International Journal of Research and Review
Vol. 10; Issue: 8; August 2023
Website: www.ijrrjournal.com
Original Research Article *E-ISSN: 2349-9788; P-ISSN: 2454-2237*

Analysis of RdRp Gene in SARS CoV-2 Variants

SM Rezvi[1,5], Desmawati Desmawati[2,5], Andani Eka Putra[3,5], Hirowati Ali[4,5]

[1]Master Program in Biomedical Sciences, [2]Department of Nutrition, [3]Department of Microbiology, [4]Department of Biochemistry, [5]Center for Diagnostic and Research on Infectious Diseases (PDRPI), Faculty of Medicine, Universitas Andalas, Padang, Indonesia

Corresponding Author: Hirowati Ali

DOI: https://doi.org/10.52403/ijrr.20230876

ABSTRACT

SARS CoV-2 was discovered in Wuhan, China, in December 2019 and has since mutated into new variants. One of the genes that mutated in SARS CoV-2 is RNA-dependent RNA polymerase (RdRp), a gene playing roles in viral replication and transcription. This study aimed to determine the variation of the RdRp gene in the SARS CoV-2 variant. We used VTM-preserved biological samples from positive nasopharyngeal swabs confirmed by qRT-PCR testing stored at -80°C. We isolated samples using the QIAamp Viral RNA Mini Kit and determined their concentration with Qubit RNA HS Assay reagents. After amplifying the RNA using qRT-PCR, we prepared the samples for sequencing with the Illumina RNA prep kit. Sequencing was carried out with the Miseq system and results were produced in FASTQ format.

This study identified 251 samples, and analysis of the RdRp gene showed 69 substitution mutations and one deletion mutation. The highest point mutation was found in the C14408T nucleotide sequence, almost in all study samples of all variants (98.4%). There were 28 missense mutations, including C14408T, which converts the amino acid proline to leucine. Four research samples that did not have mutations in C14408T were samples in the early period of the COVID-19 pandemic before various variants were reported. In summary, our study on the variation of the RdRp gene in the SARS CoV-2 variant demonstrated remarkable C14408T mutations. Utilizing this mutation is a crucial genetic indicator for tracking alterations in the SARS-CoV-2 and comprehending its evolutionary patterns.

Keywords: SARS CoV-2, COVID-19, variant, mutation, RdRp gene

INTRODUCTION

The discovery of cases of pneumonia of unknown cause in Hubei Province, Wuhan, China, in December 2019 became the beginning of the Coronavirus disease 2019 (COVID-19) pandemic. It has become a world health problem in the last two years. Since it was identified on December 30, 2019, that Severe Acute Respiratory Syndrome Coronavirus 2 (SARS CoV-2) is the cause of COVID-19, the transmission of this infection has continued to infect more than 564 million people based, on the World Health Organization (WHO) data as of July 2022 and cause more than 6.3 million deaths worldwide (1).

Symptoms that can arise from SARS-CoV-2 virus infection vary, so it requires sensitive diagnostic tools to determine the presence of this viral infection. qRT-PCR examination allows the detection of sequences of specific genes that code for viral proteins, such as RNA-dependent RNA polymerase (RdRp), nucleocapsid (N), envelope (E), and spike (S). The RdRp gene, which is part of ORF1ab, most often used in detecting SARS CoV-2, has a high analytical sensitivity (2, 3).

The RdRp gene, also known as Nsp12, is part of the nonstructural protein SARS CoV-2 to be a core component in the viral RNA replication and transcription complex. The RdRp gene has significant polymerase activity with the help of other nonstructural

proteins, namely Nsp7 and 8, Nsp12-Nsp7-Nsp8 bonds. This complex structure consists of the core catalytic unit of the RdRp gene, the Nsp7-Nsp8 heterodimer, and additional Nsp8 subunits, which accelerate viral replication. Studies report that the RdRp C14408T mutation in Europe is associated with increased mutation rates compared to genomes of Asian origin through mechanisms that have not yet been elucidated. This is the reason for further research in the analysis of RdRp mutations in virus variants that cause increased spread and death due to SARS-CoV-2 infection (4). Thus, this study is expected to discover the pattern of mutations and the effects caused by these mutations.

