© 2021 Journal of Pharmacy & Pharmacognosy Research, 9 (6), 766-779, 2021
ISSN 0719-4250
http://jppres.com/jppres

Original Article

Viroinformatics investigation of B-cell epitope conserved region in SARS-CoV-2 lineage B.1.1.7 isolates originated from Indonesia to develop vaccine candidate against COVID-19

[Investigación viroinformática de la región conservada del epítopo de células B en el linaje SARS-CoV-2 B.1.1.7 aislamientos originados en Indonesia para desarrollar una vacuna candidata contra COVID-19]

Arif N. M. Ansori[1,2#], Reviany V. Nidom[1,3*#], Muhammad K. J. Kusala[1,2], Setyarina Indrasari[1,3], Irine Normalina[1,4], Astria N. Nidom[1,3], Balqis Afifah[1,3], Kartika B. Sari[1,5], Nor L. Ramadhaniyah[1,5], Mohammad Y. Alamudi[1,3], Umi Cahyaningsih[6], Kuncoro P. Santoso[1,2], Heri Kuswanto[5], Chairul A. Nidom[1,2,3*]

[1]Coronavirus and Vaccine Formulation Research Group, Professor Nidom Foundation, Surabaya, Indonesia.
[2]Faculty of Veterinary Medicine, Universitas Airlangga, Surabaya, Indonesia.
[3]Riset AIRC Indonesia, Surabaya, Indonesia.
[4]Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia.
[5]Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia.
[6]Faculty of Veterinary Medicine, IPB University, Bogor, Indonesia. #Both authors contributed equally.
*E-mail: reviany@pnfinstitute.org, nidomca@pnfinstitute.org, nidomca@fkh.unair.ac.id

Abstract

Context: SARS-CoV-2, a member of family *Coronaviridae* and the causative agent of COVID-19, is a virus which is transmitted to human and other mammals.

Aims: To analyze the B-cell epitope conserved region and viroinformatics-based study of the SARS-CoV-2 lineage from Indonesian B.1.1.7 isolates to invent a vaccine nominee for overcoming COVID-19.

Methods: The sequences of seven Indonesian B.1.1.7 isolates, Wuhan-Hu-1 isolate, and WIV04 isolate were extracted from the GISAID EpiCoV and GenBank, NCBI. MEGA X was employed to understand the transformations of amino acid in the S protein and to develop a molecular phylogenetic tree. The IEDB was implemented to reveal the linear B-cell epitopes. In addition, PEP-FOLD3 web server was utilized to perform peptide modeling, while docking was performed using PatchDock, FireDock, and the PyMOL software. Moreover, *in silico* cloning was developed by using SnapGene v.3.2.1 software.

Results: In this study, the changes of amino acid in all seven Indonesian B.1.1.7 isolates were uncovered. Furthermore, various peptides based on the B-cell epitope prediction, allergenicity prediction, toxicity prediction from S protein to generate a vaccine contrary to SARS-CoV-2 were identified. Furthermore, the development of *in silico* cloning using pET plasmid was successfully achieved.

Conclusions: This study exhibits the transformations of amino acid in Indonesian B.1.1.7 isolates, and proposes four peptides ("LTPGDSSSGWTAG", "VRQIAPGQTGKIAD", "ILPDPSKPSKRS", and "KNHTSPDVDLG") from S protein as the candidate for a peptide-based vaccine. However, further advance trials such as *in vitro* and *in vivo* testing are involved for validation.

Keywords: COVID-19; SARS-CoV-2; vaccine design; viroinformatics.

Resumen

Contexto: SARS-CoV-2, un miembro de la familia *Coronaviridae* y el agente causante de COVID-19, es un virus que se transmite a humanos y otros mamíferos.

Objetivos: Analizar la región conservada del epítopo de células B y el estudio basado en viroinformática del linaje SARS-CoV-2 de los aislados B.1.1.7 de Indonesia para inventar una vacuna candidata para superar COVID-19.

Métodos: Las secuencias de siete aislamientos B.1.1.7 indonesios, el aislado Wuhan-Hu-1 y el aislado WIV04 se extrajeron de GISAID EpiCoV y GenBank, NCBI. Se empleó MEGA X para comprender las transformaciones de aminoácidos en la proteína S y para desarrollar un árbol filogenético molecular. El IEDB se implementó para revelar los epítopos de células B lineales. Además, se utilizó el servidor web PEP-FOLD3 para realizar el modelado de péptidos, mientras que el acoplamiento se realizó mediante PatchDock, FireDock y el software PyMOL. Además, la clonación *in silico* se desarrolló utilizando el software SnapGene v.3.2.1.

Resultados: En este estudio, se descubrieron los cambios de aminoácidos en los siete aislamientos de B.1.1.7 de Indonesia. Además, se identificaron varios péptidos basados en la predicción del epítopo de células B, la predicción de la alergenicidad, la predicción de la toxicidad de la proteína S para generar una vacuna contraria al SARS-CoV-2. Además, se logró con éxito el desarrollo de la clonación *in silico* utilizando el plásmido pET.

Conclusiones: Este estudio exhibe las transformaciones de aminoácidos en aislados B.1.1.7 de Indonesiay propone cuatro péptidos ("LTPGDSSSGWTAG", "VRQIAPGQTGKIAD", "ILPDPSKPSKRS" y "KNHTSPDVDLG") de la proteína S como candidato para una vacuna basada en péptidos. Sin embargo, para la validación se requieren más ensayos *in vitro* e *in vivo*.

Palabras Clave: COVID-19; SARS-CoV-2; vaccine design; viroinformatics.

ARTICLE INFO
Received: May 10, 2021.
Received in revised form: June 3, 2021.
Accepted: June 5, 2021.
Available Online: June 13, 2021.

INTRODUCTION

SARS-CoV-2 ranked as the seventh for the coronavirus (CoV) which has transversed the barrier between species and been able to infect human. The virus was firstly identified in China in 2019 and appeared sporadically all over China and many other nations worldwide (Harapan et al., 2020; Rodriguez-Morales et al., 2020; Wu et al., 2020). As the virus spreads massively and quickly, it was confirmed and revealed publicly by WHO in March 2020 that this contagious widespread has become a pandemic. The global health and economy have suffered by this COVID-19 abrupt contagion and rapid deployment (Astuti and Ysrafil, 2020; Huang et al., 2020). This global issue has required revolutionary scientific progression for deep investigations on SARS-CoV-2 concerning its features, mechanism of transmission, and clinical aspects for the sole purpose of ending this pandemic's catastrophic impacts (Awadasseid et al., 2021; Ganesh et al., 2021). Sadly, this virus has costed about 170 million people getting infected globally with more than 3.5 million global deaths. To be specific, Indonesia has about 1.8 million cases and around 50 thousand deaths. John Hopkins University has initiated the global data pooling with actual time updated online to pursue the real time cases of COVID-19 in numbers and geographical disbursement in which these above data are derived from (Dong et al., 2020).

