Genomic Diversity and Evolution, Diagnosis, Prevention, and Therapeutics of the Pandemic COVID-19 Disease

Running Title: **Genomic analysis, diagnosis, and management of COVID-19**

M. Nazmul Hoque[1,2], Abed Chaudhury[3], Md. Abdul Mannan Akanda[2], M. Anwar Hossain[1,4], Md Tofazzal Islam[2*]

[1]Department of Microbiology, University of Dhaka, Dhaka 1000, Bangladesh

[2]Department of Gynecology, Obstetrics and Reproductive Health, Bangabandhu Sheikh Mujibur Rahman Agricultural University (BSMRAU), Gazipur-1706, Bangladesh

[3]SoilCCompany, Charles Sturt University Campus, Orange, NSW, Australia

[2]Department of Plant Pathology, BSMRAU, Gazipur-1706, Bangladesh

[4]Jashore University of Science and Technology, Jashore 7408, Bangladesh

[2]Institute of Biotechnology and Genetic Engineering (IBGE), BSMRAU, Gazipur-1706, Bangladesh

[*]Corresponding author: Md Tofazzal Islam, e-mail: tofazzalislam@yahoo.com

 doi:10.20944/preprints202004.0359.v1

Abstract

A novel coronavirus COVID-19 was first emerged in Wuhan city of Hubei Province in China in December 2019. The COVID-19, since then spreads to 213 countries and territories, and has become a pandemic. Genomic analyses have indicated that the virus, popularly named as corona, originated through a natural process and is probably not a purposefully manipulated laboratory construct. However, currently available data are not sufficient to precisely conclude the origin of this fearsome virus. Genome-wide annotation of thousands of genomes revealed that more than 1,407 nucleotide mutations and 722 amino acids replacements occurred at different positions of the SARS-CoV-2. The spike (S) glycoprotein of SARS-CoV-2 possesses a functional polybasic (furin) cleavage site at the S1-S2 boundary through the insertion of 12 nucleotides. It leads to the predicted acquisition of 3-*O*-linked glycan around the cleavage site. Although real-time RT-PCR methods targeting specific gene(s) have widely been used to diagnose the COVID-19 patients, however, recently developed more convenient, rapid, and specific diagnostic tools targeting IgM/IgG or newly developed plug and play methods should be available for resource-poor developing countries. Some drugs, vaccines and therapies have shown great promise in early trials, however, these candidates of preventive or therapeutic agents have to pass a long path of trials before being released for the practical application against COVID-19. This review updates current knowledge on origin, genomic evolution, development of the diagnostic tools and the preventive or therapeutic remedies of the COVID-19, and discusses on scopes for further research and effective management and surveillance of COVID-19.

Key words: SARS-CoV-2, Genetic Diversity, Genome evolution, Diagnostics, Therapeutics, Vaccines

 doi:10.20944/preprints202004.0359.v1

Introduction

Emergence and reemergence of various pathogens pose global challenges for public health and human food security (Islam et al., 2016; Gao, 2018). The World Health Organization was notified of a cluster of cases of pneumonia disease of unknown etiology on 31st December 2019 in Wuhan of Hubei Province of China. Soon afterwards, the researchers investigated that disease and confirmed that it is a new coronavirus, which causes severe acute respiratory syndrome in the reported patients. Based on phylogenomics and transmission electron microscopic analyses, Zhou et al. (2020a) confirmed the pathogen as a novel coronavirus and named it as 2019-nCoV. Later, this new virus was renamed as SARS-CoV-2 or briefly COVID-19 by the Coronavirus Study Group of the International Committee on Taxonomy of Viruses (ICTV). The etiological agent of COVID-19 is the third devastating coronavirus (CoV) that severely infected human. Earlier, two similar viruses viz. severe acute respiratory syndrome (SARS-CoV) (Zaki et al., 2012; Almofti et al., 2018) and the Middle East respiratory syndrome (MERS-CoV) emerged as epidemics in 2003 and 2012, respectively (Badawi et al., 2016; Pallesen et al., 2017; ul Qamar et al., 2018). This viral disease rapidly spread to almost whole world within four months and poses a serious threat to human health globally. Considering the contagious behavior and fatality of the COVID-19, WHO declared it as a Public Health Emergency of International Concern (WHO, 2020). As of April 18, 2020, the COVID-19 has spread to 213 countries and/or territories, infecting at least 22,14,861 people of which at least 1,50,948 people died globally. The rapidly spreading human-to-human transmission of SARS-CoV-2 has been confirmed by detecting the virus in a wide range of samples including bronchoalveolar-lavage (Zhu et al., 2020; Nishiura et al., 2020), sputum (Lin et al., 2020), saliva (To et al., 2020), throat (Bastola et al., 2020) and nasopharyngeal swabs (To et al., 2020).

The SARS-CoV-2 is a positive-sense single-stranded RNA virus (+ssRNA). It belongs to the Genus *Betacoronavirus* in the family *Coronaviridae*. The *Coronaviridae* is one of the largest

 doi:10.20944/preprints202004.0359.v1

viral families having potential ability to infect, and subsequently causes diseases to a large number of mammals, birds and humans (Ahmed et al., 2020; Hemida and Abduallah, 2020). The coronaviruses manifest a wide variety of clinical signs and symptoms in the affected hosts. These signs include respiratory, nervous, enteric, and systemic health problems. Within weeks of the first outbreak of COVID-19 in Wuhan, the complete genome sequence of this novel virus was published. Open science and open data sharing approaches were practiced to rapidly tackling this deadly enemy of humans. Reminiscent of other SARS-related known coronaviruses (SARSr-CoVs), the viral RNA genome of the novel SARS-CoV-2 encodes several smaller open reading frames (ORFs) (Rota et al., 2003; Freundt et al., 2010; Cotton et al., 2013). These ORFs encode for different proteins such as the replicase polyprotein, the spike (S) glycoprotein, envelope (E), membrane (M), nucleocapsid (N) proteins, accessory proteins, and other non-structural proteins (nsp) (Ahmed et al., 2020; Islam et al., 2020; Phan, 2020; Walls et al., 2020). The genome of this RNA virus is approximately 30 kilobases (between 26,000 and 32,000 bases) in length (Abdelmageed et al., 2020; Rahman et al., 2020). It encodes for multiple structural and non-structural proteins (Ahmed et al., 2020; Rahman et al., 2020) that possess many unique features. These features make these proteins prone to frequent coding changes, thus generating new strains in a short period of time (Hemida and Abduallah, 2020; Islam et al., 2020). Rapid mutational frequencies are associated with the poor proofreading efficiency of the viral RNA polymerase, and the likelihood of recombination between different members of this family (Jackwood et al., 2012; Phan, 2020). Relatively faster spread of SARS-CoV-2 raises intriguing question whether the evolution of this virus is driven by mutations. To address these questions, several recent studies reported that substitution and/or deletion of nucleotides and amino acids (aa) at the entire genome of SARS-CoV-2 is one of the important mechanisms for virus evolution in nature (Huang et al., 2020; Islam et al., 2020; Phan, 2020; Yin, 2020). Due to the practice of open science, the research

 doi:10.20944/preprints202004.0359.v1

on SARS-CoV-2 is the fastest moving subject in the human history. In about four months, thousands of reports and data on genomics, origin, genome evolution, molecular diagnosis and vaccine or therapeutics of COVID-19 have been published (Clover B, 2020; Geo-Vax, 2020; Islam et al., 2020; Phan, 2020; Rahman et al., 2020; Shereen et al., 2020; Shanmugaraj et al., 2020; Walls et al., 2020; Zhang et al., 2020a).

Genomic analyses of the SARS-CoV-2 virus isolated from the patients of different geographic locations e.g. Asia, America and Europe represented this pathogen as a novel virus that is phylogenetically distinct from any other member of the CoVs having little evidence of local/regional adaptation, suggesting that this virus evolution is mainly driven by genetic drift and founder events (Chiara et al., 2020; Huang et al., 2020; Islam et al., 2020; Yin, 2020). Nevertheless, several reports predict possible adaptation at the nucleotide, aa, and structural heterogeneity in the viral proteins, especially the spike (S) protein (Armijos-Jaramillo et al., 2020; Islam et al., 2020; Sardar et al., 2020). Recently, Shen et al. reported even intra-host viral evolution during infection might be related to its virulence, transmissibility, and/or evolution of virus response against the host immune system (Shen et al. 2020).

Like other coronaviruses SARS-CoV-2 demonstrate versatile human-host ranges and tissue tropism (Bastola et al., 2020; Lin et al., 2020; Nishiura et al., 2020; To et al., 2020; Zhu et al., 2020). The initial attachment of the virion to the host cell is mediated by interactions between the S protein and its receptor (Pal et al., 2020). However, the virulence mechanisms of the SARS-CoV-2 is not yet clear (Khan et al., 2020; Zhou et al., 2020b). In recent studies, it is apparent that SARS-CoV-2 uses the similar mechanism that of SARS-CoV used for cell entry (Hoffmann et al., 2020; Hemida and Abduallah, 2020). Therefore, unravelling the cellular factors involved in entry of SARS-CoV-2 might give insights into the transmission of the virus and reveal the targets of therapy

 doi:10.20944/preprints202004.0359.v1

(Hoffmann et al., 2020; Hemida and Abduallah, 2020). The epidemiology of this novel virus is unknown.

The COVID-19 exhibits with a wide range of clinical signs ranging from asymptomatic patients to septic shock, and the dysfunction of multiorgans. The severity of the presentation of clinical symptoms remains as the primary basis of classifying the disease (Li et al., 2020). Initially, the clinical diagnosis of the disease can be made on the basis of epidemiological history and common symptoms of fever, fatigue, dry cough and diarrhea, but its clinical manifestations spectrum ranging from septic shock, multiorgan dysfuntion, and diarrhea (Hassan et al., 2020). The confirmatory diagnosis of COVID-19 can be made with the aid of some auxiliary examinations including nucleic acid detection by polymerase chain reaction especially real-time reverse transcription PCR (RT–PCR), computed tomography (CT)-scan, immune identification technology (Point-of-care Testing (POCT) of IgM/IgG, CRISPR-Cas and blood culture (Ai et al., 2020; Corman et al., 2020; Hindson, 2020; Kellner et al., 2020; Li et al., 2020; Wang et al., 2020a). The most common symptoms in children with confirmed SARS-CoV-2 are fever, dry cough, pharyngeal erythema and fatigue (Passanisi et al., 2020). Moreover, the clinical symptoms and signs of SARS-CoV-2 confirmed patients are highly atypical. Therefore, a convenient, rapid and specific diagnostic protocol is needed for monitoring, surveillance and management of this pandemic disease.

As an emerging pandemic virus, no effective therapeutic drug or vaccines are yet to be available for the treatment of SARS-CoV-2 patients. Currently, some supportive cares are given to the patients such as oxygen therapy, antiviral combination with antibiotic, convalescent plasma therapy, and antifungal treatment, and extra-corporeal membrane oxygenation (ECMO) (Chen et al., 2020; Holshue et al., 2020). Researchers across the globe are searching to find an antiviral drug useful in treating the infection of SARS-CoV-2. They evaluated several drugs or therapies namely,

penciclovir, ribavirin, nitazoxanide, remdesivir (GS-5734), nafamostat, favipiravir (T-750) or Avigan, EIDD-2801, hydroxychloroquine, chloroquine, and convalescent plasma (CP) therapy against the infection of SARS-CoV-2 (Duan et al., 2020; Liu et al., 2020; Martinez, 2020; Wang et al., 2020b). Furthermore, the fast evolution and high genomic disparity of RNA viruses, mutations on the receptor-binding domain (RBD) may help the new strains to get away neutralization mechanism by RBD-targeting antibodies (Rahman et al., 2020). Therefore, non-RBD functional regions of the S glycoprotein could efficiently be selected for developing effective therapeutic and prophylactic interventions against the infection of SARS-CoV-2. Several monoclonal antibodies (mAbs) targeting non-RBD regions, particularly the N-terminal domain (NTD) has recently been reported (Shang et al., 2020; Wang et al., 2019; Zhou et al., 2019). In addition to spike protein, two smaller proteins, E and M might also play important role in the viral assembly of a coronavirus, and can boost the immune response against SARS-CoV (Shi et al., 2006; Schoeman and Fielding, 2019; Shang et al., 2020).