MATERIALS & METHODS

The research was conducted at the Laboratory of the Center for Diagnostic and Research of Infectious Diseases (PDRPI) Faculty of Medicine, Andalas University, Padang, and the Health Research and Development Laboratory of the Ministry of Health of the Republic of Indonesia, Jakarta. This study sample is biological material stored in the Viral Transport Medium (VTM) from nasopharyngeal swab results with positive confirmed qRT-PCR examination stored in a freezer -80oC. Samples were isolated using QIAamp Viral RNA Mini Kit, and RNA isolates were measured in concentration using Qubit RNA HS Assay reagents. The next stage is RNA amplification using the qRT-PCR method using mBiocov-19 qRT-PCR. Library preparation of samples is performed before sequencing according to the Illumina RNA prep with enrichment (L) tagmentation kit protocol. The results of library preparation work that have produced samples with uniform concentrations will be diluted according to the provisions of the kit protocol. The sequencing process is carried out according to the Miseq system guide, with the results in FASTQ data.

Ethical approval was obtained from Faculty of Medicine, Universitas Andalas, Padang, Indonesia number 664/UN.16.2/KEP-FK/2022

STATISTICAL ANALYSIS

Data analysis was conducted to determine the type of mutation and variation of the RdRp gene using the CLC Genomics Workbench app software. These RdRp gene sequencing results were compared with GenBank reference data using the CLC Genomics Workbench app, with the GenBank reference SARS CoV-2 genome (NC_045512.2).

RESULT

Mutation points in SARS CoV-2 RdRp were identified using the CLC Genomic Workbench App by looking for single nucleotide polymorphism (SNPs), such as insertions, substitutions, or deletions. Alignment of research samples with reference SARS CoV-2 Wuhan-Hu-1 complete genome isolate (NC_045512.2), then trimming was carried out at the location of the RdRp gene (sequence 13442-16236). Mutations are characterized by bars not reaching 100%.

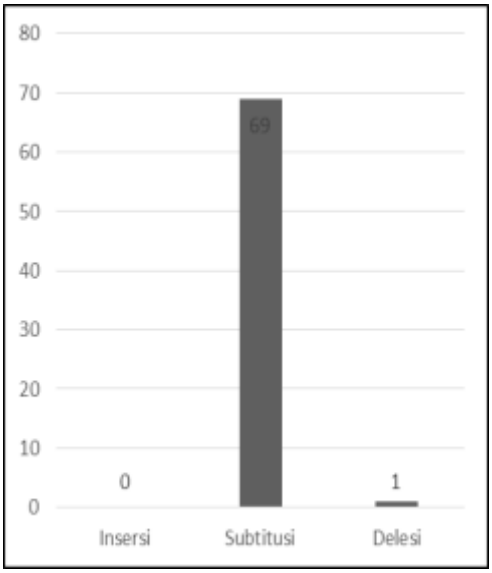

Figure 1. Types of SARS CoV-2 RdRp gene mutations. Figure 1 shows the types of mutations found in SARS CoV-2 RdRp substitution and deletion mutations, where substitution is the most common type in the study sample.

Table 1 RdRp gene variation of SARS CoV-2 non-VOI-VOC variant. The results of the analysis of SARS CoV-2 RdRp gene variation in non-VOI-VOC variants showed that most mutations were substitutes for C14408T, which caused changes in the amino acid proline to leucine in 163 study samples (97.6%).

