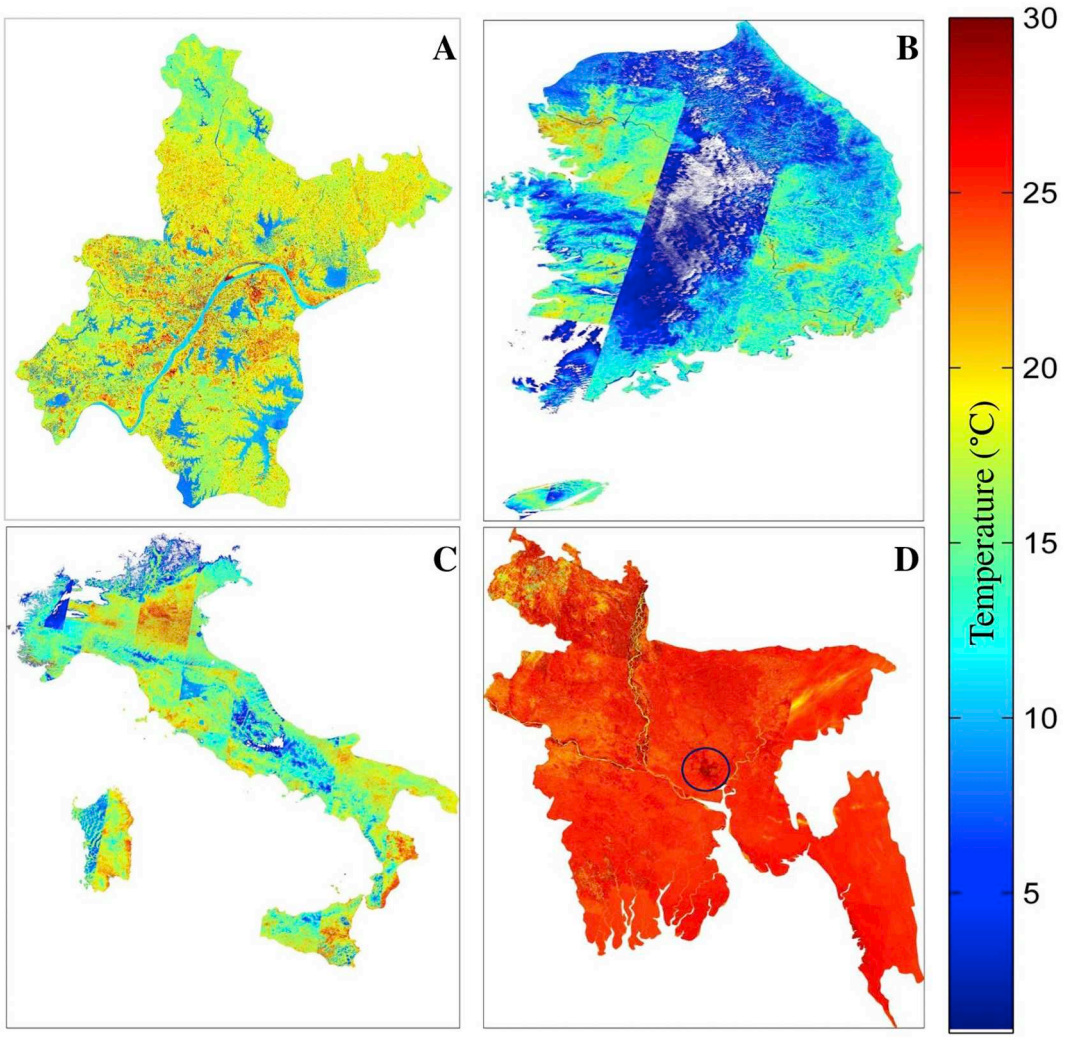

Fig. 7. Temperature (°C) for (A) Wuhan (China), (B) South Korea, (C) Italy and (D) Bangladesh during February 2020.

months. However, it is highly uncertain whether SARS-CoV-2 will behave the same way. Current research by scientists is too early to predict how the virus will respond to changing weather (National geographic, 2020).

