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.

Gene 752 (2020) 144792

Contents lists available at ScienceDirect

Gene

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

Letter to the Editor

Overwhelming mutations or SNPs of SARS-CoV-2: A point of caution

ARTICLE INFO

Keywords:
Coronavirus
Sequence
Genome
Evolution
COVID-19

ABSTRACT

The morbidity of SARS-CoV-2 (COVID-19) is reaching 3 Million landmark causing and a serious public health concern globally and it is enigmatic how several antiviral and antibody treatments were not effective in the different period across the globe. With the drastic increasing number of positive cases around the world WHO raised the importance in the assessment of the risk of spread and understanding genetic modifications that could have occurred in the SARS-CoV-2. Using all available deep sequencing data of complete genome from all over the world (NCBI repository), we identified several hundreds of point mutations or SNPs in SARS-CoV-2 all across the genome. This could be the cause for the constant change and differed virulence with an increase in mortality and morbidity. Among the 12 different countries (one sequence from each country) with complete genome sequencing data, we noted the 47 key point mutations or SNPs located along the entire genome that might have impact in the virulence and response to different antivirals against SARS-CoV-2. In this regard, key viral proteins of spike glycoprotein, Nsp1, RdRp and the ORF8 region got heavily mutated within these 3 months via person-to-person passage. We also discuss what could be the possible cause of this rapid mutation in the SARS-CoV-2.

The pandemic Corona Virus Disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is an urgent public health emergency and made a serious impact in global health and economy (https://www.cdc.gov/coronavirus/2019-ncov). The SARS-CoV-2 pandemic is the most serious hit in the last 10 years and caused more than 170,500 deaths globally and the morbidity of this viral infection reaching 3 Million landmark (www.coronavirus.gov). Epidemiologists predict the several more spike in the coronavirus infection could rise in different countries with higher density in population (Verity et al., 2020). Since the initial reports on this pneumonia-causing novel coronavirus (SARS-CoV-2) in Wuhan, China (Huang et al., 2020), mortality and morbidity are increasing exponentially around the globe despite several antiviral treatments. With the drastic increasing number of the positive cases around the world, the World Health Organization (WHO) raised the importance in the assessment of the risk of spread and understanding genetic modification that could have occurred in the SARS-CoV-2 (www.coronavirus.gov). Hence it is worthwhile to look for any mutations or SNPs in SARS-CoV-2 alone that could the cause for the constant change and virulence causing increase in continuing mortality and morbidity.

To this end, we aimed to look for mutations and SNPs in the complete genomes of SARS-CoV-2 worldwide where the sequencing data was collected using the next generation sequencing and deposited in the NCBI and all other repository. As of 24th March 2020, out of 172 countries with confirming positive cases, only 12 countries have sequenced the complete genome of SARS-CoV-2 (https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/) (Fig. 1). Furthermore, there are about 106 complete and validated sequence data sets available in the NCBI database (https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/) and (https://bigd.big.ac.cn/ncov/). Surprisingly, we noticed several hundreds of point mutations or SNPs among the different isolates from all over the world with different sequence data sets (Fig. 1A and B) and (Fig. S1A and S2). And 47 key point mutations or SNPs were located along the entire genome in the sequence just in 12 different countries (single sequence comparison), these mutations involved in the different protein-protein recognition (Fig. 1C). Point mutations or SNPs have great implications for the target drug binding and receptor binding (Puty et al., 2019). The overall mutations phylogeny shows the 3 groups (https://www.gisaid.org/) of mutations which are evolved in these 3 months. Predominantly, the mutations were also found in the different vital proteins of SARS-CoV-2 (spike glycoprotein, Nsp1, RdRp and others) (Figs. S1 and S2) and warrants epidemiologists and medical fraternity for the use of drug treatment options. This also suggests that SARS-CoV-2 is highly venerable to have quick changes and mutate even during the person-to-person transmission. This also helps to overcome the previous misconception of SARS-CoV-2 may not get mutated during person-to-person transmission (Andersen et al., 2020). The rate and number of SNPs or mutations in SARS-CoV-2 within three months of outbreak underlines the complexity of virus to handle and corroborate the quick evolution of SARS-CoV-2 (Fig. S3).