Due to the practices of open science and open data sharing approaches, the literature generating through research on COVID-19 is simply explosive. This report aims to review our current understanding on origin, genomic evolution, clinical and molecular diagnosis as well as prevention and control of the SARS-CoV-2 infection. Furthermore, this review provides meaningful information for future research and promote responses of the relevant national and international authority to tackle this pandemic disease.

Genomic composition of the SARS-CoV-2

The SARS-CoV-2 is the etiologic agent of the COVID-19 pandemic (Zhou et al., 2020b). The positively-sensed single-stranded RNA SARS-CoV-2 virus (Ahmed et al., 2020) has a genome size of 26,000 and 32,000 bases (Abdelmageed et al., 2020). The SARS-CoV-2 genome shares more

 doi:10.20944/preprints202004.0359.v1

than 80% sequence identity to the previous human coronavirus (Wu et al., 2020a). The organization and composition of novel coronavirus, SARS-CoV-2 genome comprising the structural proteins has been depicted in Figure 1. The spike (S) glycoprotein, membrane protein (M), small envelope protein (E), and nucleocapsid protein (N) are four major structural proteins of the CoVs (Figure 1A) (Ahmed et al., 2020). Transmission electron microscopic analysis revealed the typical spherical viral particles in a cell of SARS-CoV-2 with size ranging from 60-140 nm (Figure 1B). The genome of SARS-CoV-2 encodes for several smaller open reading frames (ORFs) located in both in 5′-UTR and 3′-UTR regions of the genome (Figure 2) that are assumed to express eight new proteins termed as accessory proteins (Rota et al., 2003; Freundt et al., 2010). The 5′-UTR and 3′-UTR of the CoVs play vital role in intra- and intermolecular interactions. They are functionally significant for RNA–RNA interactions, and for binding of viral and cellular proteins (Yang and Leibowitz, 2015). The first ORF at the 5´ end is Pb1ab which encodes for several non-structural proteins with size of 29844bp (7096aa), 29751bp (7073aa) and 30119bp (7078) in COVID-19, SARS-CoV; and MERS-CoV, respectively. The comparison of the spike protein at 3′ end of the genome of these three betacoronaviruses, Mousavizadeh and Ghasemi (2020) reported differences at position of 1273aa, 21493aa, and 1270aa in COVID-19, SARS-CoV, and MERS-CoV, respectively. Genetically, COVID-19 was less similar to SARS-CoV (about 79%) and MERS-CoV (about 50%) (Mousavizadeh and Ghasemi, 2020). The genomic position of envelope protein (E), membrane protein (M), and nucleocapsid protein (N) among betacoronaviruses are different as depicted in Figure 2. The accessory proteins are labelled as ORFs 1a and1b (polyprotein), 3a, 3b, 6, 7a, 7b, 8a, 8b, 9b and 10 (Figure 2). The size of these ORFs range from 39 to 274 amino acids (aa) (Marra et al., 2003; Freundt et al., 2010). These ORFs also encode for the replicase polyprotein, structural proteins, and other non-structural proteins (nsp) (Ahmed et al., 2020; Walls et al., 2020; Phan, 2020). The *orf1ab* is the largest gene in SARS-CoV-2 which encodes the polyprotein (pp1ab)

 doi:10.20944/preprints202004.0359.v1

and 15 nsps. The *orf1a* gene encodes for pp1a protein which also contains 10 nsps (Shereen et al., 2020). Distinguished differences between SARS-CoV and SARS-CoV-2 genomes particularly in absence of 8a protein and fluctuation in the number of amino acids in 8b and 3c protein in SARS-CoV-2 have been reported in several studies (Shereen et al., 2020; Wu et al., 2020a). The CoVs use their S glycoprotein, a main target for antibody neutralization, to bind their receptor, and mediate membrane fusion, and virus entry. Each monomer of homotrimeric S protein is about 180 kDa, which contains S1 and S2 subunits for mediating attachment and membrane fusion, respectively. The N- and C- terminal portions of S1 comprises two major domains S1 fold as two independent domains, the RBD and N-terminal domain (NTD) (Song et al., 2018; Ou et al., 2020; Rahman et al., 2020). While RBD of mouse hepatitis virus (MHV) is located at the NTD (Kubo et al., 1994), most of other CoVs, including SARS-CoV and MERS-CoV use C-domain to bind their receptors (Li et al., 2005; Lu et al., 2013; Ou et al., 2020). Moreover, the structural and biochemical comparison of the alpha- and beta-coronaviruses identify two notable genomic features of SARS-CoV-2 (Andersen et al., 2020; Wan et al., 2020; Walls et al., 2020; Wrapp et al., 2020; Zhou et al., 2020b). During the pathogenesis, the trimeric S protein is cleaved into S1 and S2 subunits, and the RBD of the S1 subunit directly binds to the peptidase domain (PD) of angiotensin converting enzyme 2 (ACE2) while the S2 carried out membrane fusion activity (Yan et al., 2020). Due to these novel genomic features (i) SARS-CoV-2 arises to be optimized for binding to the human ACE2 receptor; and (ii) the S protein of SARS-CoV-2 possesses a functional polybasic (furin) cleavage site at the S1–S2 boundary by way of the insertion of 12 nucleotides (Walls et al., 2020), which additionally led to the assumed acquisition of 3-*O*-linked glycans around the site.

The complete genomes of the novel SARS-CoV-2 sequenced from different patients share more than 99.9% sequence identity (Tang et al., 2020) suggesting a very recent host shift of this virus to humans (Lu et al., 2020; Tang et al., 2020; Zhou et al., 2020b). The trimeric spike (S)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2020 doi:10.20944/preprints202004.0359.v1

glycoprotein of SARS-CoV-2 and SARS-CoV are phylogenetically closely related. They showed about 77% amino-acid sequence identity (Rahman et al., 2020, Yuan et al., 2020; Zhou et al., 2020b). The genome of SARS-CoV-2 has 96.2%, 79.0% and 50.0% identity to the genomes of bat SARS-related coronavirus, SARSr-CoV; RaTG13, SARS-CoV and MERS-CoV, respectively (Tang et al., 2020; Lu et al., 2020; Zhou et al., 2020a) indicating a novel beta-coronavirus, which is likely to have originated from a bat coronavirus (Andersen et al., 2020), however, intermediate between bat and human is not clear.

Genome evolution of the SARS-CoV-2

Molecular analyses of the complete and/or nearly complete genome sequenced data of the known coronavirus from the patients from various geographical regions, researchers firmly determined that the SARS-CoV-2 originated through a natural process of evolution (Figure 3), but not a laboratory construct or a intentionally manipulated virus (Andersen et al., 2020; Zhou et al., 2020a). Around ten thousand complete and near-complete genomes of the SARS-CoV-2 have already been deposited to the global database repositories including NCBI (the National Center for Biotechnology Information), GSAID (global initiative on sharing all influenza data), and China National Center for Bioinformation 2019 Novel Coronavirus Resource (2019nCoVR) from the entire world. The genomic analyses of these sequences showed that some are genetically identical to each other, while others carry distinctive mutations (Islam et al., 2020; Phan, 2020). Analyzing 200 complete genomes of SARS-CoV-2 retrieved from the GISAID (https://www.gisaid.org/), we found that the evolution of this virus is not country or territory specific rather patient or ethnic group specific (Figure 3). The ongoing pandemic outbreak of the COVID-19 indicates its alarmingly rapid transmission across the globe. The researchers are racing round the clock to determine the exact origin of this fatal virus. Determining the origin and evolution of the SARS-CoV-2 is important for the surveillance, development of effective interventions for controlling the

242 epidemic, and prevention of the COVID-19. Analyses of the novel SARS-CoV-2 genome and
243 functional structures are needed to better understand its molecular cross-talks with human host
244 (Rahman et al., 2020; Zhang et al., 2020). Regular publication of pathogenic SARS-CoV-2 isolates
245 in open science and open data sharing model, reexamination of their origin and diversification
246 patterns are becoming possible. From the initial study on Wuhan COVID-19 outbreak to its rapid
247 spread to more than 213 countries or territories in the world, researchers suggested that this novel
248 virus is likely to have moved to human from bats through a host jump (Zhou et al., 2020a; Wu et
249 al., 2020b; Li et al., 2018; Sun et al., 2020). Despite having 77.38% and 31.93% sequence identity
250 with the S proteins of the SARS-CoV and MERS-CoV, respectively (Rahman et al., 2020), the
251 SARS-CoV-2 exhibited rich genetic diversity and frequent recombination events that might have
252 increased the potential for its cross-species transmission (Islam et al., 2020; Song et al., 2005; Sun
253 et al., 2020; Zhou et al., 2020b). In a recent study, Ou et al. (2020) reported that S protein of SARS-
254 CoV-2 shares 76% and 97% of amino acid identities with SARS-CoV and RaTG13, respectively.
255 The amino acid sequence of the RBD segment of the SARS-CoV-2 genome 74% and 90.1%
256 homologous to that of SARS-CoV and RaTG13, respectively (Ou et al., 2020).

257 Phylogenomic analysis of the recently released complete or nearly complete genomes of
258 SARS-CoV-2 to the global initiative on sharing all influenza data (GISAID)
259 (https://www.gisaid.org/) revealed that the novel coronavirus is most closely related to two severe
260 acute respiratory syndromes (SARS)-like CoV sequences that were isolated from bats during 2015
261 to 2017 (Zhang et al., 2020). This relatedness suggests that the bats' CoV and the human SARS-
262 CoV-2 shares a common ancestor. Therefore, the SARS-CoV-2 can be considered as a SARS-like
263 virus (Zhang et al., 2020). On the basis of recombination and evolutionary phylogenetic analyses,
264 Sun et al. (2020) reported that the SARS-CoV-2 shares a most recent common ancestor with
265 BetaCoV/RaTG13/2013 (EPI_ISL_402131) due to their clustering in the same position. However,

 doi:10.20944/preprints202004.0359.v1

this clustering might be associated with the convergent evolution or complex recombination events of the viral species with different evolutionary histories (Sun et al., 2020). Nonetheless, several recent reports demonstrated that the receptor-binding motif (RBM) of the beta-coronavirus genomes shares a very low sequence similarity indicating the possible association of alternative source for the RBM encoding sequence in SARS-CoV-2 (Lam et al., 2020; Sun et al., 2020; Xiao et al., 2020; Wong et al., 2020). On the other hand, Lam et al. demonstrated that the multiple putative lineages of pangolin CoV sequences shared 85.5% to 92.4% similarity to SARS-CoV-2 (Lam et al., 2020). Based on these similarities, they assumed that pangolins served as a potential intermediate host (Lam et al., 2020; Sun et al., 2020). In a phylogenetic network analysis of 160 complete human SARS-CoV-2 genomes, Forster et al. (2020) reported three central variants (A, B, and C) distinguished by aa changes, which we have named A, B, and C, with A being the ancestral type according to the bat outgroup coronavirus. The A and C types belonged to the Europeans and Americans while the B type is the most common type in East Asia (Forster et al., 2020). The genetic data indisputably revealed that SARS-CoV-2 is not derived from any previously published virus backbone (Andersen et al., 2020). Alternatively, two circumstances can plausibly explain the origin of SARS-CoV-2: (i) natural selection in humans following zoonotic transfer; and (ii) natural selection in an animal host before zoonotic transfer. However, currently available data are not sufficient enough to precisely conclude whether the virus was directly transmitted from bats to humans or indirectly through an intermediate host. It is now hard to rule out the convergent evolution as an alternative hypothesis to recombination to explain the discordant phylogenetic analysis. Inevitably, we need more sequence data to confirm the specific genetic identity and the origin of the SARS-CoV-2, which can be achieved by improved collection and monitoring of human samples across the globe, bat and other wild animal samples as well.