The coronavirus family is composed of four different genera: *Alpha-*, *Gamma-*, *Beta-*, and *Deltacoronavirus*. *Alphacoronavirus* and *Betacoronavirus* infect animals and humans, whereas *Deltacoronavirus* and *Gammacoronavirus* infect only animals (Raoult et al., 2020). Similar to SARS-CoV-1 and MERS-CoV, SARS-CoV-2 also is part of *Coronaviridae* (*Betacoronavirus*), a family that has formerly driven the previous two epidemics (Shereen et al., 2020). Referenced in some data pools, the genome of SARS-CoV-2 has been recognized with around 30 kb nucleotides in a single-stranded positive-sense RNA (NC_045512.2 - GenBank, NCBI and EPI_ISL_402124 - GISAID EpiCoV) and the similarity shared with the genome of previous variant, SARS-CoV, is quite high, around 79.5% shared similarity (Hu et al., 2020). There are four structural proteins which conceal this genome, they are spike (S), nucleocapsid (N), membrane (M), and envelope (E) (Li et al., 2020; Ou et al., 2020; Zeng et al., 2020). In addition, several researches have informed that SARS-CoV-2 ties up ACE2 receptor on the host cell. Furthermore, the receptor contributes greatly in viral entry system and closely related to pathogenicity (Harapan et al., 2020).

Various investigation is demanded to acquire molecular epidemiological data of the novel virus in Indonesia (Nidom et al., 2021). Research on molecular epidemiology is a vital instrument for observing new emerging viruses. There is an urgency to establish further studies in molecular epidemiology to comprehend the probable consequences of the disease (Nidom et al., 2020a; 2021). Currently, a novel SARS-CoV-2 lineage was detected on September 2020, called B.1.1.7. A variant of this virus was initially discovered in England among a large number of new cases. Then, this variant becomes dominant in December 2020 in Southeastern England. B.1.1.7 variant spreads globally and rapidly including in Indonesia and as predicted in the early of March 2021 (Duong, 2021). This virus variant raises concerns recognized by neutralizing the antibodies, which may be affected leading to non-maximum vaccine efficacy.

In addition, our previous study reported the four structural proteins as vaccine candidate's development and revealed the antibody-dependent enhancement (ADE) phenomenon based on the S protein gene sequences (Ansori et al., 2020; Nidom et al., 2020b; Normalina et al., 2020). The efforts to produce vaccines globally in resistance to SARS-CoV-2 by scientists with the vaccines derived from its proteins turned out to be one of the most progressive kinds and the driving force of this study is the private sectors (Callaway, 2020). Although various countries already ongoing for the COVID-19 vaccination program (Nidom et al., 2021), currently there are no approved drugs against the virus (Harisna et al., 2021). A recent study revea-

Table 1. Indonesian B.1.1.7 isolates extracted from the database (GISAID EpiCoV).

Virus Name	Accession ID	Origin	Specimen source	Sequencing technology
JK-NIHRD-MI2101540	EPI_ISL_1118931	Jakarta, Indonesia	Nasopharyngeal and oropharyngeal swab	Illumina MiSeq
JK-NIHRD-MI2101673	EPI_ISL_1118933	Jakarta, Indonesia	Nasopharyngeal and oropharyngeal swab	Illumina MiSeq
JK-NIHRD-MI2101960	EPI_ISL_1169046	Jakarta, Indonesia	Nasopharyngeal swab	Illumina MiSeq
SS-NIHRD-WGS00427	EPI_ISL_1169047	South Sumatra, Indonesia	Nasopharyngeal swab	Illumina MiSeq
KS-NIHRD-WGS00915	EPI_ISL_1169048	South Kalimantan, Indonesia	Nasopharyngeal swab	Illumina MiSeq
SU-NIHRD-WGS01098	EPI_ISL_1169049	North Sumatra, Indonesia	Nasopharyngeal swab	Illumina MiSeq
JK-NIHRD-MI211743	EPI_ISL_1239143	Jakarta, Indonesia	Nasopharyngeal and oropharyngeal swab	Illumina MiSeq

led that this virus advances to a point that will ultimately trigger running from the vaccination program worldwide (Wang et al., 2021). This study intended to analyze the B-cell epitope conserved region and viroinformatics-based study of SARS-CoV-2 lineage B.1.1.7 isolates originated from Indonesia to mature a vaccine nominee in resistance to COVID-19.

MATERIAL AND METHODS

Indonesian B.1.1.7 SARS-CoV-2 isolates retrieval

It was revealed that the Indonesian B.1.1.7 isolates S protein gene via GISAID EpiCoV database (https://www.gisaid.org/) as March 18, 2021 (Table 1). We used two viruses as reference, called Wuhan-Hu-1 derived from GenBank, NCBI (https://www.ncbi.nlm.nih.gov/sars-cov-2/) and WIV04, available in GISAID EpiCoV database (https://www.gisaid.org/). Wuhan-Hu-1 isolate was submitted by the Shanghai Public Health Clinical Center & School of Public Health, Fudan University, Shanghai, China. Meanwhile, WIV04 isolate was generated from a female retailer at Huanan Seafood Wholesale Market and submitted also by the Wuhan Institute of Virology, Chinese Academy of Sciences, China.

Viroinformatics investigation of amino acid changes, phylogenetic analysis, and S protein modeling

The changes of amino acid changes and ADE sequences of the Indonesian B.1.1.7 isolates were analyzed by using CoVsurver, GISAID EpiCoV (https://www.gisaid.org/) and MEGA X software (Pennsylvania State University, USA). Next, the visualization of the results in amino acid changes with heatmap data by utilizing GraphPad Prism software (GraphPad Software, Inc., California, USA). In this study, the researchers constructed molecular phylogenetic designing and tree visualization by employing MEGA X software (Pennsylvania State University, USA) on the S protein of the Indonesian B.1.1.7 isolates and other CoVs isolates globally with the ML approach. The molecular phylogenetic construction was deduced by using 1000 bootstrapped input datasets and cross-referenced with the Tamura–Nei substitution model (Kumar et al., 2018; Ansori et al., 2020). Then, the S protein modeling was generated by applying SWISS-MODEL (https://swissmodel.expasy.org/), a protein structure homology-modelling web server (Waterhouse et al., 2018).

Predictions of B-cell epitope, antigenicity, allergenicity, toxicity, and physiochemistry

B-cell epitopes prediction of the S protein of the Indonesian B.1.1.7 isolates was done by applying the IEDB (https://tools.iedb.org/bcell/) using default thresholds (Adianingsih and Kharisma, 2019; Dhanda et al., 2019). Previously, the conserved region of the S protein was also identified by the researchers in the previous study. Next, the predicted peptides to the VaxiJen v2.0 web server (https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) was submitted using the default threshold to resolve whether the predicted epitopes could be expected as the protective antigens that would establish an immune response (Ansori et al., 2020). In this investigation, an extensive analysis of the allergenicity prediction in the predicted peptides was conducted using AllerTOP web server (https://www.ddg-pharmfac.net/AllerTOP/) with default settings. The expected peptides were submitted to this web server as demonstrated by Abraham Peele et al. (2020). Then, the protective non-toxic antigens were projected by performing ToxinPred web server (https://crdd.osdd.net/raghava/toxinpred/). The standard thresholds used were as reported by Gupta et al. (2013). Next, physiochemical prediction of the peptides was done by applying ProtParam web server (https://web.expasy.org/protparam/) as performed by Abraham Peele et al. (2020).

