Rapid Reviews COVID-19

Review 1: "Emergence of a novel SARS-CoV-2 strain in Southern California, USA"

Takahiko Koyama¹

¹IBM Research, Computational Biology Center, USA

Published on: Feb 11, 2021

DOI: 10.1162/2e3983f5.74f370ca

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

RR:C19 Evidence Scale rating by reviewer:

- **Reliable.** The main study claims are generally justified by its methods and data. The results and conclusions are likely to be similar to the hypothetical ideal study. There are some minor caveats or limitations, but they would/do not change the major claims of the study. The study provides sufficient strength of evidence on its own that its main claims should be considered actionable, with some room for future revision.

Review:

Lately, N501Y and other strains which might have distinct characteristics of clinical importance have been discovered globally. Identifying and characterizing these variants for infectivity and potential vaccine escapes are crucial to containing the virus from further expansion. In this study, Zhang et al. have focused on CAL.C20 (452R.V1 and PANGO B.1.429), which gained momentum since last October in Southern California.

The defining haplotype for CAL.C20 consists of ORF1a: I4205V (NSP9 I65V), ORF1b: D1183Y (NSP13 D260Y), S: S13I; W152C; L452R. Among these variants, the most notable one is L452R, which is located within the receptor-binding domain (RBD) and it is reported to lead to higher RBD protein expression by Starr et al despite almost no significant enhanced binding affinity with ACE2. Furthermore, W152C might alter cysteine disulfide bonding partners and might induce major structural change in spike protein. Spike S13I and NSP9 I65V seem to be more recently joined compared with the other three variants. Although this is not a mandatory request, to accommodate more strains, the defining haplotype might better be relaxed by dropping Spike S13I and NSP I65V. Among genomes sequenced globally, a strain discovered in Mexico in early July (GISAID2: EPI_ISL_942929) is the oldest discovered in this strain; therefore, the strain has been existing for quite some time a few months earlier than the outbreak observed in Southern California.

Although the manuscript is succinct and well written, it is advised that the prevalence of the strain across the US and the rest of the world need to be updated. For instance, besides Oceania, the strain has been identified in Japanese Airport Quarantines among travelers from US and Mexico. Furthermore, authors better unify nomenclatures to either PANGO3 or NextStrain4 throughout the document, for instance, B.1.1.7 to be noted as 20I/501Y.V1 or vice versa.

The authors played an important role in identifying the emerging strain, which potentially aggravates the COVID-19 situations in other regions as well as Southern California. Only minor issues were identified in the manuscript. Therefore, I recommend just minor revisions for the acceptance.

- 1 Starr, T. N. *et al.* Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *bioRxiv : the preprint server for biology*, 2020.2006.2017.157982, doi:10.1101/2020.06.17.157982 (2020).
- 2 Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill* **22**, doi:10.2807/1560-7917.ES.2017.22.13.30494 (2017).
- 3 Rambaut, A. *et al.* A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology* **5**, 1403-1407, doi:10.1038/s41564-020-0770-5 (2020).
- 4 Emma B Hodcroft, J. H., Richard A Neher, Trevor Bedford. Year-letter Genetic Clade Naming for SARS-CoV-2 on Nextstrain.org. (2020).