 doi:10.20944/preprints202004.0359.v1

Genome-wide mutations in the SARS-CoV-2 strains

The ongoing rapid human to human transmission, and global spread of SARS-CoV-2 have raised some intriguing questions, such as whether the evolution and host adaptation of this virus are driven by mutations. In spite of having interest on several genes of the SARS-CoV-2 genome, the evolutionary biologists tend to focus on a functionally conserved but sequentially diverged genes. This is because the role for the critical gene is conserved in various species, while it is rapidly evolving, indicating that a diverging force acts on the gene, e.g., by co-evolution such as symbiosis, evolutionary arms race or others (Adachi et al., 2020). Conversely, phylogenomic analysis of three super-clades (S, V, and G) isolated from the outbreaks of distinct geographic locations (China, USA and Europe) could not clearly reflect the hypothetical ongoing adaptation of SARS-CoV-2, which alternately refer to mere genetic drift and founder effects due to rapid spreading of the virus (Chiara et al., 2020). Nonetheless, several reports predicted the possible effects of genomics mutations, amino acid (aa) variations, and structural heterogeneity in the entire genomes of different strains of SARS-CoV-2 (Table 1) (Andersen et al., 2020; Huang et al., 2020; Islam et al., 2020; Lu et al., 2020; Phan, 2020; Yin, 2020; Walls et al., 2020). Remarkably, genome-wide annotations of 2,492 complete and/or nearly complete genome sequences of the SARS-CoV-2 strains from the infected individuals (retrieving from the GISAID database on April 15, 2020) revealed 1,407 nucleotide mutations and 722 amino acid (aa) replacements at different positions of SARS-CoV-2 genomes (Table 1) (Islam et al., 2020). In another study, annotating 558 complete genomes of SARS-CoV-2 strains, Yin (2020) reported 935 aa replacements in the polyprotein (including non-structural proteins and leader sequence), and 183, 33 and 222 aa substitutions in the S, M and N proteins (Table 1) of the SARS-CoV-2 genomes, respectively. Moreover, genetic analyses of eighty-six complete or nearly complete genomes of SARS-CoV-2 revealed many mutations and deletions on coding and non-coding regions of the RNA virus (Phan, 2020). In a

 doi:10.20944/preprints202004.0359.v1

recent study, Rahman et al. reported that spike (S) glycoprotein of the SARS-CoV-2 is a homotrimer of three homologous chains, chain A, B, and C (Rahman et al., 2020), and of these chains, the chain A and C have a high degree of structural divergence in the RBD and NTD segment compared to that of chain B. Therefore, the aa replacements, mutations and/or recombination in the RBD of S protein are deemed to be associated with the host adaption and/or evolution, progression of the infection, and across species infection (Qu et al., 2020). Islam et al. reported 12 amino acid substitutions in the RBD at 331 to 524 residues of S1 subunit in different SARS-COV-2 strains of Wales, USA, Shenzhen, Hong Kong, Shanghai, Guangdong, Finland, and France (Islam et al., 2020), while Sarkar et al. identified a unique mutation in the S glycoprotein (A930V) in the Indian SARS-CoV-2 strain (Sardar et al., 2020) which was absent in other related SARS-CoV-2 strains from different geographical regions. In addition to site-specific mutations in the spike protein, several deletions in the ranged nucleotides were also reported in the polyprotein, ORF10 and 3´-UTR of the genome of SARS-CoV-2 strains reported from Japan, USA, England, Canada, Netherlands, Wuhan and Australia (Islam et al., 2020). In another study, Liu et al. reported two common deletions in the genome of SARS-CoV with a very low frequency at 23585–23599 (aa-QTQTN) positioned at the upstream of the polybasic cleavage site of S1-S2, and 23596–23617 (aa-NSPRRAR) including the polybasic cleavage site in the clinical samples and cell-isolated virus strain (Liu et al., 2004). These nucleotide deletions can influence potentially the tertiary structures and functions of the polyprotein, S, M and E proteins which may play important role in virus-host interactions for infections, pathogenesis as well as immune-modulations (Islam et al., 2020; Phan, 2020; Xu et al., 2020; Zhou et al., 2020b). However, until now, most of the studies conducted using limited number of representative complete genomes of the SARS-CoV-2 virus isolated from patients belonging to only a small number of countries. The single N501T mutation in SARS-CoV-2's spike protein may have significantly enhanced its binding affinity for ACE2 (Shereen et al.,

2020). These studies have targeting clade/group based consensus sequence, genome-wide annotations of nucleotide deletions, mutations and aa substitutions, which are continuously making the viral proteins heterogeneous, could be one of the promising tools for monitoring and tracking the epidemic of evolutionarily divergent SARS-CoV-2 strains in their gradual and local genetic variations (Islam et al., 2020).

Diagnostic tools for the COVID-19

Early detection of the pathogen in patients of a pandemic disease is the main target area of research. Initially, computed tomography (CT) imaging, whole genome sequencing, and electron microscopy were used to detect the novel coronavirus in Wuhan of Hubei Province in China in December 2019. The clinical symptoms expressed by SARS-CoV-2 patients are non-specific, and thus, cannot be used for an accurate diagnosis. Only molecular techniques are able to specifically detect specific pathogen in a convenient way. The molecular diagnostics approaches are potentially being applied in the fields of control and prevention of human and animal diseases, food safety detection, and environmental monitoring due to their high sensitivity and excellent specificity (Chertow, 2018; Li et al., 2019). A rapid, specific and convenient diagnostic protocol might play a vital role in the containment of COVID-19, helping the rapid implementation of management of the disease that limit the spread through case identification, isolation, and contact tracing. In case of the COVID-19, the complete genome sequence data of the virus was publicly available within weeks of the first outbreak in Wuhan. It helped researcher to target specific genes for the development of nucleic acid test within three weeks. Analysis of the genome of SARS-CoV-2 revealed that of the structural genes, the *S* is highly divergent with less than 78% nucleotide sequence similarity with previously described coronaviruses (Rahman et al., 2020; Wrapp et al., 2020; Zhou et al., 2020b). However, other three proteins are more conserved than the spike protein.

 doi:10.20944/preprints202004.0359.v1

Due to non-specific symptoms and gender discrimination independent of age (Jin et al., 2020), it is hard to make a conclusion/diagnose a suspected COVID-19 patient. Many patients suffer fever in early stage of the infection but a large proportion of the COVID patient had fever at the later stage of the disease. The COVID-19 patients generally suffer from cough, fatigue, sputum production and shortness of breath (Guan et al., 2020). A highly sensitive, convenient and specific diagnostic tool is essential for detecting the infected person which is one of the most important starting steps for addressing COVID-19. Due to several limitations of the requirement for culturing virus and unavailable/unvalidated serum antibody and antigen detection tests, primarily the reverse transcription-polymerase chain reaction (RT-PCR) was considered as the most useful and reliable laboratory diagnostic method for the COVID-19 worldwide. The availability of complete genome sequences of the SARS-CoV-2 early in the epidemic helped to develop specific primers and effective diagnostic laboratory protocols for COVID-19 (Chan et al., 2020b). The first real-time RT-PCR assays targeting 3 genes, nucleocapsid (*N*), envelop (*E*) and RNA-dependent RNA polymerase (*RdRp*) were developed and published on 23 January 2020 by Corman et al. (2020). The *RdRp* gene of the SARS-CoV-2 genome is highly similar to that gene of bat coronavirus RaTG13, and overall similarity between these two genome sequences is 96% (Zhou et al., 2020a). Later consistent detection of COVID-19 in saliva was published by To et al. (2020). Then several groups and countries developed many diagnostic protocols targeting or using nucleic acid tests or protein/antibody, loop-mediated amplified technique, imaging techniques (CT-scan) or CRISPR-Cas mediated technology (Table 2).

The RT-PCR method to diagnose the COVID-19 has been developed by several groups targeting different genes. For example, Chan et al. (2020) developed three methods of RT-PCR. Among the three assays, the COVID-19-*RdRp*/Hel assay had the lowest limit of detection in vitro (1.8 TCID50/ml with genomic RNA and 11.2 RNA copies/reaction with in vitro RNA transcripts).

This method was validated in testing 273 suspected patients where 15 patients were confirmed as SARS-CoV-2 positive. This method targeted the RNA-dependent RNA polymerase (*RdRp*)/helicase (Hel), *S*, and *N* genes of SARS-CoV-2 with that of the reported RdRp-P2 assay which is used in >30 European laboratories. The authors confirmed that the COVID-19-RdRp/Hel assay did not cross-react with other human pathogenic coronaviruses and respiratory pathogens in cell culture and clinical specimens (Chan et al., 2020). Therefore, this method is expected to help improve the laboratory diagnosis of COVID-19.

The non-invasive radiographic technique, CT-scan, and RT-PCR are two efficient methods for the diagnosis of this pandemic disease. Although CT-scan is more sensitive than RT-PCR, the latter technique has been widely used worldwide for the detection of COVID-19. However, considering less sensitivity of PCR with low rate of false negative, and longer time required, CT-scan imaging features were considered as clinical diagnostic criteria in Hubei Province (PCR NHcot. The Novel Coronavirus Pneumonia Diagnosis and Treatment Plants, 5th Trial version, 2020). In fact, radiographical features of SARS-CoV-2 are similar to those in community-acquired pneumonia caused by other organisms. Typical imaging features of COVID-19 patients observed in CT scan include, predominant ground-glass opacity, consolidations, smooth or irregular interlobular septal thickening, air bronchiogram and thickening of the adjacent pleura with predominant peripheral and lower lobe involvement (Shi et al., 2020). Through analysis of the data of 1014 patients in China, CT scan was found to be sensitive than RT-PCR for diagnosis of COVID-19 (Ai et al., 2020). The authors observed positive rates of RT-PCR assay and chest CT imaging in their cohort 59% (601/1014), and 88% (888/1014) for the diagnosis of the suspected patients with COVID-19, respectively. The sensitivity of chest CT imaging for COVID-19 was 97%, where RT-PCR was used as a standard reference (Ai et al., 2020) However, for worldwide

application, both techniques are expensive and required skilled manpower and high level of technical facilities.

Developing plug-and-play diagnostics to manage the SARS-CoV-2 outbreak would also be useful in preventing future epidemics. Current real-time RT-PCR techniques are not convenient, rapid, robust and also not 100% accurate (high rate of false negative). Furthermore, these methods require highly skilled personnel, sophisticated facilities and equipment that are not available in enough quantities in the developing countries. A recently developed Abbott ID Now™ COVID-19 test has been found to be very convenient, and can detect SARS-CoV-2 in 5 min only. It is a phenomenal progress but availability of this convenient method is still rather limited. Similarly, an immune assay targeting IgM/IgG antibodies in human finger pricks or venous whole blood, serum or plasma samples can detect COVID-19 within 10 min. Another rapid and simple point-of-care lateral flow immunoassay which can detect IgM and IgG antibodies simultaneously against SARS-CoV-2 virus in human blood within 15 min. It is efficient in detecting the virus at different infection stages of the patient. A research group of Peking University developed a new method for rapid construction of transcriptome sequencing library of Sequencing HEteRo RNA-DNA-hYbrid (SHERRY), which is helpful for rapid sequencing of SARS-CoV-2 (Di et al., 2020). They showed that Tn5 transposase, which randomly binds and cuts double-stranded DNA, can directly fragment and prime the RNA/DNA heteroduplexes generated by reverse transcription. The primed fragments are then subject to PCR amplification. This provides an approach for simple and accurate RNA characterization and quantification.