Moreover, scientists discovered that SARS-CoV-2 has mutated into a more aggressive disease (Tang et al., 2020). Population genetic analyses of SARS-CoV-2 genomes indicate two major lineages (designated L and S) with a higher possibility of infection by both types. The ancestral 'S-type' is suspected to be less infectious and milder in severity (~30%), while the later emerging 'L-type' is more prevalent (~70%) (Tang et al., 2020). If the evolution of SARS-CoV-2 continues, countries such as Bangladesh will be in great trouble as a human intervention may have exerted more severe selective pressure on the L type, which will help generating mutant virus that will be be more aggressive and will spread more quickly (Tang et al., 2020). Although, according to the WHO, this virus is unique in its continuous mutation (WHO, 2020a); until now, there has been no concrete evidence that the virus has already changed with regards to disease severity or infectivity. More comprehensive genomic data are required for further testing of this hypothesis.

Immunogenicity and epitope conservancy analyses of coronavirus proteins were performed to determine the potential B-cell epitopes that would interact efficiently with B lymphocytes to initiate the immune response against specific viral pathogens (Klingen et al., 2018). The study identified a total of 24 highly immunogenic B-cell epitopes from SARS-CoV-2 proteins (17 epitopes of S; 1 epitope of M; 1 epitope of E; 5 epitopes of N proteins). Conservancy analysis (Table 3) demonstrated that selected epitopes were highly conserved in other homologous strains of SARS-CoV-2, which suggested the possibility of peptide-based vaccine development from S, M, N or E proteins (Azim et al., 2019; Abdulla et al., 2019; Preethika Nair, 2020). Moreover, potential B-cell epitopes were also identified from HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HKU1, MERS-CoV, HKU4, HKU5 and BufCoV-HKU26 proteins (Supplementary Table 1). The antigenic sites of COVID-19 were also crosschecked with other coronavirus-corresponding proteins (Table 3). Higher levels of conservation among antigenic epitope sites of SARS-CoV-2 and SARS-CoV supported a similar ancestry in their pathogenicity. In contrast, antigenic epitopes from SARS-CoV-2 exhibited the minimum conservancy pattern with BufCoV-HKU26, which suggested a distinct pathogenicity pattern between SARS-CoV-2 and BufCoV-HKU26.

The national health authorities of China detected a cluster of 44 pneumonia patients of unknown etiology from December 31, 2019 to January 3, 2020 in Wuhan City, Hubei Province of China and reported this detection to the WHO Chinese office (WHO, 2020c). Moreover, on January 7, they identified the pathogen as a novel coronavirus (World Health Organization (WHO), 2020). Similar to COVID-19, the world experienced an epidemic of severe acute respiratory syndrome (SARS) in 2003 caused by a coronavirus also originating from China (Guangdong Province), which was suggested to be a potential zone of re-emergence of coronaviruses (Lai et al., 2020; WHO, 2020d). However, on March 14, 2020 WHO declared Europe as new epicentre of COVID-19 outbreak (www.who.int).

Within a very short time, COVID-19 was transmitted globally (WHO, 2020c; Wu et al., 2020; Bogoch et al., 2020; Chinazzi et al., 2020; World Health Organization (WHO), 2020). By March 20, 2020, the viruses had spread to up to 182 countries and territories, and one cruise ship named 'Diamond Princess' in Japan (Worldometers. COVID-19 Coronavirus outbreak, 2020), infecting 2,53,796 persons globally (WHO, 2020c). In contrast, SARS infected 8450 persons and caused 810 deaths globally in 33 countries (Chen et al., 2005). Moreover, MERS, another coronavirus associated with an epidemic, started in Saudi Arabia in September 2012 and spread to up to 27 countries affecting a total of 2494 persons and causing 858 deaths throughout the world up to November 2019 (WHO, 2020b). Therefore, COVID-19 has already covered approximately 5.52 times more countries than SARS and MERS (WHO, 2020c; Chen et al., 2005). In addition, up to March 20, 2020, COVID-19 has been attributed 30.03 and 101.76 times more cases, and 12.85 and 12.13 times more deaths globally than SARS and MERS, respectively (Worldometers. COVID-19 Coronavirus outbreak, 2020; Chen et al., 2005; WHO, 2020b). Although the occurrence of cases per day was high in China at the beginning of the outbreak, after February 25, 2020, the number of cases declined in China and dramatically increased outside of China. For example, all of Italy has been placed under lockdown due to COVID-19. The majority of the cases are in mainland China, followed by Italy, Spain, Iran, Germany, USA, France and Korea and the World Health Organization has formally declared the COVID-19 outbreak a pandemic in March 11, 2020 (WHO, 2020c; Jung et al., 2020).

