

ARTIKEL PENELITIAN

## Variasi Genetik *Region inding Motif* (RBM) Gen Pike SARS CoV-2 Pada Isolat Lokal Sumatera Barat

Dede Rahman Agustian<sup>1</sup>, Andani Eka Putra<sup>2</sup>, Dessy Arisanty<sup>3</sup>

1. Program Studi Magister Ilmu Biomedik Fakultas Kedokteran Universitas Andalas, Padang, Indonesia; 2. Bagian Mikrobiologi, Fakultas Kedokteran Universitas Andalas, Padang, Indonesia; 3. Bagian Biokimia, Fakultas Kedokteran Universitas Andalas, Padang, Indonesia.

**Korespondensi:** [ddrahman.md@gmail.com](mailto:ddrahman.md@gmail.com)

### Abstrak

**Tujuan:** untuk mengetahui sebaran titik mutasi pada sekuens *Region Binding Motif* (RBM) gen *Spike* SARS CoV-2 yang berasal dari isolat lokal Sumatera Barat. **Metode:** studi deskriptif eksploratif menggunakan koleksi spesimen laboratorium Pusat Diagnostik dan Riset Penyakit Infeksi (PDRPI) Fakultas Kedokteran Universitas Andalas yang diisolasi, dilanjutkan tahap *sequencing* dan analisa data secara bioinformatika. **Hasil:** terdapat tiga titik mutasi *non-synonymous* disepanjang RBM gen *Spike* SARS Co-2. **Kesimpulan:** mutasi Asn439Lys merupakan titik mutasi tersering pada RBM gen *Spike* sebesar 11% diikuti Ser477Gly dan Thr478Ile masing-masingnya 3% dan 1%.

**Kata kunci:** Mutasi; RBM; gen Spike

### Abstract

**Objective:** to determine mutation sites along *Region Binding Motif* (RBM) of *Spike* gene in local West Sumatera specimens. **Method:** explorative descriptive study that use specimen collection in Diagnostic and Research Center of Tropical Infection, Medical Faculty, Andalas University, which is extracted, sequenced, and analyzed with bioinformatic. **Result:** three non-synonymous mutation found in RBM of *Spike* gene. **Conclusion:** Asn439Lys mutation is the most common mutation found in RBM of *Spike* gene SARS-CoV-2, furthermore another mutation found in this study is Ser477Gly and Thr478Ile around 3% and 1% respectively.

**Keywords:** Mutation; RBM; Spike gene

## PENDAHULUAN

Kasus pneumonia akibat infeksi novel korona virus sangat intens dilaporkan dalam satu tahun terakhir. Pertama kali di laporkan pada akhir bulan Desember 2019 di Wuhan China, laju penyebaran *Coronavirus Disease-19* (COVID-19) hingga saat ini masih tetap tinggi. Jumlah kumulatif total kasus yang dilaporkan hingga 3 April 2020 sebanyak 127.877.462 dengan penambahan jumlah kasus baru dalam 24 jam terakhir sebanyak 509.746<sup>1</sup>. Laju penyebaran terjadi begitu cepat melewati batas geografis, dalam rentang waktu tiga bulan setelah kasus pertama Wuhan dilaporkan, total 203 negara lain ikut mengonfirmasi kasus pertamanya. Respon muncul dari badan kesehatan dunia dengan menetapkan pada 30 Januari 2020 COVID-19 sebagai *Public Health Emergency of International Concern*.

Pengendalian penyebaran masih menjadi tantangan meski telah satu tahun berlalu. Beberapa negara melaporkan kenaikan kembali setelah mengalami penurunan kasus sebelumnya. Pemulihan aktifitas ekonomi, aktifitas sosial seperti pembukaan kembali sekolah dan kantor serta reintroduksi virus dari daerah lain diduga menjadi beberapa faktor yang memicu munculnya *second wave*. Mutasi-mutasi yang baru bermunculan juga dihubungkan dengan munculnya *second wave*.