Mutation	Changes in amino acids	All variant		Sample code	Non-VOI-VOC variant	
		n (251)	%		n (167)	%
Substitution						
G13505T	C>F	1	0,4	IL84	1	0,6
C13536T		2	0,8	NV59, NV62	2	1,2
C13584T	H>D	1	0,4	IL 67	1	0,6
C13608T		1	0,4	NV40	1	0,6
C13620T		1	0,4	NV67	1	0,6
C13665T		2	0,8	NV2, NV72	2	1,2
A13712G	K>R	1	0,4	NV25	1	0,6
C13730T	A>V	4	1,6	IL1, IL8, IL45, IL65	4	2,4
C13923T		1	0,4	IL 60	1	0,6
A14032G	N>D	3	1,2	NV45, NV51, NV64	3	1,8
G14055A		1	0,4	NV36	1	0,6
C14120T	P>L	2	0,8	NV15, IL 49	2	1,2
C14322T		1	0,4	NV15	1	0,6
C14340T		1	0,4	IL 14	1	0,6
G14346T	L>F	3	1,2	NV34, IL 48, IL 80	3	1,8
C14408T	P>L	247	98,4	NV1-NV75 IL4-IL6, IL11, IL13-IL17, IL19-IL44, IL46-IL64, IL66-99	163	97,6
A14457T		1	0,4	NV22	1	0,6
C14571T		25	10,0	IL 11, IL14, IL20, IL38, IL42, IL43, IL50, IL56, IL60, IL64, IL67, IL68, IL71, IL73, IL75, IL76, IL77, IL78, IL82, IL89, IL91, IL92, IL93, IL94, IL96	25	15
C14583T		1	0,4	NV64	1	0,6
T14610C		1	0,4	NV58	1	0,6
C14703T		1	0,4	NV5	1	0,6
C14724T		1	0,4	IL 88	1	0,6
T14742C		1	0,4	IL 79	1	0,6
G14785A	A>T	1	0,4	NV63	1	0,6
C14790T		2	0,8	IL 23, IL 24	2	1,2
C14805T		1	0,4	IL99	1	0,6
G14874T	K>N	1	0,4	NV9	1	0,6
A14940G		6	2,4	NV34, IL 40, IL 48, IL 63, IL 70, IL 80, IL 81	6	3,6
G15013A	D>N	1	0,4	IL 52	1	0,6
T15021C		1	0,4	IL 78	1	0,6
T15114C		1	0,4	IL 59	1	0,6
G15181T	A>S	1	0,4	IL 49	1	0,6
T15203C	V>A	1	0,4	NV21	1	0,6
G15226A	G>S	2	0,8	IL 19, IL 90	2	1,2
C15277T	H>Y	3	1,2	NV25, IL 50, IL 71	3	1,8
G15327T	M>I	1	0,4	IL 59	1	0,6
C15368T	T>I	1	0,4	NV67	1	0,6
T15408C		1	0,4	NV55	1	0,6
G15444T	M>I	1	0,4	NV29	1	0,6
T15501C		2	0,8	NV23, NV24	2	1,2
A15605C	N>L	1	0,4	IL 84	1	0,6
G15543T		1	0,4	NV74	1	0,6
A15657G		3	1,2	NV13, NV43, NV44	3	1,8
G15672T	E>D	1	0,4	NV30	1	0,6
C15699T		1	0,4	IL 48	1	0,6
C15720T		1	0,4	NV21	1	0,6
C15810T		1	0,4	NV69	1	0,6
C15951T		3	1,2	NV2, NV37, NV72	3	1,8
T16062G	N>K	1	0,4	NV1	1	0,6
C16080T		1	0,4	IL 77	1	0,6
C16092T		1	0,4	NV64	1	0,6
G16106A	R>K	3	1,2	NV23, NV24, NV30	3	1,8

Table 2 RdRp gene variation of SARS CoV-2 delta variant. Table 2 shows two mutation points in all delta variant samples (100%). Substitutions at position C14408T and position G15451A cause the amino acid glycine to serine.

Mutation	Changes in amino acids	*All variant*		Sample code	Delta	
		n (251)	%		n (49)	%
Substitution						
A13566G		2	0,8	DE18, DE19	2	4,1
T14034C		1	0,4	DE2	1	2
C14408T	P>L	247	98,4	DE1- DE49	49	100
C15105T		1	0,4	DE32	1	2

G15243T	M>I	1	0,4	DE35	1	2
A15270T	E>D	1	0,4	DE46	1	2
C15273T		1	0,4	DE46	1	2
T15276C		1	0,4	DE46	1	2
C15277A	H>N	1	0,4	DE46	1	2
G15451A	G>S	49	19,5	DE1-DE49	49	100
T15825C		1	0,4	DE31	1	2
C16192T	P>S	2	0,8	DE44, DE45	2	4,1

Table 3 RdRp gene variation of SARS CoV-2 omicron variant, mutations were also found at the position of C14408T in the omicron variant.