The case fatality of COVID-19 as of March 19, 2020 was 4.10% and 4.00% globally and in China, respectively. This calculation agrees with the report of Wang et al. (Wang et al., 2020c), who stated that the global case fatality was close to 3%. However, the global case fatalities of SARS (9.60%) and MERS (34.40%) were higher than that of COVID-19 (WHO, 2020b; Cameron et al., 2008). By February 2020, the case fatality of China was 3.80% (Worldometers. COVID-19 Coronavirus outbreak, 2020), which supports our estimate. Despite the lower case fatality than the previous coronaviruses SARS and MERS, COVID-19 poses a great burden across the world in terms of geographical coverage, the occurrence of cases, and the cumulative death toll. It became a pandemic, causing great public panic.

Over 182 countries and territories including Bangladesh reported COVID-19 cases, with more than 2,53,796 people infected globally as of March 20, 2020 (Global Health Policy, 2020; WHO, 2020c). The COVID-19 coronavirus has spread in the neighboring countries of Bangladesh as well. Even the coronavirus is sensitive to high temperatures, but do not stop transmitting even in on summer months (Wang et al., 2020; WHO, 2020c). The government of Bangladesh should place preventive measures immediately to prevent the community spread of SARS-CoV-2. Further comprehensive studies that will combine genomic data, outbreak information, surveillance data, travel restrictions will help to generate information that will be necessary for a well-planned long-term strategy to mitigate COVID-19.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2020.104440.

Funding

No funding was received for this work.

Data availability

All data generated and analyzed during this study is included in the main manuscript or supplementary files.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the American Journal Experts (www.aje.com) for their voluntary editing services.

References

Abdulla, F., Adhikari, U.K., Uddin, M.K., 2019. Exploring T & B-Cell Epitopes and Designing Multi-Epitope Subunit Vaccine Targeting Integration Step of HIV-1 Lifecycle Using Immunoinformatics Approach. Microbial Pathogenesis. Elsevier Ltdhttps://doi.org/10.1016/j.micpath.2019.103791.

Azim, K.F., Hasan, M., Hossain, M.N., Somana, S.R., Hoque, S.F., Bappy, M.N.I., et al., 2019. Immunoinformatics approaches for designing a novel multi epitope peptide vaccine against human norovirus (Norwalk virus). Infect. Genet. Evol. 74, 103936. https://doi.org/10.1016/j.meegid.2019.103936.

Benvenuto, D., Giovanetti, M., Salemi, M., Prosperi, M., De, Flora C., Carlos, L., et al., 2020a. The global spread of 2019-nCoV : a molecular evolutionary analysis. Pathog. Glob. Health. 00, 1–4. https://doi.org/10.1080/20477724.2020.1725339.

Benvenuto, D., Giovanetti, M., Salemi, M., Prosperi, M., De Flora, C., Junior Alcantara, L.C., et al., 2020b. The global spread of 2019-nCoV: a molecular evolutionary analysis. Pathog. Glob. Health. 00, 1–4. https://doi.org/10.1080/20477724.2020.1725339.

Bogoch, I.I., Watts, A., Thomas-Bachli, A., Huber, C., Kraemer, M.U.G., Khan, K., 2020. Potential for global spread of a novel coronavirus from China. J. Travel. Med. 1–3. https://doi.org/10.1093/jtm/taaa011.

