

Large-scale analysis of SARS-CoV2 spike protein variants and their impacts in Thailand

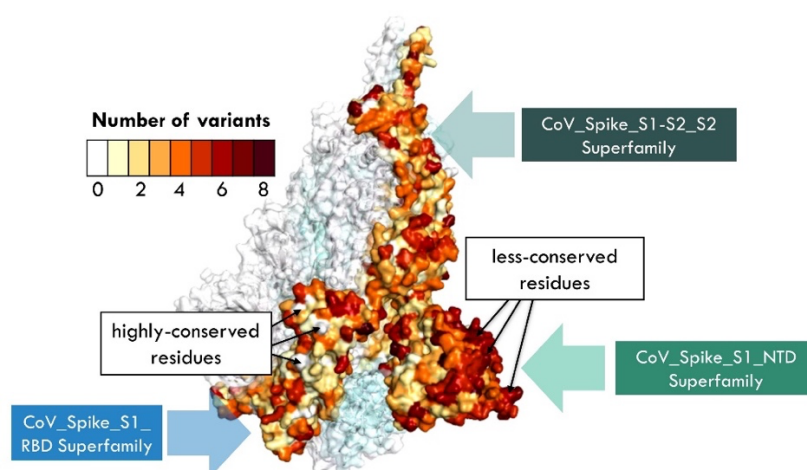
Shalip Yahangkiakan and Sirawit Ittisoponpisan *

Center for Genomics and Bioinformatics Research, Division of Biological Science, Faculty of Science, Prince of Songkla University, Songkhla, Thailand, 90112

* Correspondence: sirawit.i@psu.ac.th;

Abstract: With a large number of SARS-CoV2 sequences collected since the beginning of the COVID-19 outbreak, many genetic variations were yet to be analysed for their impacts. Moreover, little has been known about the effects of variants specific to Thai strains. In this study, we performed *in silico* analysis on 439,197 SARS-CoV2 spike protein sequences collected until 7 February 2021 from NCBI and GISAID databases to explore the distribution of variants on the SARS-CoV2 spike protein and their impacts. We identified mutation hotspots on the protein surface and several highly conserved residues located in the receptor-binding domain. These highly conserved residues could potentially serve as drug target candidates. Finally, we studied the effects of the variant of global concern Asp614Gly (D614G) and six variants exclusively found in Thailand (from 595 samples). Using 3D simulation, we showed that D614G disrupted inter-chain H bonds in the spike protein complex, while most Thai variants were structurally benign. The results from this study explain the mechanisms of variants observed in different strains of SARS-CoV2, aid variant prioritisation, and could provide implications for developing effective treatments or preventions for COVID-19.

Graphical abstract:



Copyright: © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords: coronavirus; SARS coronavirus 2; variant analysis; mutagenesis; spike protein; ACE2

Funding: This research was funded by the Faculty of Science, Prince of Songkla University.

Acknowledgements: The authors would like to thank Prof. Michael Sternberg and Dr Alessia David from Structural Bioinformatics Group, Imperial College London, UK, for their supporting help.