The recent outbreak of the novel coronavirus COVID-19 can be diagnosed using qPCR, but inadequate access to reagents and equipment has slowed disease detection. To rapidly diagnose the disease, Zhang group of MIT developed a test paper for rapid detection of SARS-CoV-2 in one hour by using SHERLOCK (**S**pecific **H**igh Sensitivity **E**nzyme **R**eporter Un**LOCK**ing)

technology. This technology may be used widely after clinical trials (Zhang et al., 2020b). The one hour long SHERLOCK COVID-19 diagnosis method has three major steps: (i) 25 min incubation – isothermal amplification of the extracted nucleic acid sample using a commercially available recombinase polymerase amplification (RPA) kit; (ii) 30 min incubation – detection of pre-amplified viral RNA sequence using Cas13; (iii) 2 min incubation – visual read out of the detection result by eye using a commercially-available paper dipstick. This technique used synthetic COVID-19 *S* and *ORF1ab* genes for the diagnosis and no clinical specimen has yet been tested.

In the process of the development of new technique, an exciting improvement is the DZ-Lite SARS-CoV-2 CLIA IgM and IgG tests established by Diazyme, USA. This technique has received FDA EUA approval (https://bit.ly/2UXlils). The molecular principle of this test is a chemiluminescence immunoassay (CLIA) that run on an automated Diazyme DZ-Lite 3000 Plus chemiluminescence analyzer with a throughput of 50 tests/h. Similarly, Snibe, China, has developed automated CLIA tests on MAGLUMI CLIA analyzers for the detection of IgG and IgM in the patient sample in 30 min (https://bit.ly/2JXGMZm). The major advantages of automated CLIA analyzers based COVID-19 assays compared to rapid LFIA tests is the very high throughput of samples that can be analyzed and the ability to perform more clinical tests for other biomarkers, such as C-reactive protein (CRP), which also need to be monitored in COVID-19 suspects. The rapid, convenient, low cost and specific serological and automated tests are urgently needed to be distributed worldwide especially in the developing countries for testing higher number of patients to tackle this highly contagious disease.

Antivirals for the pandemic COVID-19 virus: vaccines and therapeutics

Though each and every country and region with or without having COVID-19 outbreak is following strict safeguard or prevention, and control strategies which have been reported either at

 doi:10.20944/preprints202004.0359.v1

national-level (case-related population level and general population level) and international-level (CDC, Centers for Disease Control and Prevention; WHO, world health organization; FDA, Food and Drug Administration), the SARS-CoV-2 outbreak has already spread across the globe covering more than 213 countries or territories. Despite several public health measures such as case isolation, identification and follow-up of contacts, environmental disinfection, social distance, and use of personal protective equipment have been introduced (Wei and Ren, 2020), in the absence of any antivirals (Kalita et al., 2020; Rahman et al., 2020; Wang et al., 2020b), the disease is spreading at an alarming rate. The outbreak of novel COVID-19 disease causes severe pneumonia-like respiratory infections in the unavailability of any recommended antiviral or immunomodulatory therapy (Kalita et al., 2020; Rahman et al., 2020; Wang et al., 2020b). The new cases of active acute infections are being added to the open COVID-19 database every day (Figure 4), as the case count globally skyrockets. The genome of SARS-CoV-2 is being continuously sequenced, deposited to the repository databases, and large-scale culture of SARS-CoV-2 are also being regularly performed throughout the world which ultimately paved the way for the development of diagnostic tests, and research into vaccine candidate design and therapeutics (Rahman et al., 2020; Islam et al., 2020, Kalita et al., 2020; Zhou et al., 2020b). Therefore, researchers from across the globe are desperately working round the clock to find ways to slow the spread of the novel coronavirus and to find an effective treatment to control this fatal viral disease. Though, more than 200 clinical trials of COVID-19 treatments or vaccines that are either ongoing or recruiting patients (Zhou et al., 2020b), till now no recommended therapeutic drug or vaccines are available for the treatment of COVID–19. The treatment is mainly supportive and through a few repurposed drugs (Rahman et al., 2020; Rehman et al., 2020).

At the outset of the epidemic in Wuhan, China, COVID-19 confirmed patients were treated with interferons-α nebulization, broad-spectrum antibiotics, and few antiviral drugs to reduce the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2020 doi:10.20944/preprints202004.0359.v1

481 viral load (Shereen et al., 2020; Wang et al., 2020b), however, only remdesivir (GS-5734) has
482 shown promising impact against the virus (Wang et al., 2020b). Since then, various other antiviral
483 drugs including nafamostat, nitazoxanide, ritonavir, aak1, baricitinib, arbidol, ribavirin,
484 penciclovir, chloroquine, favipiravir (T-750) or avigan, hydroxychloroquine and chloroquine
485 EIDD-2801 are being tested in clinical trials (Martinez, 2020; Liu et al., 2020; Wang et al., 2020b),
486 and some of these drugs exhibited moderate results when tested against infection in patients and *in-*
487 *vitro* clinical isolates (Shereen et al., 2020). In an open-label non-randomized clinical trial, Gautret
488 et al. reported that a combined therapy of hydroxychloroquine and azithromycin reduced the
489 detection of viral RNA compared to control (Gautret et al., 2020). In another randomized,
490 controlled and open-label hospital-based trial with confirmed cases of SARS-CoV-2 infection,
491 which causes the respiratory illness COVID-19, Cao et al. reported that two protease inhibitors,
492 lopinavir and ritonavir failed completely to cure this disease (Cao et al., 2020). The orally
493 bioavailable ribonucleoside analog, β-D-N4-hydroxycytidine (NHC, EIDD-1931), is a broad-
494 spectrum antiviral drug against various unrelated RNA viruses including influenza, Ebola, CoV,
495 and Venezuelan equine encephalitis virus (VEEV) (Reynard et al., 2015; Agostini et al., 2019;
496 Toots et al., 2019). This proven NHC/EIDD-2801 against multiple coronaviruses showed potential
497 antiviral activity against SARS-CoV-2, and recommended for future zoonotic coronaviruses
498 (Sheahan et al., 2020).

499 Immunoprophylaxis through passive transfer of antibodies is regarded as an effective method
500 for clinical treatment of infectious diseases. For example, the use of versatile class of
501 pharmaceuticals commonly known as monoclonal antibodies (mAbs) is a new era in infectious
502 disease prevention. This passive immunization overcomes many drawbacks associated with serum
503 therapy and intravenous immunoglobulins preparations in terms of specificity, purity, low risk of
504 blood-borne pathogen contamination and safety (ter Meulen., 2006; Shanmugaraj et al., 2020).

Several earlier studies reported the successful generation of neutralizing antibodies in mice against SARS-CoV through experimental vaccination or passive transfer of mAb and subsequent reduction of viral replication (Traggiai et al., 2004; Sui et al., 2005; ter Meulen., 2006). Thus, mAbs with potent neutralizing activity against SARS-CoV-2 infections could become promising candidates for both prophylactic and therapeutic interventions (Shanmugaraj et al., 2020; Zhou et al., 2020b). The RBD segment of the S glycoprotein is one of the prime important immunogenic sites carrying both T-cell and B-cell epitopes (Rahman et al., 2020), and antibody binding to this surface is likely to block viral entry into cells (Qu et al., 2020). Antibody generation targeting the RBD and/or NTD of the S, M, and E proteins could be an effective preventive measure against of SARS-CoV-2 infections (Rahman et al., 2020).

Vaccines are the most effective and economical means to prevent and control the infectious viral diseases (Zhang et al., 2020). However, yet to date, no vaccine is available against COVID-19, while previous vaccines or strategies used to develop a vaccine against SARS-CoV might be effective. The development of an effective vaccine against SARS-CoV-2 infection is urgently required. Whole-cell killed or live-attenuated vaccines present multiple antigenic components to the host and can thus potentially induce diverse immunologic effectors against pathogen (Zhang et al., 2020a). Traditionally mature preparatory technology such as inactivated and/or live-attenuated vaccine could become the first SARS-CoV-2 vaccine put into clinical applications (Regla-Nava et al., 2015; Shang et al., 2020). However, these inactivated and attenuated virus vaccines have a wide range of disadvantages and side effects including inappropriate for highly immunosuppressed individuals (Shang et al., 2020), phenotypic or genotypic reversion is possible and can still cause some disease (Regla-Nava et al., 2015). Alternatively, putative protective antigen/peptides vaccine candidate for SARS-CoV-2 should be considered on the basis immunogenicity. Unlike inactivated or live-attenuated and viral vectored vaccines, subunit vaccine contains specific viral antigenic

fragments without including any components of infectious viruses; hence such vaccines are free of any potential harmful immune responses, making them promising vaccine candidates for SARS-CoV-2 (Wang et al., 2020c). The conserved epitopes of the SARS-CoV-2 genomes may allow structure-based vaccine design not only for this novel coronavirus, but also for cross-protective antibody responses against future coronavirus epidemics and pandemics (Rahman et al., 2020; Yuan et al., 2020). Moreover, subunit vaccines may be target specific, well-defined neutralizing epitopes with improved immunogenicity and/or efficacy (Zhang et al., 2020a; Wang et al., 2020c).

With the advancement in immunoinformatics and computational biology, it is now possible to accelerate the drug discovery pipeline and vaccine development (Rahman et al., 2020; Zhang et al., 2020a), and these methods have surpassed the conventional methods. Thus, few vaccines are in the pipeline against SARS-CoV-2, and of them, the mRNA-based vaccine prepared by the US National Institute of Allergy and Infectious Diseases against SARS-CoV-2 is under phase 1 trial (McKay, 2020). Another vaccine based on INO-4800-DNA will be soon available for human testing (Inovio IP, 2020). The Centre for Disease Control and Prevention (CDC), China is working to develop an inactivated virus vaccine (Cheung, 2020). An mRNA-based vaccine's sample prepared by Stermirna Therapeutics will be available soon (Xinhua, 2020). The GeoVax and BravoVax (Wuhan, China) is working to develop a Modified Vaccina Ankara (MVA) based vaccine (Geo-Vax, 2020). In addition, the Clover Biopharmaceuticals is trying to develop a recombinant 2019-nCoV S protein subunit-trimer based vaccine (Clover B, 2020).

Numerous studies have been published related to the B and T-cell epitope-based vaccine development using *in silico* immunoinformatics methods (Feng et al., 2020; Rahman et al., 2020; Wang et al., 2020c; Zhang et al., 2020a). Being an RNA virus, genome-wide nucleotide mutations and amino-acid mutations and/or substitutions (Table 1) have already been reported in different

SARS-CoV-2 strains from across the globe (Huang et al., 2020; Islam et al., 2020; Phan, 2020; Yin, 2020; Wang et al., 2020d). Therefore, it is critical to develop vaccines with strong efficacy and safety targeting this COVID-19 virus to prevent its infection in humans. The structural divergence in the RBD and NTD segments of the S protein in SARS-CoV-2 is main focus of epitope-based chimeric vaccine candidate designing, selection, and development (Rahman et al., 2020). Therefore, instead of single epitope, multi-epitope based vaccines targeting the full-length S protein and its structural domains (RBD, NTD, S1 and S2 subunits), M, E and N proteins can play a great role in fighting against this SARS-Cov-2 virus (Rahman et al., 2020; Zhang et al., 2020a). Furthermore, multi-epitope-based recombinant vaccines have already been developed for several emerging diseases like MERS and SARS. The potential efficacies of these chimeric vaccines have been reported in MERS-CoV and SARS-CoV (Almofti et al., 2018; ul Qamar et al., 2019; Yong et al., 2019).

Even as the hunt for a vaccine to treat Covid-19 continues, a classic adaptive immunotherapy known as convalescent plasma (CP) therapy that was successfully applied over the past two decades in the treatment of SARS, MERS, and 2009 H1N1 pandemic with satisfactory efficacy and safety (Cheng et al., 2005; Hung et al., 2009; Ko et al., 2018) holds good promise. In a recent pilot study, Duan et al. reported that CP therapy was found to be well tolerated and could potentially improve the clinical outcomes through neutralizing viremia in severe COVID-19 cases (Duan et al., 2020). One dose of CP with a high concentration of neutralizing antibodies can rapidly reduce the viral load, and tends to improve clinical outcomes. However, the optimal dose and treatment time point, as well as the definite clinical benefits of CP therapy should be further investigated in randomized clinical studies. Although many news reports described the discovery of vaccines by several groups against SARS-CoV-2, it may take more than a year or so to complete several levels

of examinations of those promising candidates including human trials before releasing for the practical applications.