Cahyono, A.B., Saptarini, D., Pribadi, C.B., Armono, H.D., 2017. Estimation of sea surface temperature (SST) using Split window methods for monitoring industrial activity in coastal area. Appl. Mech. Mater. 862, 90–95. https://doi.org/10.4028/www.scientific.net/amm.862.90.

Cameron, M.J., Bermejo-Martin, J.F., Danesh, A., Muller, M.P., Kelvin, D.J., 2008. Human immunopathogenesis of severe acute respiratory syndrome (SARS). Virus Res. 133, 13–19. https://doi.org/10.1016/j.virusres.2007.02.014.

CDC, 2020. Lesson 3: Measures of Risk, Section 3: Mortality Frequency Measures. https://www.cdc.gov/csels/dsepd/ss1978/lesson3/section3.html.

Chan, K.H., Peiris, J.S.M., Lam, S.Y., Poon, L.L.M., Yuen, K.Y.S.W., 2011. The effects of temperature and relative humidity on the viability of the SARS coronavirus. Adv. Virol. 2011. https://doi.org/10.1155/2011/734690.

Chen, Z., Zhang, L., Qin, C., Ba, L., Yi, C.E., Zhang, F., et al., 2005. Recombinant modified Vaccinia virus Ankara expressing the spike glycoprotein of severe acute respiratory syndrome coronavirus induces protective neutralizing antibodies primarily targeting the receptor binding region. J. Virol. 79, 2678–2688. https://doi.org/10.1128/jvi.79.5.2678-2688.2005.

Chinazzi, M., Davis, J.T., Gioannini, C., Pastore, A., Rossi, L., Xiong, X., et al., 2020. Preliminary Assessment of the International Spreading Risk Associated with the 2019 novel Coronavirus (2019-nCoV) outbreak in Wuhan City. pp. 1–7.

Chojnacki, S., Cowley, A., Lee, J., Foix, A., Lopez, R., 2017. Programmatic access to bioinformatics tools from EMBL-EBI update: 2017. Nucleic Acids Res. 45, W550–W553. https://doi.org/10.1093/nar/gkx273.

City population, 2020. CHINA: Provinces and Major Cities. http://www.citypopulation.de/en/china/cities/ (Retrieved 09 March 2020. 2020).

Countrymeters, 2020. Chaina Population. https://countrymeters.info/en/China.

Desjardins, C.A., Perfectti, F., Bartos, J.D., Enders, L.S., 2010. The genetic basis of interspecies host preference differences in the model parasitoid Nasonia. Heredity (Edinb). 104, 270–277. https://doi.org/10.1038/hdy.2009.145.The.

ECDC, 2020. Outbreak of Acute Respiratory Syndrome Asssociated with a Novel Coronavirus, Wuhan, China; First Update. Rapid Risk Assess.

Flower, D.R., Doytchinova, I., Zaharieva, N., Dimitrov, I., 2017. Immunogenicity prediction by VaxiJen: a ten year overview. J. Proteomics Bioinform. 10, 298–310. https://doi.org/10.4172/jpb.1000454.

Giasuddin, M., Mahmud, M., Al, Asari M.A.M., Akter, S., 2017. Occurrence of foot and mouth disease (FMD) during 2014-2016 in cattle of Sirajganj district, Bangladesh. Jahangirnagar Univ. J. Biol. Sci. 6, 45–49. https://doi.org/10.3329/jujbs.v6i1.33730.

Giovanetti, M., Benvenuto, D., Angeletti, S., Ciccozzi, M., 2020. The first two cases of 2019-nCoV in Italy: where they come from? J. Med. Virol. 1–4. https://doi.org/10.1002/jmv.25699.

Global Health Policy, 2020. COVID-19 Coronavirus Tracker. https://www.kff.org/global-health-policy/fact-sheet/coronavirus- tracker/-Updated as of April 29. .

Graham, R.L., Baric, R.S., 2010. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. J. Virol. 84, 3134–3146. https://doi.org/10.1128/jvi.01394-09.