Peptide construction, molecular docking, and refinement process

PEP-FOLD 3.5 web server was utilized to look for the epitopes 3D structural findings (https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/). PEP-FOLD is a *de novo* approach intended to determine the peptide structures from the sequences of amino acid (Joshi et al., 2020). In this research, crystal structure of the human B-cell receptor (BCR) Fab fragment (2.90 Å) was retrieved, which was deposited in Protein Data Bank (PDB; https://www.rcsb.org/) (PDB ID: 5IFH) with the total structure weight about 46.98 kDa according to Minici et al. (2017). Next, the visualization of the structure was performed by applying PyMOL software v2.4.1 (Schrödinger, Inc, USA) (Luqman et al., 2020) and removing the water molecule and other molecules for preparing the protein before further molecular docking analysis (Harisna et al., 2021). Peptide-protein molecular docking and refinement processes were conducted by applying PatchDock web server (https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php) and FireDock web server (https://bioinfo3d.cs.tau.ac.il/FireDock/php.php) to reveal global energy score (Maiti and Banerjee, 2021).

Immune simulation and *in silico* cloning

The immune simulation was tested by employing the C-ImmSim web server (https://kraken.iac.rm.cnr.it/C-IMMSIM/). The simulation volume and steps were designated at 1000 based on the study by Abraham Peele et al. (2020). We used J-CAT tool (https://www.jcat.de/) to analyze codon optimization of vaccine constructs by using *E. coli* as a source organism (Abraham Peele et al., 2020). Here, this investigation appointed the expression vector of pET-28a(+) for cloning, and generated its nucleotides sequences from the Addgene vector database (https://www.addgene.org/vector-database/) (Kamens, 2015). Then, in seeking the *in silico* cloning of peptide-based vaccine component was performed using SnapGene v3.2.1 software (GSL Biotech LLC, California, USA) (Abraham Peele et al., 2020; Le Bert et al. 2020).

RESULTS

Viroinformatics investigation of amino acid changes, phylogenetic analysis, and S protein modeling

This research demonstrated all the amino acid changes in full-length genome of the Indonesian B.1.1.7 isolates (Fig. 1A, Table 2, and Annex 1) and visualized the three-dimensional S protein amino acid changes (Fig. 1B). This investigation identified that various changes of amino acid did occurred in S protein of Indonesian B.1.1.7 isolates, such as

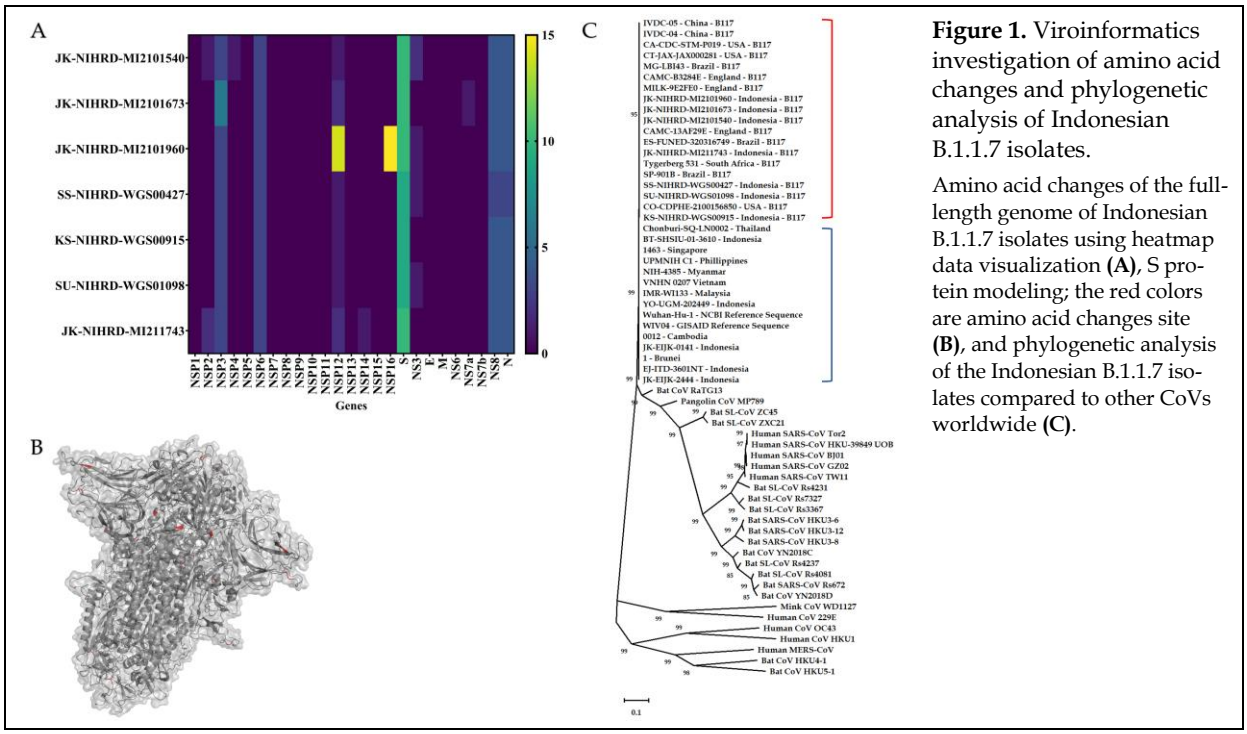

Figure 1. Viroinformatics investigation of amino acid changes and phylogenetic analysis of Indonesian B.1.1.7 isolates.

Amino acid changes of the full-length genome of Indonesian B.1.1.7 isolates using heatmap data visualization **(A)**, S protein modeling; the red colors are amino acid changes site **(B)**, and phylogenetic analysis of the Indonesian B.1.1.7 isolates compared to other CoVs worldwide **(C)**.

Table 2. Viroinformatics investigation of amino acid changes from Indonesian B.1.1.7 isolates S protein.

Virus name	Amino acid changes
JK-NIHRD-MI2101540	V143C, Y144F, T286I, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H
JK-NIHRD-MI2101673	H69 (deletion), V70 (deletion), Y144 (deletion), N501Y, A570D, D614G, P681H, T716I, S982A, D1118H
JK-NIHRD-MI2101960	H69 (deletion), V70 (deletion), Y144 (deletion), N501Y, A570D, D614G, P681H, T716I, S982A, D1118H
SS-NIHRD-WGS00427	H69 (deletion), V70 (deletion), N501Y, A570D, D614G, P681H, T716I, S982A, D1118H
KS-NIHRD-WGS00915	H69 (deletion), V70 (deletion), N501Y, A570D, D614G, P681H, T716I, S982A, D1118H
SU-NIHRD-WGS01098	H69 (deletion), V70 (deletion), N501Y, A570D, D614G, P681H, T716I, S982A, D1118H

H69 (deletion), V70 (deletion), V143C, Y144 (deletion) or Y144F, T286I, N501Y, A570D, T716I, D614G, P681H, S982A, and D1118H. In addition, it revealed the ADE sequences of all isolates had the same pattern as the previous study, namely "^{611}LYQDVNC^{617}". Moreover, it successfully constructed the phylogenetic examination of the Indonesian SARS-CoV-2 isolates based on the S protein gene (Fig. 1C).