Conclusions and perspectives

Distinguished from MERS-CoV and SARS-CoV, the ongoing pandemic of 2019-nCoV disease is novel coronavirus which first emerged in the Wuhan city of China. This new virus was renamed as SARS-CoV-2 by the coronavirus study group, the ICTV. Genome sequencing is important for the taxonomy, designing primers and probe for the diagnostics of the pathogen. By January 10, 2020, complete whole genome sequence of the bronchoalveolar lavage (BAL) fluid revealed that the pathogen is a novel betacoronavirus B1 lineage (Zhou et al., 2020a). Later, complete or nearly complete genome sequences of a large number of strains of COVID-19 have been published and all the research data on this new virus are publicly available. The genomic features described in this review is based on the recent reports of the infectiousness and transmissibility of SARS-CoV-2 in humans. Currently, evidence shows that SARS-CoV-2 is not purposefully manipulated, however, controversy still remained to prove or disprove the other theories of the origin of this virus described in this article. Genome-wide annotations of a wider range of sequences (50-2500) revealed huge number of mutations throughout the SARS-CoV-2 genome, which includes both mismatch and deletion mutations both in translated and untranslated regions. Moreover, the identification of the conformational changes in mutated protein structures and untranslated cis-acting elements is of significance for studying the virulence, pathogenicity and transmissibility of SARS-CoV-2. Regardless of the available theories of origin of SARS-CoV- through natural selection, the ongoing surveillance of pneumonia in humans and other animals is clearly of utmost importance. Thus, combining the assessment of the epidemiologic origin, evolutionary genomic features with clinical and laboratory findings could facilitate early diagnosis

 doi:10.20944/preprints202004.0359.v1

of COVID-19 pneumonia like syndrome. The discovery of specific diagnostic tool targeting specific genes of the genome of COVID-19 within weeks of the outbreak of the disease in China was a phenomenal research success which has been playing vital role in tackling this highly contagious disease. Although real-time RT-PCR methods targeting specific genes have widely been used to diagnose the COVID-19 patients, however, recently developed more convenient, rapid, and specific diagnostic tools targeting IgM/IgG or newly developed plug and play methods should be available especially for the resource-poor developing countries. Therefore, identification of the etiological agent using prudent and cutting-edge diagnostic tools, as well as sequencing of the virus from very early cases, would similarly be highly informative. Several approaches for vaccines and antivirals targeting human coronaviruses are in developmental stages, which could be safely and effectively used against the current as well as future epidemics. Though controversial, convalescent plasma therapy has been reported to show limited cross-neutralization activity in COVID-19 patients. We can assume that potential targets for development of drugs and multiepitope-based chimeric peptide vaccines against this newly emerging lineage B beta-CoV, SARS-CoV-2 will be available soon. However, vaccine delivery modality and immunization strategy should be ensured through rapid human and animal-based trails before commercialization. We expect researchers and/or companies who are working round the clock will bring a new SARS-CoV-2-based vaccine from isolated virus particle or gene sequences to clinical testing within short time. This review summarizes new knowledge on genomics, genome evolution, developed diagnostic methods and progress in development of vaccine or therapeutic remedies of the COVID-19, and discusses on scopes for further research and effective management and surveillance of COVID-19.

AUTHORS CONTRIBUTIONS

TI: Involved in conceived the idea, drafted and edited the manuscript; MNH: Wrote manuscript, prepared Figures and Tables; AC, MAMA and MAH: Critically edited the manuscript.

CONFLICT OF INTEREST

The authors declared no conflict of interests

ETHICAL STATEMENT

This review article has no ethical issues

DATA AVAILABILITY STATEMENT

All data used in this manuscript are available in the manuscript as Figures and Tables.

References

Abdelmageed, M. I. et al. (2020). Design of multi epitope-based peptide vaccine against E protein of human 2019-nCoV: An immunoinformatics approach. *BioRxiv*.

Adachi, S., Koma, T., Doi, N., Nomaguchi, M., and Adachi, A. (2020). General Commentary: Origin and evolution of pathogenic coronaviruses. *Front. Immunol*. doi: 10.3389/fimmu.2020.00811.

Adhikari, S. (2020). Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review. *Infect. Dis. Poverty* 9, 29, https://doi.org/10.1186/s40249-020-00646-x.

Ahmed, S. F., Quadeer, A. A., and McKay, M. R. (2020). Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies. *Viruses* 12(3), 254.

650 Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W. et al. (2020). Correlation of chest CT
651 and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of
652 1014 cases. Radiology, http://doi.org/10.1148/radiol.2020200642.
653 Agostini, M. L., Pruijssers, A. J., Chappell, J. D., Gribble, J., Lu, X., Andres, E. L. et al. (2019).
654 Small-Molecule Antiviral β-d-N4-Hydroxycytidine Inhibits a Proofreading-Intact
655 Coronavirus with a High Genetic Barrier to Resistance. *J. Virol.* 93(24).
656 Almofti, Y. A., Abd-elrahman, K. A., Gassmallah, S. A. E., and Salih, M. A. (2018). Multi
657 Epitopes Vaccine Prediction against Severe Acute Respiratory Syndrome (SARS)
658 Coronavirus Using Immunoinformatics Approaches. *Am. J. Microbiol. Res.* 6(3), 94-
659 114.
660 Andersen, K.G. et al. (2020). The proximal origin of SARS-CoV-2. *Nat. Med.*
661 https://doi.org/10.1038/s41591-020-0820-9.
662 Anonymous (2020a). Division of Viral Diseases. CDC 2019-Novel Coronavirus (2019-nCoV)
663 Real-Time RT-PCR Diagnostic Panel; Division of Viral Diseases, U.S. Centers for
664 Disease Control and Prevention, 2020.
665 Anonymous (2020b). Specific Primers and Probes for Detection 2019 Novel Coronavirus;
666 China National Institute For Viral Disease Control and Prevention, 2020.
667 Anonymous (2020c). Detection of 2019 Novel Coronavirus (2019-nCoV) in Suspected
668 Human Cases by RT-PCR; School of Public Health, Hong Kong University.
669 Armijos-Jaramillo, V., Yeager, J., Muslin, C., and Perez-Castillo, Y. (2020). SARS-CoV-2,
670 an evolutionary perspective of interaction with human ACE2 reveals undiscovered
671 amino acids necessary for complex stability. *bioRxiv*,
672 https://doi.org/10.1101/2020.03.21.001933.
673 Badawi, M. M. et al. (2016). *In Silico* Prediction of a Novel Universal Multi-Epitope Peptide
674 Vaccine in the Whole Spike Glycoprotein of MERS CoV. *Am. J. Microbiol. Res.* 4(4),
675 101-21.
676 Bastola, A., Sah, R., Rodriguez-Morales, A. J., Lal, B. K., Jha, R., Ojha, H. C. et al. (2020).
677 The first 2019 novel coronavirus case in Nepal. *The Lancet Infect. Dis.* 20(3), 279-280.
678 Cao, B., Wang, Y., Wen, D., Liu, W., Wang, J., Fan, G. et al. (2020). A trial of lopinavir–
679 ritonavir in adults hospitalized with severe Covid-19. *N. Engl. J. Med.*
680 NEJMoa2001282.

Chan, J. F. W. et al. (2020a). Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/Hel real-time reverse transcription-polymerase chain reaction assay validated in vitro and with clinical specimens. *J. Clin. Microbiol.* doi:10.1128/JCM.00313-20.

Chan, J. F. W. et al. (2020b). Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg. Microbes Infect.* 9:221–36

Chen, N. et al. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 395:507–13.

Cheng, Y. et al. (2005). Use of convalescent plasma therapy in SARS patients in Hong Kong. *Eur. J. Clin. Microbiol. Infect. Dis.* 24, 44–46.

Chertow, D. S. (2018). Next-generation diagnostics with CRISPR. *Science* 360, 381-382.

Chiara, M., Horner, D. S. and Pesole, G. (2020). Comparative genomics suggests limited variability and similar evolutionary patterns between major clades of SARS-Cov-2. *bioRxiv*. https://doi.org/10.1101/2020.03.30.016790.

Clover, B. (2020). Clover initiates development of recombinant subunit- trimer vaccine for wuhan coronavirus (2019-ncov). [cited 2020 6 March]; Available from: http://www.cloverbiopharma.com/index.php?m=content&c=index&a=show&catid=11&id=40.

Corman, V. M. et al. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill.* 25(3): 2000045. doi: 10.2807/1560-7917.ES.2020.25.3.2000045.

Details - Public Health Image Library(PHIL) https://phil.cdc.gov/Details.aspx?pid=23354 (accessed Mar 27, 2020).

Di, L., Fu, Y., Sun, Y., Li, J., Liu, L., Yao, J. et al. (2020). RNA sequencing by direct tagmentation of RNA/DNA hybrids. *Proc. Natl. Acad. Sci.* USA, https://doi.org/10.1073/pnas.1919800117.

Duan, K., Liu, B., Li, C., Zhang, H., Yu, T., Qu, J. et al. (2020). Effectiveness of convalescent plasma therapy in severe COVID-19 patients. *Proc. Natl. Acad. Sci.* USA.

Feng, Y., Qiu, M., Zou, S., Li, Y., Luo, K., Chen, R. et al. (2020). Multi-epitope vaccine design using an immunoinformatics approach for 2019 novel coronavirus in China (SARS-CoV-2). *bioRxiv* doi: https://doi.org/10.1101/2020.03.03.962332.

Forster, P., Forster, L., Renfrew, C., & Forster, M. (2020). Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc. Natl. Acad. Sci.* USA https://doi.org/10.1073/pnas.2004999117.

Freundt, E. C., Yu, L., Goldsmith, C. S., Welsh, S., Cheng, A., Yount, B. et al. (2010). The open reading frame 3a protein of severe acute respiratory syndrome-associated coronavirus promotes membrane rearrangement and cell death. *J. Virol.* 84(2), 1097-1109.

Gao, G. F. (2018). From 'A'IV to 'Z'IKV: attacks from emerging and reemerging pathogens. *Cell* 172: 1157-1159.

Gautret, P. et al. (2020). Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. *Int. J. Antimicrob. Agents* 105949.

Geo-Vax, (2020). Geovax and bravovax (wuhan, china) to collaborate on development of coronavirus vaccine. [cited 2020 3 March]; Available from: https://www.geovax.com/news/geovax-and-bravovax-wuhan-china-to-collaborate-ondevelopment-of-coronavirus-vaccine.

Hassan, S., Sheikh, F. N., Jamal, S., Ezeh, J. K., and Akhtar, A. (2020). Coronavirus (COVID-19): a review of clinical features, diagnosis, and treatment. *Cureus* 12(3).

Hemida, M. G., and Abduallah, M. M. B. (2020). The SARS-CoV-2 outbreak from a one health perspective. *One Health*, 100127.

Hindson, J. (2020). COVID-19: faecal–oral transmission? *Nat. Reviews Gastroent. Hepatol.* 1-1.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S. et al. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. https://doi.org/10.1016/j.cell.2020.02.052.