Gundampati, R.K., Chikati, R., Kumari, M., Sharma, A., Pratyush, D.D., Jagannadham, M.V., et al., 2012. Protein-protein docking on molecular models of Aspergillus Niger RNase and human actin: novel target for anticancer therapeutics. J. Mol. Model. 18, 653–662. https://doi.org/10.1007/s00894-011-1078-4.

Haider, A., Simons, D., Osman, A.Y., Ntoumi, F., Zumla, A., Kock, R.N.Y., 2020a. Passengers' destinations from China: low risk of novel coronavirus (2019-nCoV) transmission into Africa and South America. Epidemiol. Infect. 148, e41. https://doi.org/10.1017/S0950268820000424.

Haider, N., Yavlinsky, A., Simons, D., Osman, A.Y., Ntoumi, F., Zumla, A., et al., 2020b.

Passengers' destinations from China: low risk of novel coronavirus (2019-nCoV) transmission into Africa and South America. Epidemiol. Infect.(363). https://doi.org/10.1017/S0950268820000424.
Hasan, M., Azim, K.F., Begum, A., Khan, N.A., Shammi, T.S., Imran, A.S., et al., 2019a. Vaccinomics strategy for developing a unique multi-epitope monovalent vaccine against Marburg marburgvirus. Infect. Genet. Evol. 70, 140–157. https://doi.org/10.1016/j.meegid.2019.03.003.
Hasan, M., et al., 2019b. Reverse vaccinology approach to design a novel multi-epitope subunit vaccine against avian influenza A (H7N9) virus. Microb. Pathog. 130, 19–37. https://doi.org/10.1051/matecconf/201712107005.
Hasan, M., Islam, S., Chakraborty, S., Mustafa, A.H., Azim, K.F., Joy, Z.F., et al., 2019c. Contriving a chimeric polyvalent vaccine to prevent infections caused by herpes simplex virus (type-1 and type-2): an exploratory immunoinformatic approach. J. Biomol. Struct. Dyn. 0, 1–18. https://doi.org/10.1080/07391102.2019.1647286.
Hui, D.S., I Azhar, E., Madani, T.A., Ntoumi, F., Kock, R., Dar, O., Ippolito, G., Mchugh, T.D., Memish, Z.A., 2020. Drosten C ZA. International journal of infectious diseases the continuing 2019-nCoV epidemic threat of novel coronaviruses to global health — the latest 2019 novel coronavirus outbreak in. Int. J. Infect. Dis. 91, 264–266. https://doi.org/10.1016/j.ijid.2020.01.009.
Jacob, ME GM, 2008. Neuroepidemiology. Handbook of Clinical Neurology. https://www.sciencedirect.com/topics/pharmacology https://doi.org/10.1016/s1474-4422(08)70083-8.
Jung, S., Akhmetzhanov, A.R., Hayashi, K., Linton, N.M., Yang, Y., Yuan, B., et al., 2020. Real-time estimation of the risk of death from novel coronavirus (COVID-19) infection: inference using exported cases. J. Clin. Med. 9, 523. https://doi.org/10.3390/JCM9020523.
Kampf, G., Todt, D., Pfaender, S., Steinmann, E., 2020. Persistence of coronaviruses on inanimate surfaces and its inactivation with biocidal agents. J. Hosp. Infect. 104, 246–251. https://doi.org/10.1016/j.jhin.2020.01.022.
Klingen, T.R., Reimering, S., Guzmán, C.A., McHardy, A.C., 2018. In Silico vaccine strain prediction for human influenza viruses. Trends Microbiol. 26, 119–131. https://doi.org/10.1016/j.tim.2017.09.001.
Kramer, A., Schwebke, I., Kampf, G., 2006. How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. BMC Infect. Dis. 6, 1–8. https://doi.org/10.1186/1471-2334-6-130.
Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549. https://doi.org/10.1093/molbev/msy096.
L DeLano, W., 2002. Pymol: an open-source molecular graphics tool. Newsl. Protein Crystallogr. 40.
Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J., Hsueh, P.-R., 2020. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges. Int. J. Antimicrob. Agents 2, 105924. https://doi.org/10.1016/j.ijantimicag.2020.105924.