We next looked at individual mutations that occurred in the three different prime proteins (i) outer membrane spike glycoprotein, (ii) Non-structural protein (Nsp1) and (iii) RNA-dependent RNA Polymerase (RdRp) (Vankadari and Wilce, 2020; Wrapp et al., 2020; Narayanan et al., 2008; Miller et al., 2020) (Fig. 1B–D). Foremost in spike glycoprotein, the mutations or SNPs occurred are surface exposed and in the solvent-accessible regions some involved in the host receptor (ACE3 and CD26) binding and Furin cleavage site (Fig. S1C and D)

Abbreviations: SNP, single nucleotide polymorphism; ACE2, angiotensin-converting enzyme 2; CD26, cluster of differentiation 26; WHO, World Health Organization; NCBI, National Center for Biotechnology Information; Nsp1, nucleoporin NSP1; Rdrp, RNA-dependent RNA polymerase

https://doi.org/10.1016/j.gene.2020.144792
Received 24 March 2020; Received in revised form 22 April 2020; Accepted 18 May 2020
Available online 20 May 2020

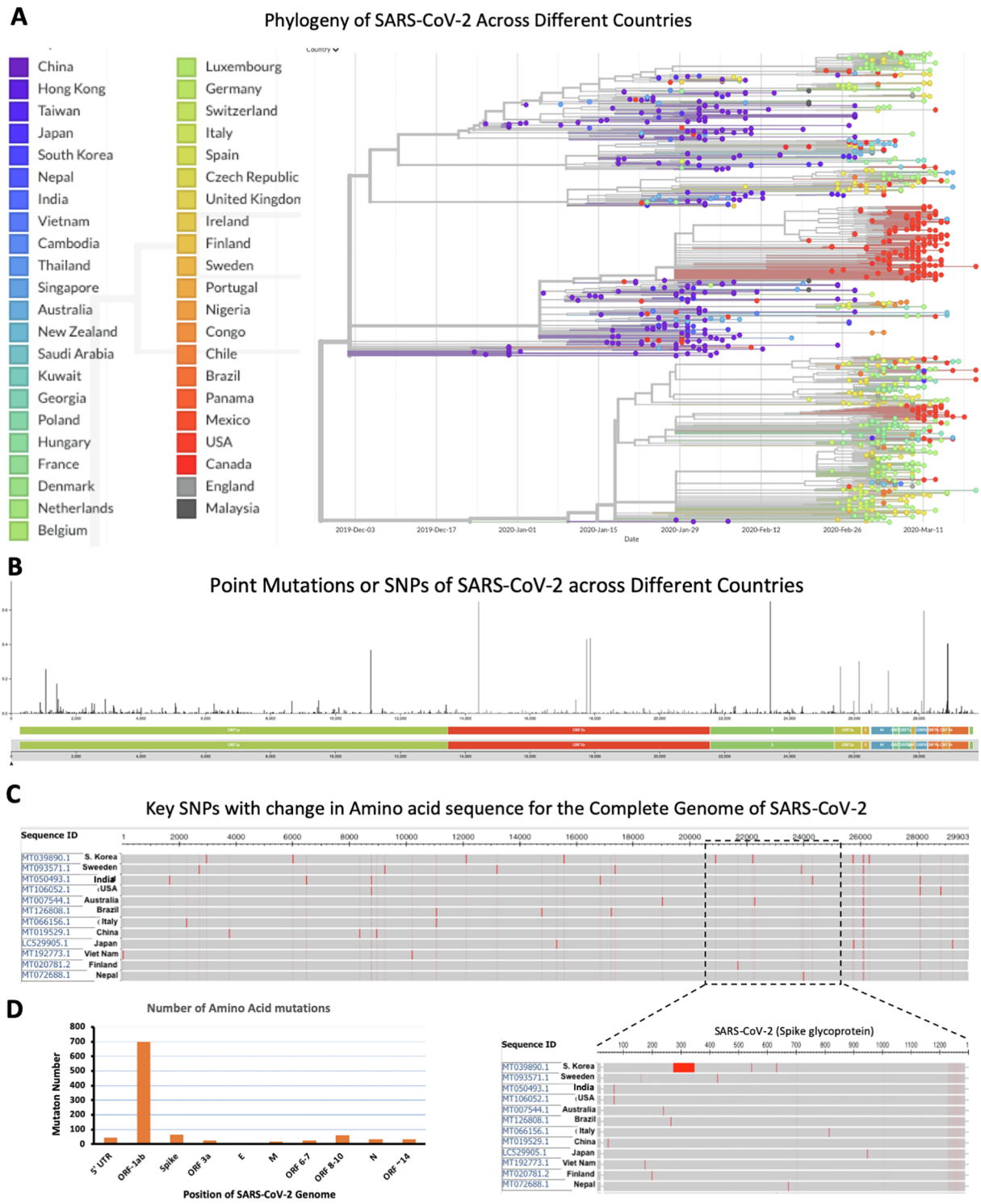

Fig. 1. (A) Phylogenetic tree showing the evolution of SARS-CoV-2 from the initial origin China (2020/01/17). The tree represents the mutations or SNPs that resulted in the evolution of current SARS-CoV-2 in the last three months. Individual countries are colored in as shown in the color key. (B) Position and number of SNPs across are the genome is denoted with bar graph. (C) The key mutations with change in amino acid observed across the whole genome of 12 countries are listed and highlighted in red lines. Enlarged view showing the mutations occurred in spike glycoprotein. (D) Bar graph depicting the number of mutations and position is complete genome.