738 Holshue, M. L., DeBolt, C., Lindquist, S., Lofy, K. H., Wiesman, J., Bruce, H. et al. (2020).
739 First case of 2019 novel coronavirus in the United States. *N. Engl. J. Med.*
740 NEJMoa2001191.
741 Huang, J. M., Jan, S. S., Wei, X., Wan, Y. et al. (2020). Evidence of the Recombinant Origin
742 and Ongoing Mutations in Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-
743 CoV-2). *bioRxiv*.
744 Hung, F. et al. (2009). Convalescent plasma treatment reduced mortality in patients with severe
745 pandemic influenza A (H1N1) 2009 virus infection. Clin. Infect. Dis. 52, 447–456.
746 Inovio, I. P. (2020). Inovio selected by cepi to develop vaccine against new coronavirus inovio.
747 [cited 2020 29 January]; Available from: <http://ir.inovio.com/news-and-
748 media/news/press-release-details/2020/Inovio-Selectedby-CEPI-to-Develop-Vaccine-
749 Against-newCoronavirus/default.aspx>.
750 Islam, M. T., Croll, D., Gladieux, P., Soanes, D. M., Persoons, A., Bhattacharjee, P., Hossain,
751 M. S., Gupta, D. R., Rahman, M. M. and Mahboob, M. G. (2016). Emergence of wheat
752 blast in Bangladesh was caused by a South American lineage of *Magnaporthe oryzae*.
753 BMC Biol. 14, 84. doi:10.1186/s12915-016-0309-7
754 Islam, M.R., Hoque, M.N., Rahman, M.S., Puspo, J.A., Akhter, M., Akter, S. et al. (2020).
755 Genome Wide Analysis of Severe Acute Respiratory Syndrome Coronavirus-2
756 Implicates World-Wide Circulatory Virus Strains Heterogeneity. Preprints 2020,
757 2020040137, doi: 10.20944/preprints202004.0137.v1.
758 Jackwood, M. W., Hall, D., and Handel, A. (2012). Molecular evolution and emergence of
759 avian gammacoronaviruses. *Infect. Genet. Evol.* 12(6), 1305-1311.
760 Jin, J. M. et al. (2020). Gender differences in patients with COVID-19: Focus on severity and
761 mortality. *Front. Public Health* doi: 10.3389/fpubh.2020.00152.
762 Kalita, P., Padhi, A. K., Zhang, K. Y., and Tripathi, T. (2020). Design of a Peptide-Based
763 Subunit Vaccine against Novel Coronavirus SARS-CoV-2. Preprints
764 (www.preprints.org), doi:10.20944/preprints202003.0433.v1.

765 Kellner, M. J., Koob, J. G., Gootenberg, J. S., Abudayyeh, O. O., and Zhang, F. (2019).
766 SHERLOCK: Nucleic Acid Detection with CRISPR Nucleases. *Nat. Protoc.* 14 (10),
767 2986–3012.
768 Khan, S., Siddique, R., Shereen, M. A., Ali, A., Liu, J., Bai, Q. et al. (2020). The emergence
769 of a novel coronavirus (SARS-CoV-2), their biology and therapeutic options. *J.*
770 *Clinical Microbiol.* doi: 10.1128/JCM.00187-20.
771 Ko H. et al. (2018). Challenges of convalescent plasma infusion therapy in Middle East
772 respiratory coronavirus infection: A single centre experience. *Antivir. Ther.* 23, 617–
773 622.
774 Lam, T. T. Y. et al. (2020) Identification of 2019-nCoV related coronaviruses in Malayan
775 pangolins in southern China. *bioRxiv* https://doi.org/10.1101/2020.02.13.945485.
776 Letko, M., Marzi, A., and Munster, V. (2020). Functional assessment of cell entry and receptor
777 usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* 5(4),
778 562-569.
779 Li, F., Li, W., Farzan, M., and Harrison, S. C. (2005). Structure of SARS coronavirus spike
780 receptor-binding domain complexed with receptor. *Science* 309, 1864-1868.
781 Li, G. et al. (2018). Origin, genetic diversity, and evolutionary dynamics of novel porcine
782 circovirus 3. *Adv. Sci.* (Weinh) 5, 1800275.
783 Li, Y., Li, S., Wang, J., and Liu, G. (2019). CRISPR/Cas systems towards next-generation
784 biosensing. *Trends Biotechnol.* 37, 730-743.
785 Li, Z. et al. (2020). Development and clinical application of a rapid IgM-IgG combined
786 antibody test for SARS-CoV-2 infection diagnosis. *J. Med. Virol.* 2020.
787 Lin, X., Gong, Z., Xiao, Z., Xiong, J., Fan, B., and Liu, J. (2020). Novel coronavirus
788 pneumonia outbreak in 2019: computed tomographic findings in two cases. *Korean J.*
789 *Radiol.* 21(3), 365-368.
790 Liu, J. et al. (2020). Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in
791 inhibiting SARSCoV-2 infection *in vitro*. *Cell Discov* 6,16, 10.1038/s41421-020-
792 0156-0.
793 Lu, G. et al. (2013). Molecular basis of binding between novel human coronavirus MERS-
794 CoV and its receptor CD26. *Nature* 500, 227–231.

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H. et al. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet* 395(10224), 565-574.

Marra, M. A. et al. (2003). The genome sequence of the SARS-associated coronavirus. *Science* 300(5624), 1399-1404.

Martinez, M. A. (2020). Compounds with therapeutic potential against novel respiratory 2019 coronavirus. *Antimicrob. Agents Chemother.* doi:10.1128/AAC.00399-20.

McKay, B. L. P. (2020). Drugmakers rush to develop vaccines against china virus. The Wall Street Journal. Available online: https://www. wsj. com/articles/drugmakers-rush-to-develop-vaccines-against-china-virus-11579813026 (accessed on 23 March 2020).

Mousavizadeh, L., & Ghasemi, S. (2020). Genotype and phenotype of COVID-19: Their roles in pathogenesis. *J. Microbiol. Immunol. Infect.* https://doi.org/10.1016/j.jmii.2020.03.022.

Nao, N., Shirato, K., Katano, H., Matsuyama S., and Takeda, M. (2020). Detection of Second Case of 2019-nCoV Infection in Japan; National Institute of Infectious Diseases, Department of Virology III, Japan.

Nishiura, H., Linton, N. M., and Akhmetzhanov, A. R. (2020). Initial cluster of novel coronavirus (2019-nCoV) infections in Wuhan, China is consistent with substantial human-to-human transmission. *J. Clin. Med.* 9(2), 488; https://doi.org/10.3390/jcm9020488.

Ou, J., Zhou, Z., Zhang, J., Lan, W., Zhao, S., Wu, J. et al. (2020). RBD mutations from circulating SARS-CoV-2 strains enhance the structural stability and human ACE2 affinity of the spike protein. *bioRxiv*.

Pal M, Berhanu G, Desalegn C, et al. (2020). Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. *Cureus* 12(3): e7423. doi:10.7759/cureus.7423.

Pallesen, J. et al. (2017). Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl. Aca. Sci.* 114(35), E7348-E7357.

Passanisi, S., Lombardo, F., Salzano, G. and Pajno, G. B. (2020). Are children most of the submerged part of SARS-CoV-2 iceberg? *Front. Pediatr.* doi: 10.3389/fped.2020.00213.

Phan, T. (2020). Genetic diversity and evolution of SARS-CoV-2. *Infect. Genet. Evol.* 81, 104260.

Rahman, M. S., Hoque, M. N., Islam, M. R., Akter, S., Rubayet-Ul-Alam, A. S. M., Siddique, M. A. et al. (2020). Epitope-based chimeric peptide vaccine design against S, M and E proteins of SARS-CoV-2 etiologic agent of global pandemic COVID-19: an *in silico* approach. *bioRxiv* doi: https://doi.org/10.1101/2020.03.30.015164.

Regla-Nava, J. A. et al. (2015). Severe acute respiratory syndrome coronaviruses with mutations in the E protein are attenuated and promising vaccine candidates. *J. Virol.* 89, 3870–3887.

Rehman, H. M., Mirza, M. U., Saleem, M., Froeyen, M., Ahmad, S., Gul, R. et al. (2020). A Putative Prophylactic Solution for COVID-19: Development of Novel Multiepitope Vaccine Candidate against SARS-COV-2 by Comprehensive Immunoinformatic and Molecular Modelling Approach. *Preprints* 2020030242, doi: 10.20944/preprints202003.0242.v1.

Reynard, O., Nguyen, X. N., Alazard-Dany, N., Barateau, V., Cimarelli, A., and Volchkov, V. E. (2015). Identification of a new ribonucleoside inhibitor of Ebola virus replication. *Viruses* 7(12), 6233-6240.

Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P. et al. (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 300(5624), 1394-1399.

Sardar, R., Satish, D., Birla, S., and Gupta, D. (2020). Comparative analyses of SAR-CoV2 genomes from different geographical locations and other coronavirus family genomes reveals unique features potentially consequential to host-virus interaction and pathogenesis. *bioRxiv* doi: https://doi.org/10.1101/2020.03.21.001586.

Schoeman, D., and Fielding, B. C. (2019). Coronavirus envelope protein: current knowledge. *Virol J.* 16, 69.

Shang, W., Yang, Y., Rao, Y., and Rao, X. (2020). The outbreak of SARS-CoV-2 pneumonia calls for viral vaccines. *npj Vaccines* 5(1), 1-3.

Shanmugaraj, B., Siriwattananon, K., Wangkanont, K., and Phoolcharoen, W. (2020). Perspectives on monoclonal antibody therapy as potential therapeutic intervention for Coronavirus disease-19 (COVID-19). *Asian Pac. J. Allergy Immunol.* 38(1), 10-18.

Sheahan, T. P., Sims, A. C., Zhou, S., Graham, R. L., Pruijssers, A. J., Agostini, M. L. et al. (2020). An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. *Science Translatl. Med.* eabb5883, doi: 10.1126/scitranslmed.abb5883.

Shen, Z. et al. (2020). Genomic diversity of SARS-CoV-2 in Coronavirus Disease 2019 patients. *Clinical Infect. Dis.* (2020).

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., and Siddique, R. (2020). COVID-19 infection: origin, transmission, and characteristics of human coronaviruses. *J. Adv. Res.*

Shi, S. Q. et al. (2006). The expression of membrane protein augments the specific responses induced by SARS-CoV nucleocapsid DNA immunization. *Mol. Immunol.* 43, 1791–1798.

Shi, H., Han, N., Jiang, Y., Cao, O., Alwalid, J., Gu et al. (2020). Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *The Lancet Infect. Dis.* 20, 425-434.

Song, H. D., Tu, C. C., Zhang, G. W., Wang, S. Y., Zheng, K., Lei, L. C. et al. (2005). Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. *Proc. Natl. Acad. Sci.* 102(7), 2430-2435.

Song, W., Gui, M., Wang, X., and Xiang, Y. (2018). Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathog*. 14(8), e1007236.

Sui, J., Li, W., Roberts, A., Matthews, L. J., Murakami, A. et al. (2005). Evaluation of human monoclonal antibody 80R for immunoprophylaxis of severe acute respiratory syndrome by an animal study, epitope mapping, and analysis of spike variants. *J. Virol.* 79: 5900–5906.

Sun, J., He, W. T., Wang, L., Lai, A., Ji, X., Zhai, X. et al. (2020). COVID-19: epidemiology, evolution, and cross-disciplinary perspectives. *Trends Mol. Med.* https://doi.org/10.1016/j.molmed.2020.02.008.