Mutation	Changes in amino acids	*All variant*		Sample code	Omicron	
		n (251)	%		n (35)	%
Substitution						
C13684G		2	0,8	OM22, OM30	2	5,7
C14408T	P>L	247	98,4	OM1- OM35	35	100
T14910C		1	0,4	OM26	1	2,9
C14913T		1	0,4	OM10	1	2,9
C15240T		20	8,0	OM1, OM2, OM3, OM5-OM15, OM18, OM19, OM22-OM25	20	57,1
T15521A	F>Y	1	0,4	OM11	1	2,9
C15714T		15	6,0	OM4, OM16, OM17, OM20, OM21, OM26-OM35	15	42,9
Deletions						
G15745		2	0,8	OM22, OM30	2	5,7

In addition, there were two silent mutations found in the omicron variant in the C15240T position in as many as 20 samples (57.1%) and C15714T in 15 study samples (42.9%), as well as deletion mutations in the nucleotide sequence 15745 (G15745) in 2 omicron variant research samples.

DISCUSSION

Specific mutations in the Nsp12 protein (RdRp) were followed by increased mutations in Europe and America compared to genomes that did not have mutation points in Nsp12. Most mutations were found in the position of nucleotide 14408, which affects the binding of Nsp12 with Nsp14, which increases the likelihood of further mutations. Nsp14 has a role in Proofreading during RNA synthesis; in this case, it plays a role in correcting errors during synthesis (5). This study showed the results of the most mutation analysis in the form of substitutions at 69 points and deletions at one nucleotide position. The most substitutions occurred in the change of cytosine to uracil, as many as 31 samples (44.3%), followed by the substitution of uracil to cytosine 15.7% and guanine to uracil 12.9%.

The results showed the most mutations, almost in all study samples, in C14408T, which caused missense mutations, the amino acid proline to leucine. The alternation of proline to leucine at position 4715 (P4715L) RdRp of the polyprotein ORF1ab appears in the enzyme RdRp as a substitute for proline with leucine at position 323 (P323L). This enzyme acts as a catalyst in viral RNA replication and can proofreading, with mutations at this point, can affect changes in the ability of appropriate viral replication to contribute to viral infectivity and disease severity. The C14408T mutation that causes mutations in P323L was associated with an overall increase in mutations in the SARS-CoV-2 genome (6).

Mutations in P323L often occur along with nucleotide changes in the spike protein resulting in changes in aspartic acid to glycine at position 614 (D614G). Ilmjärv S et al. mentioned that changes in one of these frequent mutations, P323/G614 or L323/D614, showed relatively slower emergence when compared to mutants in both positions (L323/G614) with very fast and efficient emergence (7).

RdRp mutations are also associated with changes in M and E proteins, in the study of Eskier D et al. showed a significant relationship between changes in RdRp nucleotides at positions 13730, 14408, 14805, 15324 with changes in proteins M and E. These proteins are part of the genome that evolves relatively slowly. Hence, the influence of RdRp mutations on

changes in M and E proteins strengthens the hypothesis that RdRp mutations contribute significantly to the evolution of the SARS CoV-2 genome (8). The results showed that there were mutations in the position of 13730 in as many as four samples, 14408 as many as 247 samples, and 14805 as many as 1 sample. Viruses with mutant RdRp become more resistant to antiviral drugs. In addition, mutant RdRp can cause amino acid changes in other positions. RdRp mutations can cause acceleration or slow down the replication process, which will later affect viral load and virulence (8).

A typical mutation was found in this study in all delta variants of the G15451A nucleotide sequence, resulting in a missense mutation of the amino acid change of glycine to serine. This mutation was also identified at a time of increase in delta variant cases in South Africa by whole genome analysis of delta variant sequences, showing that 100% of the sequences analyzed had mutations at this point. The G15451A mutation is not found in previous variants, be it alpha, beta or gamma. This mutation affects the results of qRT-PCR examination using the target gene RdRp, reducing the amplification value of the test results. It significantly caused a single nucleotide mismatch at the penultimate base of the forward primer binding site for the WHO-recommended RdRp gene (9). In addition, silent mutations were found in the omicron variant nucleotide sequences C15240T and C15714T. The C15240T mutation was also reported in EPI_ISL_437308 isolates from Ankara, Turkey (10).

