Borneo Journal of Resource Science and Technology (2023), 13(2): 103-110
DOI: https:doi.org/10.33736/bjrst.5853.2023

Analysis of Seven Human Respiratory Coronavirus (CoV) S Proteins from a Bioinformatics Approach

LEONARD WHYE KIT LIM* & HUNG HUI CHUNG

Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia
*Corresponding author: limwhyekitleonard@gmail.com
Received: 12 July 2023 Accepted: 11 August 2023 Published: 31 December 2023

ABSTRACT

The coronavirus disease 2019 (COVID-19) has caused a huge pandemic repercussion across the globe and it is mainly contributed by the human severe acute respiratory syndrome coronavirus (SARS-CoV-2). There are seven human respiratory coronaviruses identified to date, namely HCoV-229E, HCoV-NL63, HCoV-OC43, HCoV-HKU1, MERS-CoV, SARS-CoV and SARS-CoV-2. A recently published bioinformatic human CoV comparison only covered four human CoV. Therefore, in this study, a bioinformatics approach-based analyses route was taken to dissect the S proteins of all the available (seven) human respiratory coronaviruses publicly available in the GenBank database. The antigenic epitope amount is postulated to be the most accurate bioindicator among all in determining the severity of a particular human respiratory coronavirus. Other powerful bioinformatic indicators are global similarity index, maximum likelihood phylogenetic analysis as well as domain analysis. The data generated in this study can be channelled to the vaccine and antiviral drug development to combat the current and future spread of the human respiratory coronaviruses.

Keywords: Antigenic epitope, COVID-19, SARS-CoV, S protein

INTRODUCTION

The human severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the major contributor to coronavirus disease 2019 (COVID-19), which had caused unprecedented challenges worldwide with its rapid spread and severe impact on mortalities, public health system and economies globally (Lim, 2023). This single-stranded RNA-enveloped virus is categorised under the order Nidovirales and family Coronaviridae, whereby the family encompasses alpha-, beta-, gamma- and delta-coronavirus genera (Cicaloni *et al.,* 2022). The alpha-coronavirus can infect both animals and human whereas the beta-coronavirus such as SARS-CoV, HCoV-HKU1, MERS-CoV and SARS-CoV-2 infect mainly human. The remaining two genera primarily infect animals. This virus contains four structural proteins, namely spike (S), membrane (M), nucleocapsid (N) and envelope (E). The non-structural proteins found within the virus include ORF3a, ORF8, ORF7a, Nsp3, Nsp6, Nsp4 and Nsp12 (Cicaloni *et al*., 2022).

The surface of the viral envelope is where the S proteins are situated and it can be divided into S1 and S2 regions based on their functionalities. The receptor binding domain (RBD) that functions in the binding of viral-receptor is housed within the S1 region whereas the homology region structural domains 1 and 2 (HR1 and HR2) responsible for viral fusion is located within the S2 region (Xia *et al*., 2020). Attachment of S proteins to the receptors on the host cells lead to fusion of host cell and virus membranes, subsequently replicase will be translated and the viral replicase complex will be assembled. This made S protein a potential target for vaccine development (Xia *et al*., 2020). A recently published bioinformatic human CoV comparison only covered four human CoV (Niu *et al*., 2023), namely SARS-CoV-2, HCoV-HKU1, MERS as well as SARS. Therefore, there is a research gap as not all seven human CoV were included in the analysis. Only the four aforementioned human CoV were studied using the methods employed in this study, this creates a gap in knowledge in several viral aspects such as the antigenic epitope characteristics, similarity, physical and chemical properties as