Predictions of B-cell epitope, antigenicity, allergenicity, toxicity, and physiochemistry

This study identified the B-cell epitope prediction, antigenicity prediction, allergenicity prediction, and toxicity prediction analysis that are shown in Table 3 and Fig. 2. Moreover, it also established the physiochemical prediction for the expected peptides, such as grand average of hydropathicity (GRAVY), aliphatic index, theorical pI, molecular weight, and instability index (Table 4).

Peptide construction, peptide-protein molecular docking, and refinement process

PEP-FOLD 3.5 web server were employed to create the three-dimensional epitope structures (Fig. 3). This study revealed BCR (resolution 2.90 Å) from PDB (PDB ID: 5IFH). The heavy chain and light chain of the BCR were distinguished by coloring selection (Fig. 3). In addition, the results of molecular docking and refinement process were derived by performing PatchDock web server and FireDock web server (Fig. 4 and Table 5). Furthermore, this study also uncovered various variables, such as hydrogen bonds, atomic contact energy, attractive and repulsive Van der Walls, and global energy. The strong binding energy was exhibited by Pep3, Pep2, Pep1, and Pep4, respectively.

Table 3. B-cell epitopes and other prediction analyses in the Indonesian B.1.1.7 isolates.

No	Peptide sequence	Length and position	Antigenicity	Allergenicity	Toxicity
1	RTQLPPAYTNS	11 (21-31)	Yes	Yes	Non
2	SGTNGTKRFDN	11 (71-81)	Yes	Yes	Non
3	LTPGDSSSGWTAG	13 (249-261)	Yes	Non	Non
4	VRQIAPGQTGKIAD	14 (407-420)	Yes	Non	Non
5	NNLDSKVGG	9 (439-447)	Yes	Yes	Non
6	YQAGSTPCNGV	11 (473-483)	Non	Non	Non
7	YGFQPTNGVGYQ	12 (495-506)	Yes	Yes	Non
8	TVCGPKKSTN	10 (523-532)	Non	Yes	Non
9	QTQTNSPRRARSV	13 (675-687)	Non	Non	Non
10	IYKTPPIKDF	10 (788-797)	Non	Yes	Non
11	ILPDPSKPSKRS	12 (805-816)	Yes	Non	Non
12	PAQEKNFTT	9 (1069-1077)	Non	Yes	Non
13	VYDPLQPELDSF	12 (1137-1148)	Non	Yes	Non
14	KNHTSPDVDLG	11 (1157-1167)	Yes	Non	Non
15	FDEDDSEPVL	10 (1256-1265)	Non	Non	Non

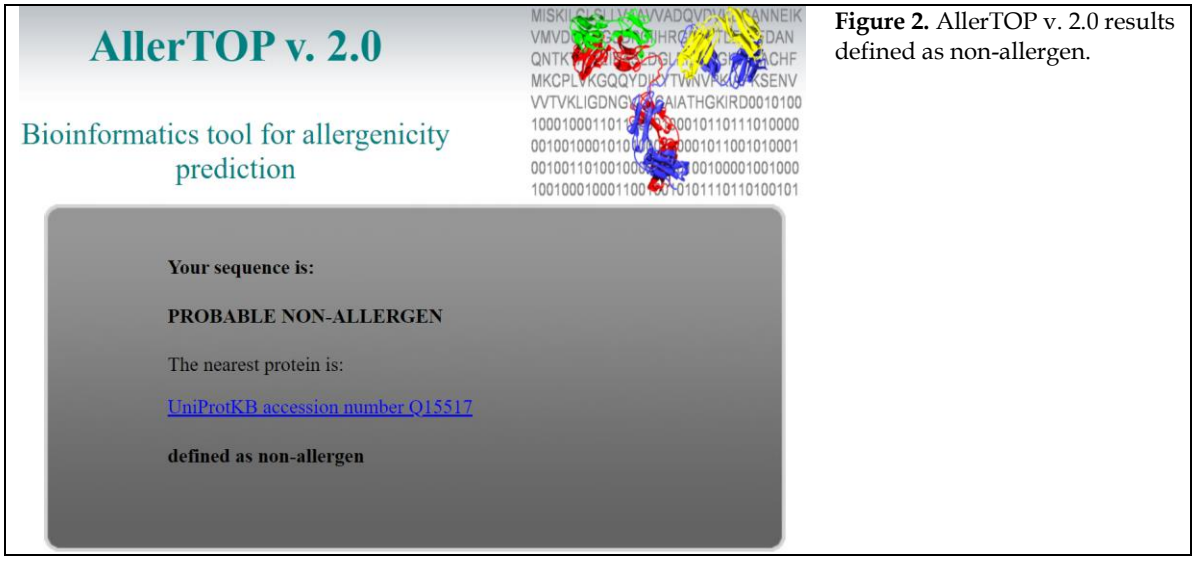

Figure 2. AllerTOP v. 2.0 results defined as non-allergen.

Table 4. The results of physiochemical prediction in peptides using ProtParam web server.

Name	Peptide sequence	Molecular weight	Theorical pI	Instability index	Aliphatic index	Grand average of hydropathicity (GRAVY)
Pep1	LTPGDSSSGWTAG	1235.27	3.80	51.61	37.69	-0.415
Pep2	VRQIAPGQTGKIAD	1453.66	8.72	12.54	90.71	-0.371
Pep3	ILPDPSKPSKRS	1324.54	9.99	124.58	65	-1.225
Pep4	KNHTSPDVDLG	1182.26	5.21	21.66	61.82	-1.191

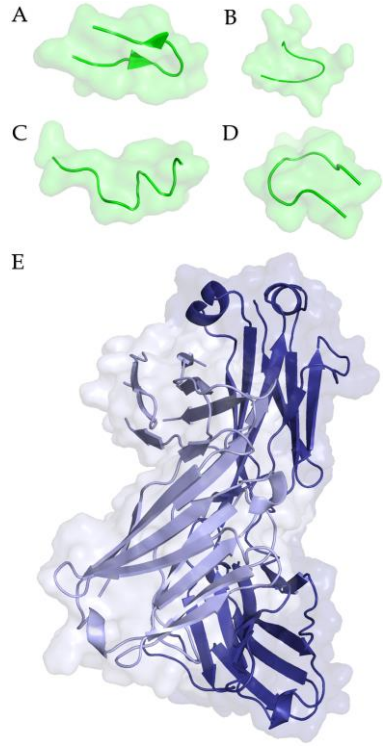

Figure 3. Visualization of peptides construction by PEP-FOLD 3.5 web server and B-cell receptor (BCR).

Pep1 **(A)**, Pep2 **(B)**, Pep3 **(C)**, and Pep4 **(D)**. BCR with PDB ID: 5IFH **(E)**; Light blue color is heavy chain and deep blue color is light chain of BCR.