Larsen, J.E.P., Lund, O., Nielsen, M., 2006. Improved method for predicting linear B-cell epitopes. Immun. Res. 2, 2. https://doi.org/10.1186/1745-7580-2-2.
Lau, S.K.P., Tsang, A.K.L., Ahmed, S.S., Alam, M.M., Ahmed, Z., Wong, P., et al., 2016. First genome sequences of buffalo coronavirus from water buffaloes in. New Microb. New Infect. 11, 54–56. https://doi.org/10.1016/j.nmni.2016.02.011.
Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., et al., 2020. Effects of Temperature Variation and Humidity on the Mortality of COVID-19 in Wuhan. 101101/2020031520036426. .
National geographic. https://www.nationalgeographic.com/science/2020/02/what-happens-to-coronavirus-covid-19-in-warmer-spring-temperatures/.Retrieved 09 March 2020. 2020.
Phan, T., 2020. Genetic diversity and evolution of SARS-CoV-2. Infect. Genet. Evol. 81, 104260. https://doi.org/10.1016/j.meegid.2020.104260.
Pharmaceutical-technology, 2020. Coronavirus in China: the Outbreak, Measures, and Impact. https://www.pharmaceutical-technology.com/features/coronavirus-affected-countries-china-outbreak/.
Preethika Nair, S.U., 2020. Significance of RNA sensors in activating immune system in emerging viral diseases. In: Dynamics of Immune Activation in Viral Diseases. Springer, Singapore, pp. 229–242.
Reliefweb, 2020. Asia Pacific: Coronavirus Outbreak - Information Bulletin. https://reliefweb.int/report/china/asia-pacific-coronavirus-outbreak-information-bulletin.
Rothe, C., Schunk, M., Sothmann, P., Bretzel, G., Froeschl, G., Wallrauch, C., Zimmer, T., Thiel, V., Janke, C., Guggemos, W.S.M., 2020. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. New Engl J Med. https://doi.org/10.1056/NEJMc2001468.
Rowe, J.D., East, N.E., 1997. Risk factors for transmission and methods for control of caprine arthritis-encephalitis virus infection. Vet. Clin. North Am. Food Anim. Pract. 13, 35–53. https://doi.org/10.1016/S0749-0720(15)30363-7.
Sajadi, M.M., Habibzadeh, P., Vintzileos, A., Miralles-wilhelm, F., Amoroso, A., 2020. Temperature, Humidity, and Latitude Analysis to Predict Potential Spread and Seasonality for COVID-19. https//ssrn.com/abstract = 3550308 or http//dx.doi.org/102139/ssrn3550308, March 05. .
Sharmin, R., Bashar, A., Khademul, M., 2014. A Highly Conserved WDYPKCDRA Epitope in the RNA Directed RNA Polymerase Of Human Coronaviruses can be Used as Epitope-Based Universal Vaccine Design. 15. pp. 1–10. https://doi.org/10.1186/1471-2105-15-161.
Shirato, K., Nao, N., Katano, H., Takayama, I., Saito, S., Kato, F., et al., 2020. Development of genetic diagnostic methods for novel coronavirus 2019 (nCoV-2019) in Japan. Jpn. J. Infect. Dis. https://doi.org/10.7883/yoken.JJID.2020.061. JJID. 2020.061.
Sooryanarain, H., Elankumaran, S., 2015. Environmental role in influenza virus outbreaks. Ann. Rev. Anim. Biosci. 3, 347–373. https://doi.org/10.1146/annurev-animal-022114-111017.
Syariz, M.A., Jaelani, L.M., Subehi, L., Pamungkas, A., Koenhardono, E.S., Sulisetyono, A., 2015. Retrieval of sea surface temperature over Poteran Island water of Indonesia with Landsat 8 TIRS image: a preliminary algorithm. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 40, 87–90. https://doi.org/10.5194/isprsarchives-XL-2-W4-87-2015.
Tang, X., Wu, C., Li, X., Song, Y., Yao, X., Wu, X., Duan, Y., Zhang, H., Wang, Y.Q.Z., Cui, J.L.J., 2020. On the origin and continuing evolution of SARS-CoV-2. Natl. Sci. Rev. https://doi.org/10.1093/nsr/nwaa036.