884 Tang, X., Wu, C., Li, X., Song, Y., Yao, X., Wu, X. et al. (2020). On the origin and continuing
885 evolution of SARS-CoV-2. *Nat. Sci. Rev.* nwaa036,
886 https://doi.org/10.1093/nsr/nwaa036.
887 ter Meulen, J. et al. (2006). Human Monoclonal Antibody Combination against SARS
888 Coronavirus: Synergy and Coverage of Escape Mutants. *PLoS Med.* 3(7): e237.
889 https://doi.org/10.1371/journal.pmed.0030237.
890 To, K. K. W. et al. (2020). Consistent detection of 2019 novel coronavirus in saliva. Clinical
891 infectious diseases: an official publication of the Infectious Diseases Society of
892 America. *ciaa149* https://doi.org/10.1093/cid/ciaa149.
893 Toots, M., Yoon, J. J., Cox, R. M., Hart, M., Sticher, Z. M., Makhsous, N. et al. (2019).
894 Characterization of orally efficacious influenza drug with high resistance barrier in
895 ferrets and human airway epithelia. *Science Translatl. Med.* 11(515).
896 Traggiai, E. et al. (2004) An efficient method to make human monoclonal antibodies from
897 memory B cells: Potent neutralization of SARS coronavirus. *Nat. Med.* 10: 871–875.
898 Udugama, B. et al. (2020) Diagnosing COVID-19: the disease and tools for detection. *ACS*
899 *Nano* DOI: 10.1021/acsnano.0c02624.
900 ul Qamar, M. T., Saleem, S., Ashfaq, U. A., Bari, A., Anwar, F., and Alqahtani, S. (2019).
901 Epitope-based peptide vaccine design and target site depiction against Middle East
902 Respiratory Syndrome Coronavirus: an immune-informatics study. *J. Translational*
903 *Med.* 17(1), 362.
904 Vashist, S. K. (2020). *In Vitro* Diagnostic Assays for COVID-19. *Recent Adv. Emerg. Trends.*
905 *Diag.* 2020, 10(4), 202; https://doi.org/10.3390/diagnostics10040202
906 Velavan, T. P., and Meyer, C. G. (2020). The COVID-19 epidemic. *Trop. Med. Int. Health*
907 25(3), 278-280.
908 Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020).
909 Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*.
910 Wan, Y., Shang, J., Graham, R., Baric, R. S., and Li, F. (2020). Receptor recognition by the
911 novel coronavirus from Wuhan: an analysis based on decade-long structural studies of
912 SARS coronavirus. *J. Virol.* 94(7).
913 Wang, C., Liu, Z., Chen, Z., Huang, X., Xu, M., He, T., and Zhang, Z. (2020d). The
914 establishment of reference sequence for SARS-CoV-2 and variation analysis. *J. Med.*
915 *Virol*. PMID: 32167180, doi: 10.1002/jmv.25762.

916 Wang, M. et al. (2020b). Remdesivir and chloroquine effectively inhibit the recently emerged
917 novel coronavirus (2019-nCoV) *in vitro*. *Cell Res* 269-271. 10.1038/s41422-020-0282-
918 0.
919 Wang, N. et al. (2019). Structural Definition of a Neutralization-sensitive Epitope on the
920 MERS-CoV S1-NTD. *Cell Rep.* 28(13), 3395-405.
921 Wang, N., Shang, J., Jiang, S., and Du, L. (2020c). Subunit vaccines against emerging
922 pathogenic human coronaviruses. *Front. Microbiol.* 11, 298.
923 Wang, X. et al. (2020a). Clustered Regularly Interspaced Short Palindromic Repeats/Cas9-
924 Mediated Lateral Flow Nucleic Acid Assay. ACS Nano 2020, 14 (2), 2497–2508.
925 Wei, Q., and Ren, Z. (2020). Disinfection measures for pneumonia foci infected by novel
926 coronavirus in 2019. *Chin. J. Disinfect.* 37, 59-62.
927 WHO. (2020). Statement on the second meeting of the International Health Regulations (2005)
928 Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV).
929 2020. https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-
930 meeting-of-the-international-health-regulations-(2005)-emergency-committee-
931 regarding-the-outbreak-of-novel-coronavirus-(2019-ncov). Accessed 1 Feb 2020.
932 Wong, M.C. et al. (2020). Evidence of recombination in coronaviruses implicating pangolin
933 origins of nCoV-2019. *bioRxiv* https://doi.org/ 10.1101/2020.02.07.939207 57.
934 Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O. et al. (2020).
935 Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*
936 367(6483), 1260-1263.
937 Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P. et al. (2020a). Genome composition
938 and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host*
939 *Microbe*. 27(3), 325-328.
940 Wu, F. et al. (2020b). A new coronavirus associated with human respiratory disease in China.
941 *Nature* 579, 265–269(2020).
942 Xiao, K. et al. (2020) Isolation and characterization of 2019- nCoV-like coronavirus from
943 Malayan pangolins. *bioRxiv* https://doi.org/10.1101/ 2020.02.17.951335.
944 Xinhua, (2020). China fast-tracks novel coronavirus vaccine development xinhua. [cited 202
945 29 January]; Available from: http://www.xinhuanet.com/english/2020-
946 01/28/c_138739378.htm.

947 Xu, Y. et al. (2020). Characteristics of pediatric SARS-CoV-2 infection and potential evidence
948 for persistent fecal viral shedding. *Nat. Med.* 26, 502–505.
949 Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., & Zhou, Q. (2020). Structural basis for the
950 recognition of SARS-CoV-2 by full-length human ACE2. *Science*, 367(6485), 1444-
951 1448.
952 Yang, D., and Leibowitz, J. L. (2015). The structure and functions of coronavirus genomic 3′
953 and 5′ ends. *Virus Res.* 206, 120-133.
954 Yin, C. (2020). Genotyping coronavirus SARS-CoV-2: methods and implications. *arXiv*
955 preprint *arXiv* 2003.10965.
956 Yong, C.Y. et al. (2019). Recent Advances in the Vaccine Development Against Middle East
957 Respiratory Syndrome-Coronavirus. *Front. Microbiol.* 10, 1781.
958 Yuan, M. et al. (2020). A highly conserved cryptic epitope in the receptor-binding domains of
959 SARS-CoV-2 and SARS-CoV. *Science*, eabb7269; DOI: 10.1126/science. abb7269.
960 Zaki, A. M., Van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. et al. (2012). Isolation of
961 a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.*
962 367(19), 1814-1820.
963 Zhai, P., Ding, Y., Wu, X., Long, J., Zhong Y., and Li, Y. (2020) The epidemiology, diagnosis
964 and treatment of COVID-19. *Int. J. Antimicrobial Agents*
965 http://doi:10.1016/j.ijantimicag.2020.105955.
966 Zhang, J., Zeng, H., Gu, J., Li, H., Zheng, L., and Zou, Q. (2020a). Progress and Prospects on
967 Vaccine Development against SARS-CoV-2. *Vaccines* 8(2), 153.
968 Zhang, F., Jonathan, Abudayyeh, O. O., and Gootenberg, J. S. (2020b). A protocol for
969 detection of COVID-19 using CRISPR diagnostics.
970 https://www.broadinstitute.org/files/publications/special/COVID19.
971 Zhou, H. et al. (2019). Structural definition of a neutralization epitope on the N-terminal
972 domain of MERS-CoV spike glycoprotein. *Nat. Commun.* 10, 3068.
973 Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. (2020a). A pneumonia outbreak
974 associated with a new coronavirus of probable bat origin. *Nature*, doi: 10.1038/s41586-
975 020-2012-7. PubMed PMID: 32015507.
976 Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W., and Cheng, F. (2020b). Network-based
977 drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Dis.* 6(1), 1-
978 18.

 doi:10.20944/preprints202004.0359.v1

Zhu, N. et al. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* 382, 727-733. DOI:10.1056/NEJMoa2001017.

Table 1: Genome-wide nucleotide mutations and amino-acid mutations and substitutions in SARS-CoV-2 strains. The number in the parentheses indicated the missense mutations.

Genome-site/position	No. of amino-acid replacements	No. of nucleotide mutations	References
Polyprotein (nsp)	404	655	Islam et al. (2020)
	757		Yin (2020)
Leader sequence	178		Yin (2020)
Spike (S) glycoprotein	114	173	Islam et al. (2020)
	14 (8)		Phan (2020)
	183		Yin (2020)
	7	11	Wang et al. (2020d)
	13		Huang et al. (2020)
	18		Lu et al. (2020)
	6		Andersen et al. (2020)
Membrane (M) protein	15	30	Islam et al. (2020)
	2 (1)		Phan (2020)
	33		Yin (2020)
		5	Wang et al. (2020d)
	2		Huang et al. (2020)
Envelop (E) protein	10	25	Islam et al. (2020)
	2		Huang et al. (2020)
Nucleocapsid (N) protein	76	109	Islam et al. (2020)
	7 (4)		Phan (2020)
	6	17	Wang et al. (2020d)
	5		Huang et al. (2020)
	222		Yin (2020)
Open-reading frames (ORFs)			
ORF1a	44		Huang et al. (2020)
ORF1ab	48 (29)		Phan (2020)
ORF1ab	8		Huang et al. (2020)
ORF1ab	6	43	Wang et al. (2020d)
	48	64	Islam et al. (2020)
ORF3a	49		Yin (2020)
	7		Huang et al. (2020)
	6	6	Wang et al. (2020d)
ORF6	5	8	Islam et al. (2020)
ORF7a	21	30	Islam et al. (2020)
	2		Huang et al. (2020)
ORF7b	4	9	Islam et al. (2020)
ORF8	15	19	Islam et al. (2020)
	8		Huang et al. (2020)
	34	34	Wang et al. (2020d)
ORF10	10	16	Islam et al. (2020)
	1		Huang et al. (2020)
5´-UTR		105	Islam et al. (2020)
		8	Phan (2020)
3´-UTR		158	Islam et al. (2020)
	3		Phan (2020)
3´-to-5´exonuclease	62		Yin et al. (2020)
Spacer region		6	Islam et al. (2020)
	6		Phan (2020)

Here nsp, non-structural proteins; ORF, open open-reading frames; UTR, untranslated region.

1005 **Table 2:** Diagnostic protocols developed for SARS-CoV-2
1006

Type of clinical sample	Method/platform (technology)	Target gene/Biomarker	Who developed	References
Upper and lower respiratory specimens*	Real-Time RT-PCR	*N* gene	U.S. CDC	Anonymous (2020a)
Upper and lower respiratory specimens*	Real-Time RT-PCR	*ORF1ab* and *N* gene	China, CDC	Anonymous (2020b)
Respiratory specimens	Real-Time RT-PCR	*RdRp*, *E* and *N* genes	Multicountries: Germany, The Netherlands, China, France and UK	Corman et al. (2020)
Respiratory specimens	Real-Time RT-PCR	*RdRp*/Hel, *S* and *N* genes	Hong Kong, China	Chan et al. (2020a)
Saliva	Real-Time RT-PCR	*S* gene	Hong Kong	To et al. (2020)
Human clinical specimen	Real-Time RT-PCR	*ORF1b-nsp*14 and *N* genes	Hong Kong University	Anonymous (2020c)
Pharyngeal swab	Real-Time RT-PCR	*N* gene	National Institute of Infectious Diseases in Japan	Nao et al. (2020)
Serum	CRISPR-Cas (RPA)	Nucleic acid biomarker	China	Wang et al. (2020a)
Nasopharyngeal swabs	CRISPR-Cas (RT-RPA)	Nucleic acid biomarker	USA	Kellner et al. (2020)
Synthetic COVID-19 virus RNA fragment	CRISPR-based SHERLOCK (dipstick)	*ORF1ab* and *S* genes	MIT, USA	Zhang et al. (2020b)
Throat, nasal, nasopharyngeal or	ID NOW™ COVID-19	*RdRp* gene	Abbott	https://bit.ly/3b0W8bd

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2020 doi:10.20944/preprints202004.0359.v1

oropharyngeal swabs				
Human finger pricks or venous whole blood, serum, and plasma	Immunoassay	IgM/IgG	BioMedomics, USA	https://bit.ly/2UXh5OF
Human finger pricks or venous whole blood, serum, and plasma	Immunoassay	IgM/IgG	China	Li et al. (2020)
Human finger pricks or venous whole blood, serum, and plasma	Immunoassay	IgM/IgG	Diazyme	https://bit.ly/2UXlils

* nasopharyngeal or oropharyngeal swabs, sputum, lower respiratory tract aspirates, bronchoalveolar lavage, and nasopharyngeal wash/aspirate or nasal aspirate; RPA, recombinase polymerase amplification.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2020 doi:10.20944/preprints202004.0359.v1