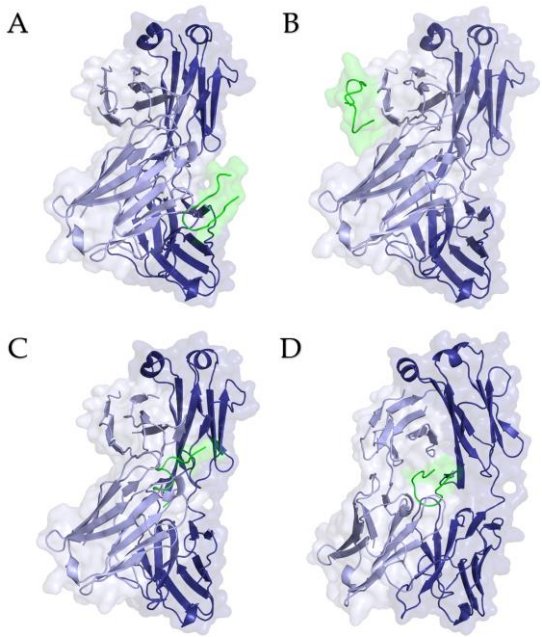

Figure 4. Visualization of peptide-protein molecular docking by PatchDock web server and FireDock web server.

Pep1-BCR **(A)**, Pep2-BCR **(B)**, Pep3-BCR **(C)**, and Pep4-BCR **(D)**. Light blue color is heavy chain and deep blue color is light chain of BCR.

Table 5. Peptide-protein interactions in detail after refinement process using FireDock web server.

Name	Global energy	Attractive van der Waals	Repulsive van der Waals	Atomic contact energy	Hydrogen bonds
Pep1	-42.40	-22.16	11.57	-10.43	-1.84
Pep2	-44.38	-25.45	12.83	-7.48	-0.65
Pep3	-44.78	-24.96	8.17	1.49	-4.32
Pep4	-36.56	-30.15	6.23	-2.94	-4.67

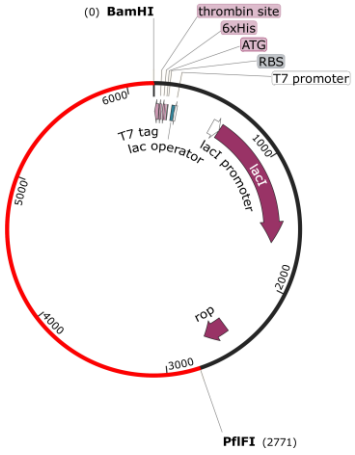

Figure 5. *In silico* cloning's schematic representation of vaccine candidate within pET28a(+) expression vector. The total size of the construct is 6180 bp and the red color is gene interest (S protein).

Immune simulation and *in silico* cloning

C-ImmSim web server demonstrated the responses of cells and models from the memory of immune cells whether it has any successful and effective immune or not by applying a procedure which improved its half-life. Furthermore, its simulation results established reliability with real immune responses. Moreover, the improving B-cell population was described by the expression of immunoglobulins, which produced a reduced concentration of antigen. Additionally, the pET28a(+) vector was employed to clone the construction of vaccine by performing SnapGene v3.2.1 software (Fig. 5).

DISCUSSION

Six types of coronaviruses were acknowledged to be causative agents which were able to transmit to humans at the end of 2019. SARS-CoV-2 was identified as the seventh and appeared in China (Gunadi et al. 2020). To date, it was stated by Johns Hopkins University, USA that there are about 140 million people getting transmitted with this novel virus worldwide (Dong et al. 2020). Previously, Turista et al. (2020) reported the COVID-19 pandemic in Indonesia. In line with this, Setiawaty et al. (2020) described the virological and clinical features of the first eleven COVID-19 cases in Indonesia. Furthermore, this study retrieved the first seven Indonesian B.1.1.7 isolates from the database (Table 1). These B.1.1.7 isolates were submitted by the National Institute of Health Research and Development, Republic of Indonesia and detected using nasopharyngeal swab and also nasopharyngeal and oropharyngeal swab methods. Therefore, this investigation conducted a viroinformatics investigation of B-cell epitope conserved region in SARS-CoV-2 lineage B.1.1.7 isolates originated from Indonesia.

Currently, many studies have focused on the S protein of the deadly virus (Watanabe et al., 2020; Weisblum et al., 2020; Zhang et al., 2020). However, the urgency in investigating S protein is very significant an extensive computational-based study (Normalina et al., 2020). This study in viroinformatics investigation has discovered amino acid changes of the full-length genome in Indonesian B.1.1.7 isolates (Fig. 1 and Table 2). Furthermore, this study also identified that various transformations of amino acid occurred in S protein of Indonesian B.1.1.7 isolates, such as H69 (deletion), V70 (deletion), V143C, Y144 (deletion) or Y144F, A570D, S982A, T286I, N501Y, D614G, P681H, T716I, and D1118H. In fact, many studies reported that these unprecedented variants of SARS-CoV-2 appeared in many countries, such as South Africa, UK, and Brazil (Grubaugh et al. 2021; Nonaka et al. 2021).

Consistently, a study showed that B.1.1.7 is known to be uncompromising to any counteracting efforts performed by most monoclonal antibodies against the N-terminal domain from the spike protein and relatively impenetrable to a few monoclonal antibodies against RBD of the novel

virus. However, it is less resistant to plasma from individuals who have recovered from COVID-19 or sera from individuals who have been vaccinated against the novel virus (Wang et al., 2021). In line with this, the mutation rates in RNA viruses are much higher than in most other microorganisms. Any surge of virulence and any heightened prospects of adaptative evolution can be triggered by a growth in mutation rate (Yu, 2020). Consequently, these features are able to encourage the possibility of zoonotic viral pathogens, which can accommodate human-to-human transmission and allows them to intensify their virulence (Nidom et al., 2021). Thus, the urgency for investigating the S protein from Indonesian B.1.1.7 isolates comes as the top priority.

Furthermore, scientists have proven that this novel virus gnome indeed globally experiences mutations. Previously, a genetic analysis was conducted by Sallam et al. (2021) and exhibited that among the novel virus sequences from the North Africa and Middle East many mutations turn up in the S protein. Moreover, for the evolution of viruses in nature, nucleotide substitution is believed to be one of the most important mechanisms. In addition, a study of SARS-CoV-2 outbreak in Uruguay conducted by Elizondo et al. (2021) has made significant contribution in leading improved understanding on the patterns of this novel virus and its regional pandemic features in Latin America. In addition, a study by Khailany et al. (2020) had successfully accomplished recovering 94 SARS-CoV-2 genomes and inspecting their molecular variations among them. Furthermore, Zhang et al. (2020) stated that the S protein mutations are associated with the virulence of the virus. Thus, this study is the first to report the analysis of first seven Indonesian B.1.1.7 isolates. These data might be a groundbreaking in supporting further researches in the establishment of recent biological aspects of SARS-CoV-2 in Indonesia. In addition, without any experimental data, our interpretation using the limited data of SARS-CoV-2 isolates in Indonesia will potentially benefit the continuous and forthcoming research. Le Page et al. (2021) stated that the research on S protein mutations was very fundamental considering its impacts on the development of global vaccination frameworks against this novel virus.