The sequencing results in this study showed genome variations in the virus. In this case, the analysis was Nsp12 / RdRp. Genome variation due to mutations is an adaptation process that occurs in viruses when infecting humans (11, 12).

CONCLUSION

The results of the RdRp gene analysis study on SARS CoV-2 virus variants show that changes in the RdRp gene in the study sample were found at 70 points with 28 missense mutations. Most of the mutations were found in C14408T. This finding emphasizes the potential role of this mutation as a genomic marker for monitoring the dynamic changes in the SARS-CoV-2 population. It highlights its importance in understanding the evolutionary dynamics of the virus.

Declaration by Authors

Ethical Approval: Approved

Acknowledgement: We would like to extend our gratitude to dr. Mutia Lailani for proofreading this manuscript.

Source of Funding: None

Conflict of Interest: The authors declare no conflict of interest.

REFERENCES

1. WHO. WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/: https://covid19.who.int/. 2022
2. Feng W, Newbigging AM, Le C, Pang B, Peng H, Cao Y, et al. Molecular Diagnosis of COVID-19: Challenges and Research Needs. Anal Chem. 2020;92, 10196-10209.
3. Böger B, Fachi MM, Vilhena RO, Cobre AF, Tonin, FS, Pontarolo R. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19. Am J Infect Control. 2021;49, 21-29.
4. Jiang Y, Yin W, Xu HE. RNA-dependent RNA polymerase: Structure, mechanism, and drug discovery for COVID-19. Biochemical and Biophysical Research Communications. 2021;538, 47-53.
5. Kadam SB, Sukhramani GS, Bishnoi P, Pable AA. Barvkar, VT. SARS-CoV-2, the pandemic coronavirus: Molecular and structural insights. J Basic Microbiol. 2021;61, 180-202.
6. Biswas SK, Mudi SR. Spike protein D614G and RdRp P323L: the SARS-CoV-2 mutations associated with severity of COVID-19. Genomics Inform. 2020;18, e44.
7. Ilmjärv S, Abdul F, Acosta-Gutiérrez S, Estarellas C, Galdadas, I, Casimir M, et al. Concurrent mutations in RNA-dependent RNA polymerase and spike protein emerged as the epidemiologically most successful

SARS-CoV-2 variant. Scientific Reports. 2021;11, 13705.
8. Eskier D, Karakülah G, Suner A, Oktay Y. RdRp mutations are associated with SARS-CoV-2 genome evolution. PeerJ. 2020;8, e9587.
9. Valley-Omar Z, Marais G, Iranzadeh A, Naidoo M, Korsman S, Maponga T, et al. The reduced amplification efficiency of the RNA-dependent-RNA-polymerase target enables tracking of the Delta SARS-CoV-2 variant using routine diagnostic tests. Journal of Virological Methods. 2022;302, 114471.
10. DemİR AB, Benvenuto D, Abacioğlu H, Angeletti S, Ciccozzi M. Identification of the nucleotide substitutions in 62 SARS-CoV-2 sequences from Turkey. Turk J Biol. 2020;44, 178-184.
11. ZhaoN, Zhou N, Fan H, Ding J, Xu X, Dong X, et al. Mutations and Phylogenetic Analyses of SARS-CoV-2 Among Imported COVID-19 From Abroad in Nanjing, China. Front Microbiol. 2022;13, 851323.
12. Forster P, Forster L, Renfrew C, Forster, M. Phylogenetic network analysis of SARS-CoV-2 genomes. Proc Natl Acad Sci U S A. 2020;117, 9241-9243.

How to cite this article: SM Rezvi, Desmawati Desmawati, Andani Eka Putra, Hirowati Ali. Analysis of RdRp gene in SARS CoV-2 variants. *International Journal of Research and Review*. 2023; 10(8): 580-585. DOI: *https://doi.org/10.52403/ijrr.20230876*

******