(Vankadari and Wilce, 2020; Wrapp et al., 2020; Coutard et al., 2020). These mutations could make the virus more resistant to receptor binding inhibitors and some of the antibodies. These warrants epidemiologists, medical fraternity and pharmacologists to use the antivirals and antibodies which are compatible and promising to their strain of SARS-CoV-2 and not to be generalised in some cases. Furthermore, we also noticed a number of mutations or SNPs in other key proteins Nsp1 and RdRp (Fig. 1B and S1A and S2). Interestingly, most mutations were also surface exposed and show the high possibility to impede the antiviral drugs.

Although there was only 12 out of 172 countries complete genome sequence is available for now, genome analysis data is adequate to

prove the rapid rate of mutations and SNPs striking in SARS-CoV-2. The rate at which SARS-CoV-2 getting mutated and evolved was least noticed till date and it needs further attention and research for the cause for the rapid genetic alterations and how it can be regulated, which is a serious concern. The notable mutations in the viral vital proteins (spike glycoprotein, Nsp1 and RdRp) and at the host or drug integration region raise the speculation on the changes in the cell entry and response to the different antiviral treatments. Besides, several SNPs were also noticed in the other regions such as ORF-1ab (codes for 17 structural proteins) and ORF-8/10 (Fig. S1A). Furthermore, NSP1 and ORF8 are two the particular hot spot area, where the mutations and deletions were noticed in the early onset of infection late January and the rate of mutations were quickly increased (~35%) then declined or stabilised in the later passages of the virus (Fig. S1B). It is also interesting to notice that ORF8 in SARS-CoV-2 is not split into 8a and 8b as observed in SARS-CoV-1 (Chan et al., 2020; Lau et al., 2015). However, our understanding of the structure and functional importance of ORF8 is limited, which need further attention. We also like to extend our discussion that the observed contact changes in the genome of SARS-CoV-2 is not only limited to mutation or SNPs but also there are several nucleotide deletions and insertions were also noticed across the genome. These observations has great importance for assessing the increased virulence and efficacy of antivirals and raises possible cause for the overwhelming mutations in SARS-CoV-2. One such hypothesis could be the extensive use of a various combination of antiviral drugs that could have to lead the sequential mutations in the virus (Manrubia et al., 2005; Peck and Lauring, 2018; Sanjuan et al., 2010). This study also warrants the importance of sequencing the whole genome of SARS-CoV-2 after several passages and key mutations or SNPs should be noted for the effective drug designing and treatment options such as antiviral and immune therapy. With the exponential SARS-CoV-2 public-health emergency, it is essential to know the epidemiology and evolution of the virus to be prepared for the combat.

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.

Acknowledgements

We thank GISAID EpiFlu™ (www.gisaid.org/) and NGDC (www.bigd.big.ac.cn/) for giving permission and agreement of using the sequencing data and graphical images used for this manuscript. I also acknowledge Joseph Polidano of Monash University for editing and proof reading the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2020.144792.

References

Verity, R., et al., 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect. Dis.
Huang, C., et al., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395 (10223), 497–506.
Puty, T.C., et al., 2019. Evaluation of the impact of single-nucleotide polymorphisms on treatment response, survival and toxicity with cytarabine and anthracyclines in patients with acute myeloid leukaemia: a systematic review protocol. Syst. Rev. 8 (1), 109.
Andersen, K.G., et al., 2020. The proximal origin of SARS-CoV-2. Nat. Med. 26, 450–452.
Vankadari, N., Wilce, J.A., 2020. Emerging WuHan (COVID-19) coronavirus: glycan shield and structure prediction of spike glycoprotein and its interaction with human CD26. Emerg. Microbes Infect. 9 (1), 601–604.
Wrapp, D., et al., 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367 (6483), 1260–1263.
Narayanan, K., et al., 2008. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. J. Virol. 82 (9), 4471–4479.
Miller, C.M., et al., 2020. Systemic expression of a viral RdRP protects against retrovirus infection and disease. J. Virol. 16 (9), 94.
Coutard, B., et al., 2020. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antiviral Res. 176, 104742.
Chan, J.F., et al., 2020. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerg. Microbes Infect. 9 (1), 221–236.
Lau, S.K., et al., 2015. Severe acute respiratory syndrome (SARS) coronavirus ORF8 protein is acquired from SARS-related coronavirus from greater horseshoe bats through recombination. J. Virol. 89 (20), 10532–10547.
Manrubia, S.C., et al., 2005. High mutation rates, bottlenecks, and robustness of RNA viral quasispecies. Gene 347 (2), 273–282.
Peck, K.M., Lauring, A.S., 2018. Complexities of viral mutation rates. J. Virol. 92 (14), 1031–1117.
Sanjuan, R., et al., 2010. Viral mutation rates. J. Virol. 84 (19), 9733–9748.

Naveen Vankadari
Monash Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Victoria 3800, Australia
E-mail address: Naveen.vankadari@monash.edu.