Under those above circumstances, COVID-19 vaccine development has started in many research centers and pharmaceutical industries following the announcement of SARS-CoV-2 agent and its full genome recognized. Recently, the available assemble data stated that COVID-19 vaccine candidates were grouped into the following types: protein-based, epitope, inactivated or live-attenuated virus, virus-like particle, nucleic acid-based, and viral vectors (Nidom et al., 2021). Today, more than one year after the prevalence of novel coronavirus, vaccine and antiviral products are still in progress due to the pandemic paradigm development with several medication options and vaccines are in clinical trials globally (Callaway, 2020). Furthermore, scientists considered traversing the new concepts and latest cultivation in each type of vaccine to formularize a potent vaccine against COVID-19. In the same manner, genomics has encouraged the revolutionized researches related to the vaccine development through the reverse vaccinology with its predominant capability in sequencing the genome completely from any deadly microorganism driving some *in silico* screenings to obtain the protective antigens prior to any further hypothesis testing researches.

Meanwhile, epitope prediction studies have been accomplished for some viruses, such as the Zika virus (Adianingsih and Kharisma, 2019). This current investigation developed a peptide-based vaccine using B-cell epitope prediction, toxicity prediction, allergenicity prediction, and antigenicity prediction. In addition, it also performed physiochemical prediction, peptide construction, molecular docking, immune simulation, and *in silico* cloning for the predicted peptides. Thus, it proposes "LTPGDSSSGWTAG", "VRQIAPGQTGKI AD", "ILPDPSKPSKRS", and "KNHTSPDVDLG" peptides isolated from S protein as candidate for a peptide-based vaccine to fight against to SARS-CoV-2. Moreover, the notable epitope prediction methods are significantly fundamental in some biotechnological and clinical applications, such as therapeutic antibody and vaccine initiation, or theoretical studies of immune systems.

Moreover, this study was multidisciplinary collaboration between Indonesian and Swiss researchers in various adjuvanted pandemic influenza formulation strategies which suggested that a liposome-based adjuvant containing the saponin or QS21 is one of the most important and promised adjuvant in the near future (Lemoine et al., 2021). On the contrary, commonly used in vaccines, the pharmacodynamics in the injected form of aluminum is not well-acknowledged, especially related to on how the variation of time frames will affect the accumulation and how determinants like genetics and environments will influence the viral clearance on detoxification process (McFarland et al., 2020). Briefly, the Professor Nidom Foundation (PNF) is a research and education organization in Indonesia that has extensive experience in the surveillance and characterization of various zoonotic viruses (Lemoine et al., 2021).

Additionally, elucidating the transmission routes, relationships, and origin of the causative agents from any emerging infections is crucial to understand the possible approach for interference and their biological actions. Many scientists have established the molecular phylogenetic analysis of SARS-CoV-2 with the previous coronaviruses by applying the specific or whole-genome sequences to comprehend the recombination events and an evolutionary chronicle (Ansori et al., 2020; Sallam et al., 2021). According to the recently available genome isolates of coronaviruses, the whole-genome phylogenetic tree designates that SARS-CoV-2 is closest to *Rhinolophus affinis* coronavirus RaTG13, followed by the one found in pangolin's coronavirus. Consistent with this, a group of researchers stated they accepted the idea that SARS-CoV-2 has been issued in bats and may have used pangolins as an intermediate host before transmitting it to humans. Moreover, the SARS-CoV-2 genome contributes its genome for around 80% to SARS-CoV (Li et al., 2020; Shereen et al., 2020).

Lastly, this study promoted the molecular phylogenetic tree and revealed the relationship between Indonesian B.1.1.7 isolates, other isolates, and other groups of coronaviruses originating from humans, pangolins, and bats. Here, this investigation reports the first molecular phylogenetic tree of seven Indonesian B.1.1.7 isolates based on the S protein gene. It also utilizes molecular phylogenetic analysis as the basic of virus research which includes taxonomy, evolution, and origin studies (Turista et al., 2020). It is principal to investigate the likelihood of SARS-CoV-2 intermediate hosts to understand and contain the transmission of COVID-19 (Gunadi et al., 2020). In addition, this study advises that further surveillance investigations should be carried out on various mammals in their natural environment, including pangolins and bats, especially in Asia, to control the chance of future zoonotic transmissions.

CONCLUSIONS

This study exhibits the transformations of amino acid in Indonesian B.1.1.7 isolates. Additionally, this study proposes four peptides ("LTPGDSSSGWTAG", "VRQIAPGQTGKIAD", "ILPDPSKPSKRS", and "KNHTSPDVDLG") from S protein as the candidate for a peptide-based vaccine. However, further advance trials such as *in vitro* and *in vivo* testing are involved for validation.

CONFLICT OF INTEREST

The authors declare no conflicts of interests.

ACKNOWLEDGMENTS

This investigation was fully facilitated by the Professor Nidom Foundation, Surabaya, Indonesia (Grant Number: 005/PNF-RF/07/2020) and PMDSU Scholarship Program (Batch III), Directorate General of Higher Education, Ministry of Education and Culture of the Republic of Indonesia (Grant Number: 27/E1/KPT/2020). Authors acknowledge the participants for their great works from the originating and submitting laboratories of GISAID database on which the isolates were generated from. Additionally, authors thank Work of Words by Sartika, Indonesia for editing the manuscript.

REFERENCES

Abraham Peele K, Srihansa T, Krupanidhi S, Ayyagari VS, Venkateswarulu TC (2020) Design of multi-epitope vaccine candidate against SARS-CoV-2: A *in silico* study. J Biomol Struct Dyn 2020: 1-9.

Adianingsih OR, Kharisma VD (2019) Study of B cell epitope conserved region of the Zika virus envelope glycoprotein to develop multi-strain vaccine. J Appl Pharm Sci 9: 98-103.

Ansori ANM, Kusala MKJ, Normalina I, Indrasari S, Alamudi MY, Nidom RV, Santoso KP, Rachmawati K, Nidom CA (2020) Immunoinformatic investigation of three structural protein genes in Indonesian SARS-CoV-2 isolates. Sys Rev Pharm 11(7): 422-434.

Astuti I, Ysrafil (2020) Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response. Diabetes Metab Syndr 14(4): 407-412.

Awadasseid A, Wu Y, Tanaka Y, Zhang W (2021) SARS-CoV-2 variants evolved during the early stage of the pandemic and effects of mutations on adaptation in Wuhan populations. Int J Biol Sci 17(1): 97-106.

Callaway E. (2020) The race for coronavirus vaccines: A graphical guide. Nature 580: 576-577.

Dhanda SK, Mahajan S, Paul S, Yan Z, Kim H, Jespersen MC, Jurtz V, Andreatta M, Greenbaum JA, Marcatili P, Sette A, Nielsen M, Peters B (2019) IEDB-AR: Immune epitope database-analysis resource in 2019. Nucleic Acids Res 47(W1): W502-W506.

Dong E, Du H, Gardner L (2020) An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis S1473-3099(20): 30120-30121.