The Daily Star, 2020. Emergency Response Plan Ready for Coronavirus. https://www.thedailystar.net/frontpage/news/emergency-response-plan-ready-govt-1876468.
Thompson, J.D., Gibson, T.J., Higgins, D.G., 2003. Multiple sequence alignment using ClustalW and ClustalX. Curr. Protoc. Bioinformatics 00, 2.3.1–2.3.22. https://doi.org/10.1002/0471250953.bi0203s00.
Vita, R., Overton, J.A., Greenbaum, J.A., Ponomarenko, J., Clark, J.D., Cantrell, J.R., et al., 2015. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 43, D405–D412. https://doi.org/10.1093/nar/gku938.
Wang, W., Xia, M., Chen, J., Deng, F., Yuan, R., Zhang, X., et al., 2016. Data set for phylogenetic tree and RAMPAGE Ramachandran plot analysis of SODs in Gossypium raimondii and G. arboreum. Data Br. 9, 345–348. https://doi.org/10.1016/j.dib.2016.05.025.
Wang, J., Tang, K., Feng, K., Lv, W., 2020a. High Temperature and High Humidity Reduce the Transmission of COVID-19. https//ssrn.com/abstract = 3551767 or http//dx.doi.org/102139/ssrn3551767, March 09. .
Wang, M., Jiang, A., Gong, L., Luo, L., Guo, W., Li, C., Zheng, J., Li, C., Yang, B., Zeng, J., Chen, Y., Zheng, K.L.H., 2020b. Temperature Significantly Change COVID-19 Transmission in 429 Cities. pp. 1–9. https://doi.org/10.1101/2020.02.22.20025791.
Wang, C., Horby, P.W., Hayden, F.G., Gao, G.F., 2020c. A novel coronavirus outbreak of global health concern. Lancet. 395, 470–473. https://doi.org/10.1016/S0140-6736(20)30185-9.
WHO World Health Organization: Virus May Not Disappear in Summer. https://www3.nhk.or.jp/nhkworld/en/news/20200307_07/.
WHO, 2020b. Middle East Respiratory Syndrome Coronavirus (MERS-CoV): MERS Monthly Summary November 2019. https://www.who.int/emergencies/mers-cov/en/.
WHO, 2020c. (World Health Organization): Coronavirus Disease (COVID-2019) Situation Reports. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/.
WHO, 2020d. WHO. SARS (Severe Acute Respiratory Syndrome). https://www.who.int/ith/diseases/sars/en/.
Wikipedia, 2020. Provinces of China. https://en.wikipedia.org/wiki/Provinces_of_China.
World Health Organization (WHO), 2020. Novel Coronavirus (2019-nCoV). WHO Bull. 1–7.
Worldometers. COVID-19 Coronavirus outbreak, 2020. https://www.worldometers.info/coronavirus/.
Wu, J.T., Leung, K., Leung, G.M., 2020. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet. 395, 689–697. https://doi.org/10.1016/S0140-6736(20)30260-9.
Xu, D., Zhang, Y., 2011. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. Biophys. J. 101, 2525–2534. https://doi.org/10.1016/j.bpj.2011.10.024.
Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinform. 9, 1–8. https://doi.org/10.1186/1471-2105-9-40.
Zhang, H.Zeng, Y. Transmissibility of COVID-19 and its Association with Temperature and Humidity. https://www.researchsquare.com/article/rs-17715/v1.
Zhang, J., Liang, Y., Zhang, Y., 2011. Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. Structure. 19, 1784–1795. https://doi.org/10.1016/j.str.2011.09.022.
Zhou, L., Med, M., Tong, Y., Ph, D., Ren, R., Med, M., et al., 2020. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. pp. 1–9. https://doi.org/10.1056/NEJMoa2001316.
Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al., 2020. A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. 727–733. https://doi.org/10.1056/nejmoa2001017.
https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-mission-briefing-on-covid-19—13-march-2020, Accessed date: 29 April 2020.