Figures

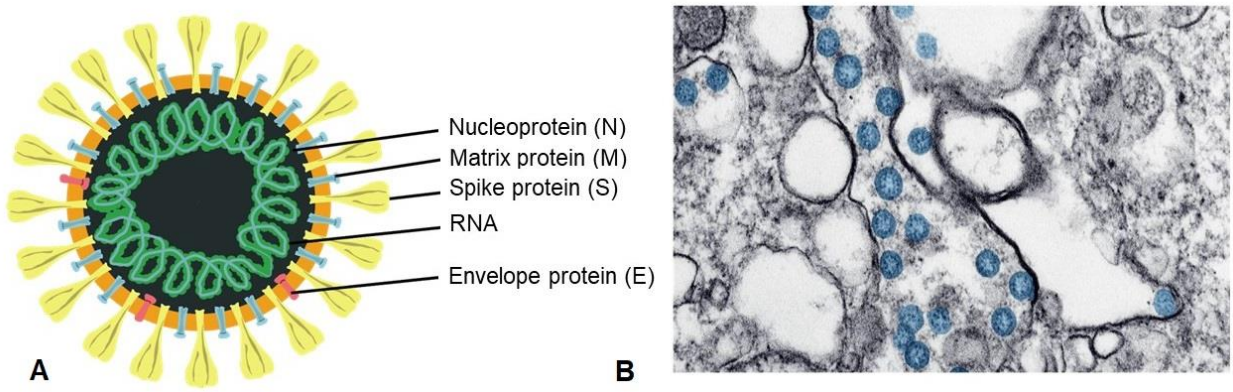


FIGURE 1 | SARS-CoV-2 morphology. (A) Illustrated of the viral structure with its structural viral proteins (adapted from Udugama et al. 2020). (B) Spherical viral particles of SARS-CoV-2 in a cell is visualized by the transmission electron microscopy (Details - Public Health Image Library(PHIL) https://phil.cdc.gov/Details.aspx?pid=23354 (accessed April 15, 2020). The virus (60-140 nm) is colorized in blue (adapted from the US Centers for Disease Control).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2020 doi:10.20944/preprints202004.0359.v1

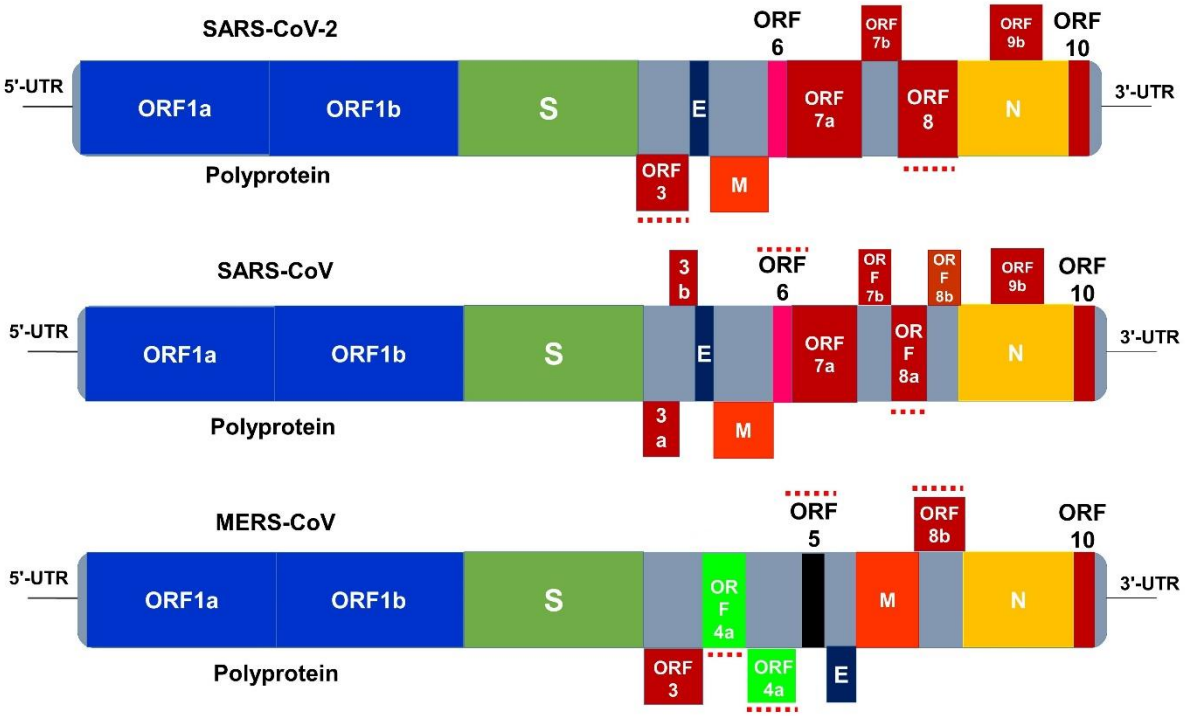
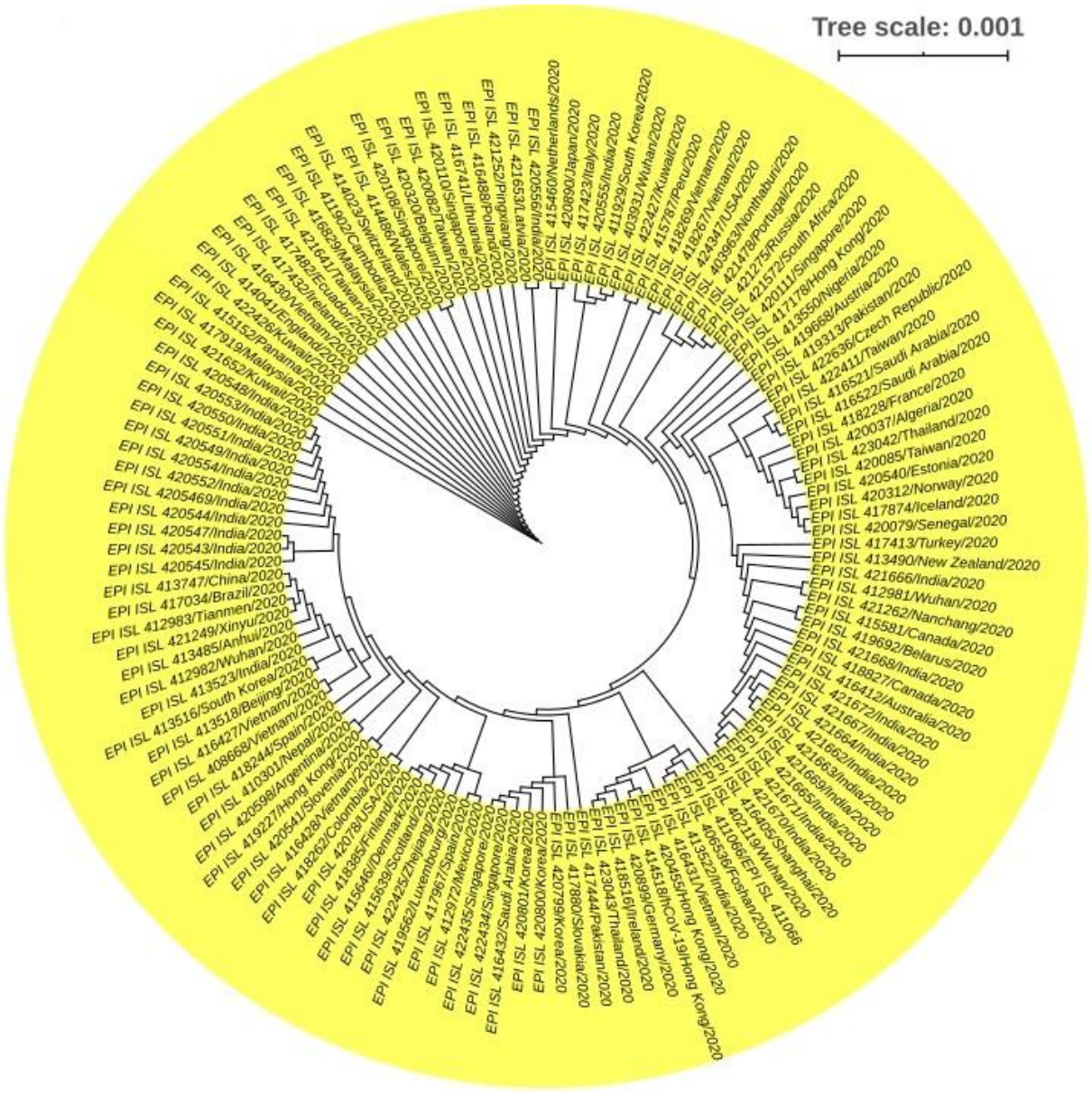


FIGURE 2 | Genome organization of SARS-CoV-2, SARS-CoV and MERS-CoV. The genome of these three viruses comprises the 5′-untranslated region (5′-UTR), polyprotein with open reading frame (orf) 1a/b (blue box) representing non-structural proteins (nsp) for replication, structural proteins including S glycoprotein (dark green box), envelop (E) (dark blue box), membrane (M) (orange box), and nucleocapsid (N) (yellow box) proteins, accessory proteins such as orf 3a/b (red boxes), 5 (black box), 6 (pink box), 7a/b, 8a/b, 9b and 10 (red boxes) , and the 3′-untranslated region (3′-UTR). The doted red lines (both in above and under) are the protein which show key differences among SARS-CoV-2, SARS-CoV and MERS-CoV. The nsps and orfs lengths are not drawn in scale (adapted from Islam et al., 2020; Shereen et al., 2020).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2020 doi:10.20944/preprints202004.0359.v1



1067 **Figure 3 | Phylogenetic tree of SARS-CoV-2.** 200 complete genome sequences of SARS-CoV-2
1068 retrieved from global initiative on sharing all influenza data (GISAID) (https://www.gisaid.org/)
1069 from different countries were used to build this tree. The sequences were aligned using MAFFT
1070 online server (Katoh et al., 2002), and a maximum likelihood tree was built with iTOL (interactive
1071 Tree Of Life). Each node represents a single strain which is found to be patient and/or sample
1072 specific, and not clustered according to geographical locations. Tree scale 0.01, represents days
1073 before the time of lastly sampled genomes by scale*365.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 April 2020 doi:10.20944/preprints202004.0359.v1

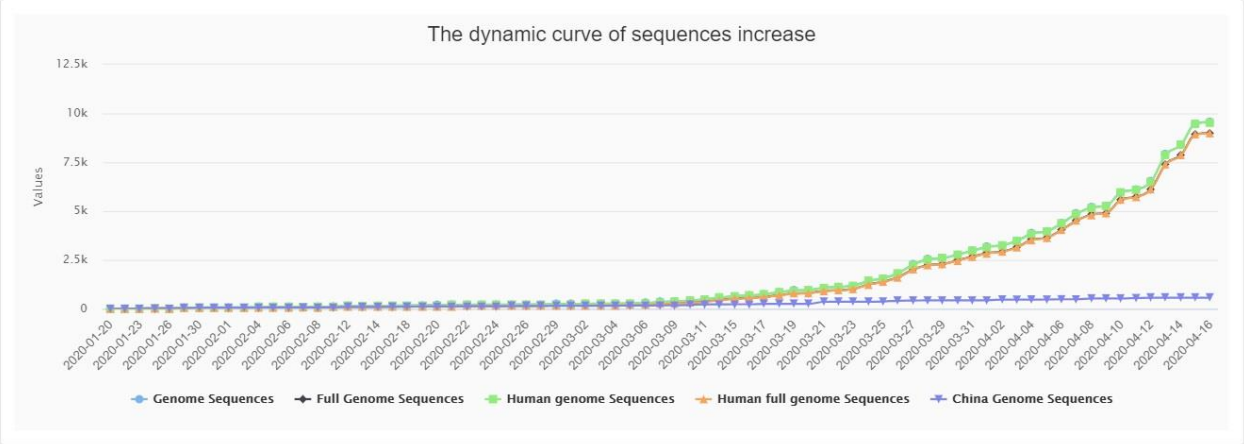


FIGURE 4 | The dynamic curve showing daily increase in complete genome sequences of SARS-CoV-2 strain (s) from different patients across the globe, and being submitted to the reference databases. The data were collected from China National Center for Bioinformation 2019 Novel Coronavirus Resource (2019nCoVR) with available sequences from different countries (as on April 16, 2020).