Duong D (2021) What's important to know about the new COVID-19 variants? CMAJ 193(4): E141-E142.

Elizondo V, Harkins GW, Mabvakure B, Smidt S, Zappile P, Marier C, Maurano MT, Perez V, Mazza N, Beloso C, Ifran S, Fernandez M, Santini A, Perez V, Estevez V, Nin M, Manrique G, Perez L, Ross F, Boschi S, Zubillaga MN, Balleste R, Dellicour S, Heguy A, Duerr R (2021) SARS-CoV-2 genomic characterization and clinical manifestation of the COVID-19 outbreak in Uruguay. Emerg Microbes Infect 10(1): 51-65.

Ganesh B, Rajakumar T, Malathi M, Manikandan N, Nagaraj J, Santhakumar A, Elangovan A, Malik YS (2021) Epidemiology and pathobiology of SARS-CoV-2 (COVID-19) in comparison with SARS, MERS: An updated overview of current knowledge and future perspectives. Clin Epidemiol Glob Health 10: 100694.

Grubaugh ND, Hodcroft EB, Fauver JR, Phelan AL, Cevik M (2021) Public health actions to control new SARS-CoV-2 variants. Cell 184(5): 1127-1132.

Gunadi, Wibawa H, Marcellus, Hakim MS, Daniwijaya EW, Rizki LP, Supriyati E, Nugrahaningsih DAA, Afiahayati, Siswanto, Iskandar K, Anggorowati N, Kalim AS, Puspitarani DA, Athollah K, Arguni E, Nuryastuti T, Wibawa T (2020) Full-length genome characterization and phylogenetic analysis of SARS-CoV-2 virus strains from Yogyakarta and Central Java, Indonesia. PeerJ 8: e10575.

Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R; Open Source Drug Discovery Consortium, Raghava GP (2013) *In silico* approach for predicting toxicity of peptides and proteins. PLoS One 8(9): e73957.

Harapan H, Itoh N, Yufika A, Winardi W, Keam S, Te H, Megawati D, Hayati Z, Wagner AL, Mudatsir M (2020) Coronavirus disease 2019 (COVID-19): A literature review. J Infect Public Health 13(5): 667-673.

Harisna AH, Nurdiansyah R, Syaifie PH, Nugroho DW, Saputro KE, Firdayani, Prakoso CD, Rochman NT, Maulana NN, Noviyanto A, Mardliyati E (2021) *In silico* investigation of potential inhibitors to main protease and spike protein of SARS-CoV-2 in propolis. Biochem Biophys Rep 26: 100969.

Hu T, Liu Y, Zhao M, Zhuang Q, Xu L, He Q (2020) A comparison of COVID-19, SARS and MERS. PeerJ 8: e9725.

Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395(10223): 497-506.

Joshi A, Joshi BC, Mannan MA, Kaushik V (2020) Epitope based vaccine prediction for SARS-COV-2 by deploying immuno-informatics approach. Inform Med Unlocked. 19: 100338.

Kamens J (2015) The Addgene repository: An international nonprofit plasmid and data resource. Nucleic Acids Res 43: D1152-D1157.

Khailany RA, Safdar M, Ozaslan M (2020) Genomic characterization of a novel SARS-CoV-2. Gene Rep 19: 100682.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6): 1547-1549.

Le Bert N, Tan AT, Kunasegaran K, Tham CYL, Hafezi M, Chia A, Chng MHY, Lin M, Tan N, Linster M, Chia WN, Chen MI, Wang LF, Ooi EE, Kalimuddin S, Tambyah PA, Low JG, Tan YJ, Bertoletti A (2020) SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. Nature 584(7821): 457-462.

Le Page M (2021) Threats from new variants. New Sci 249(3316): 8-9.

Lemoine CH, Nidom RV, Ventura R, Indrasari S, Normalina I, Santoso KP, Derouet F, Barnier-Quer C, Borchard G, Collin N, Nidom CA (2021) Better pandemic influenza preparedness through adjuvant technology transfer: Challenges and lessons learned. Vaccines 9(5): 461.

Li H, Liu SM, Yu XH, Tang SL, Tang CK (2020) Coronavirus disease 2019 (COVID-19): Current status and future perspectives. Int J Antimicrob Agents 55(5): 105951.

Luqman A, Kharisma VD, Ruiz RA, Götz F (2020) *In silico* and *in vitro* study of trace amines (TA) and dopamine (DOP) interaction with human alpha 1-adrenergic receptor and the bacterial adrenergic receptor QseC. Cell Physiol Biochem 54(5): 888-898.

Maiti S, Banerjee A (2021) Epigallocatechin gallate and theaflavin gallate interaction in SARS-CoV-2 spike-

protein central channel with reference to the hydroxychloroquine interaction: Bioinformatics and molecular docking study. Drug Dev Res 82(1): 86-96.

McFarland G, La Joie E, Thomas P, Lyons-Weiler J (2020) Acute exposure and chronic retention of aluminum in three vaccine schedules and effects of genetic and environmental variation. J Trace Elem Med Biol 58: 126444.

Minici C, Gounari M, Übelhart R, Scarfò L, Dühren-von Minden M, Schneider D, Tasdogan A, Alkhatib A, Agathangelidis A, Ntoufa S, Chiorazzi N, Jumaa H, Stamatopoulos K, Ghia P, Degano M (2017) Distinct homotypic B-cell receptor interactions shape the outcome of chronic lymphocytic leukaemia. Nat Commun 8: 15746.

Nidom RV, Ansori ANM, Indrasari S, Normalina I, Kusala MKJ, Saefuddin A, Nidom CA (2020a) Recent updates on COVID-19 vaccine platforms and its immunological aspects: A review. Sys Rev Pharm 11(10): 807-818.

Nidom RV, Indrasari S, Normalina I, Kusala MKJ, Ansori ANM, Nidom CA (2020b) Investigation of the D614G mutation and antibody-dependent enhancement sequences in Indonesian SARS-CoV-2 isolates and comparison to Southeast Asian isolates. Sys Rev Pharm 11(8): 203-213.

Nidom RV, Indrasari S, Normalina I, Nidom AN, Afifah B, Dewi L, Putra AK, Ansori ANM, Kusala MKJ, Alamudi MY, Nidom CA (2021) An updated investigation prior to COVID-19 vaccination program in Indonesia: Full-length genome mutation analysis of SARS-CoV-2. bioRxiv 2021.01.26.426655.

Nonaka CKV, Franco MM, Gräf T, de Lorenzo Barcia CA, de Ávila Mendonça RN, de Sousa KAF, Neiva LMC, Fosenca V, Mendes AVA, de Aguiar RS, Giovanetti M, de Freitas Souza BS (2021) Genomic evidence of SARS-CoV-2 reinfection involving E484K spike mutation, Brazil. Emerg Infect Dis 27(5): 1522-1524.

Normalina I, Indrasari S, Nidom RV, Kusala MKJ, Alamudi MY, Santoso KP, Rachmawati K, Nidom CA (2020) Characterization of the spike glycoprotein and construction of an epitope-based vaccine candidate against Indonesian SARS-CoV-2: *In silico* study. Sys Rev Pharm 11(7): 404-413.

Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, Guo L, Guo R, Chen T, Hu J, Xiang Z, Mu Z, Chen X, Chen J, Hu K, Jin Q, Wang J, Qian Z (2020) Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. Nat Commun 11(1): 1620.

Raoult D, Zumla A, Locatelli F, Ippolito G, Kroemer G (2020) Coronavirus infections: Epidemiological, clinical and immunological features and hypotheses. Cell Stress 4(4): 66-75.

Rodriguez-Morales AJ, Cardona-Ospina JA, Gutiérrez-Ocampo E, Villamizar-Peña R, Holguin-Rivera Y, Escalera-Antezana JP, Alvarado-Arnez LE, Bonilla-Aldana DK, Franco-Paredes C, Henao-Martinez AF, Paniz-Mondolfi A, Lagos-Grisales GJ, Ramírez-Vallejo E, Suárez JA, Zambrano LI, Villamil-Gómez WE, Balbin-Ramon GJ, Rabaan AA, Harapan H, Dhama K, Nishiura H, Kataoka H, Ahmad T, Sah R; Latin American Network of Coronavirus Disease 2019-COVID-19 Research (LANCOVID-19) (2020) Clinical, laboratory and imaging features of COVID-19: A systematic review and meta-analysis. Travel Med Infect Dis 34: 101623.

Sallam M, Ababneh NA, Dababseh D, Bakri FG, Mahafzah A (2021) Temporal increase in D614G mutation of SARS-CoV-2 in the Middle East and North Africa. Heliyon 7(1): e06035.

Setiawaty V, Kosasih H, Mardian Y, Ajis E, Prasetyowati EB, Siswanto, Karyana M, SARS-CoV-Reference Laboratory Ministry of Health Indonesia (2020) The identification of first COVID-19 cluster in Indonesia. Am J Trop Med Hyg 103(6): 2339-2342.

Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R (2020) COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. J Adv Res 24: 91-98.

Turista DDR, Islamy A, Kharisma VD, Ansori ANM (2020) Distribution of COVID-19 and phylogenetic tree construction of SARS-CoV-2 in Indonesia. J Pure Appl Microbiol 14: 1035-1042.

Wang P, Nair MS, Liu L, Iketani S, Luo Y, Guo Y, Wang M, Yu J, Zhang B, Kwong PD, Graham BS, Mascola JR, Chang JY, Yin MT, Sobieszczyk M, Kyratsous CA, Shapiro L, Sheng Z, Huang Y, Ho DD (2021) Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. Nature 593(7857): 130-135.

Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M (2020) Site-specific glycan analysis of the SARS-CoV-2 spike. Science 369(6501): 330-333.

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T (2018) SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Res 46(W1): W296-W303.

Weisblum Y, Schmidt F, Zhang F, DaSilva J, Poston D, Lorenzi JC, Muecksch F, Rutkowska M, Hoffmann HH, Michailidis E, Gaebler C, Agudelo M, Cho A, Wang Z, Gazumyan A, Cipolla M, Luchsinger L, Hillyer CD, Caskey M, Robbiani DF, Rice CM, Nussenzweig MC, Hatziioannou T, Bieniasz PD (2020) Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. Elife 9: e61312.

Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, Meng J, Zhu Z, Zhang Z, Wang J, Sheng J, Quan L, Xia Z, Tan W, Cheng G, Jiang T (2020) Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. Cell Host Microbe 27(3): 325-328.

Yu J (2020) From mutation signature to molecular mechanism in the RNA world: A case of SARS-CoV-2. Genomics Proteomics Bioinformatics 30: S1672-0229(20)30092-9.

Zeng W, Liu G, Ma H, Zhao D, Yang Y, Liu M, Mohammed A, Zhao C, Yang Y, Xie J, Ding C, Ma X, Weng J, Gao Y, He H, Jin T (2020) Biochemical characterization of SARS-CoV-2 nucleocapsid protein. Biochem Biophys Res Commun 527(3): 618-623.

Zhang L, Jackson CB, Mou H, Ojha A, Peng H, Quinlan BD, Rangarajan ES, Pan A, Vanderheiden A, Suthar MS, Li W, Izard T, Rader C, Farzan M, Choe H (2020) SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. Nat Commun 11(1): 6013.

AUTHOR CONTRIBUTION:

Contribution	Ansori ANM	Nidom RV	Kusala MKJ	Indrasari S	Normalina I	Nidom AN	Afifah B	Sari KB	Ramadhaniyah NL	Alamudi MY	Cahyaningsih U	Santoso KP	Kuswanto H	Nidom CA
Concepts or ideas	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Design	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Definition of intellectual content	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Literature search	x	x	x	x	x	x	x	x	x					
Experimental studies	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Data acquisition	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Data analysis	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Statistical analysis	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Manuscript preparation	x	x	x	x	x	x	x	x	x					
Manuscript editing	x	x	x	x	x									
Manuscript review	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Citation Format: Ansori ANM, Nidom RV, Kusala MKJ, Indrasari S, Normalina I, Nidom AN, Afifah B, Sari KB, Ramadhaniyah NL, Alamudi MY, Cahyaningsih U, Santoso KP, Kuswanto H, Nidom CA (2021) Viroinformatics investigation of B-cell epitope conserved region in SARS-CoV-2 lineage B.1.1.7 isolates originated from Indonesia to develop vaccine candidate against COVID-19. J Pharm Pharmacogn Res 9(6): 766–779.

Annex 1. The amino acid changes in the full-length genome of the Indonesian B.1.1.7 isolates.

No.	Genes	JK-NIHRD-MI2101540	JK-NIHRD-MI2101673	JK-NIHRD-MI2101960	SS-NIHRD-WGS00427	KS-NIHRD-WGS00915	SU-NIHRD-WGS01098	JK-NIHRD-MI211743
1	NSP1	0	0	0	0	0	0	0
2	NSP2	1	0	0	0	0	0	2
3	NSP3	3	6	3	3	3	3	3
4	NSP4	1	0	0	0	0	0	0
5	NSP5	0	0	0	0	0	0	0
6	NSP6	3	3	3	3	3	3	3
7	NSP7	0	0	0	0	0	0	0
8	NSP8	0	0	0	0	0	0	0
9	NSP9	0	0	0	0	0	0	0
10	NSP10	0	0	0	0	0	0	0
11	NSP11	0	0	0	0	0	0	0
12	NSP12	1	2	14	1	1	1	2
13	NSP13	0	0	0	0	0	0	0
14	NSP14	0	0	0	0	0	0	1
15	NSP15	0	0	0	0	0	0	0
16	NSP16	0	0	15	0	0	0	0
17	S	10	10	10	9	9	9	10
18	NS3	2	0	1	1	0	1	0
19	E	0	0	0	0	0	0	0
20	M	0	0	0	0	0	0	0
21	NS6	0	0	0	0	0	0	0
22	NS7a	0	1	0	0	0	0	0
23	NS7b	0	0	0	0	0	0	0
24	NS8	4	4	4	3	4	4	4
25	N	4	4	4	3	4	4	4