Variasi gejala pada pasien positif COVID-19 mengakibatkan penyebaran SARS-CoV-2 menjadi semakin sulit untuk dihentikan. Proporsi asimtomatis dan presimtomatis COVID-19 sangat tinggi. Suatu penelitian pada kapal yang

terjangkit didapati 81% pasien bersifat asimtomatis<sup>2</sup>. Hal ini menunjukkan besarnya potensi *silent transmission* COVID-19 ditengah masyarakat.

Berdasarkan informasi nCoV-19 (SARS-CoV-2) yang diperoleh dari *database* GISAID pada Januari 2020, sekuens yang pertama kali di unggah sudah memiliki banyak perbedaan dengan sekuens yang akhir-akhir ini di unggah (6 April 2021). Hingga saat ini diyakini terdapat tiga varian yang sedang mendominasi secara global yaitu varian A, B dan C yang dibedakan berdasarkan perubahan asam amino. Protein *Spike* tepatnya pada RBM merupakan jalan masuk bagi virus untuk dapat menginfeksi sel epitel melalui interaksinya terhadap reseptor *human Angitensin Converting Enzyme* (hACE). Perubahan pada komposisi protein *Spike* memiliki potensi besar merubah daya patogenitas virus dan menjadi tantangan dalam pengembangan vaksin.

Belum ada standar baku dalam pengobatan COVID-19 mengakibatkan pandemi tampaknya belum akan segera berakhir. Kebanyakan obat hanya bersifat suportif dimana terapi oksigen menjadi intervensi utama dalam penatalaksanaan pasien COVID-19.

## METODE

Penelitian dilakukan di Laboratorium Pusat Diagnostik dan Riset Penyakit Infeksi Fakultas Kedokteran Universitas Andalas (PDRPI FK Unand) Padang, Laboratorium Bioinformatika FK Unand, Padang dan Laboratorium Penelitian dan Pengembangan Kesehatan Kementerian Kesehatan Republik Indonesia (Litbangkes Kemenkes RI), Jakarta. Penelitian

berlangsung pada bulan Maret - Mei 2021. Populasi penelitian ini adalah koleksi *Viral Transport Medium* (VTM) hasil pemeriksaan swab nasofaring yang terdeteksi positif di Laboratorium PDRPI FK Unand, Padang. Sampel pada penelitian ini adalah VTM hasil pemeriksaan swab nasofaring yang terdeteksi positif di Laboratorium PDRPI FK Unand yang memenuhi kriteria inklusi. Alat yang digunakan pada penelitian ini berupa Biosafety cabinet, autoclave, heating block, mikropipet 10-1000  $\mu$ L, multichannel mikropipet, tip pipet mikro ukuran 10-1000  $\mu$ L, microcentrifuge tube, timer, vortex mixer, microwave, timbangan digital, spin down machine, mesin sentrifus, plate shaker, thermal cycler, magnetic stand, perangkat elektroforesis, freezer, gel documentation system, miseq illumine sequencer, alat pelindung diri standar level 3, atanol absolut, QIamp Viral RNA, 96 well PCR plate, microseal B adhesive film, Illumina RNA enrichment kit, alcohol 80%, RNase zap, Qubit RNA/DNA assay, PCR kit mBIOCOV, Tris Boric acid EDTA (TBE) buffer, loading dye, DNA ladder. Sampel terpilih dalam penelitian ini akan diisolasi sesuai prosedur kit *QIAamp Viral RNA mini kit*. Selama proses pengerjaan sampel selalu diupayakan *on ice* untuk menjaga kualitas RNA SARS-COV-2. Tahapan isolasi meliputi *lisis, binding, washing dan elution*. Hasil isolasi akan segera diuji menggunakan *Qubit HS assay*, proses pengerjaan mengikuti protocol *Qubit HS assay*. Sampel dengan nilai lebih dari 100 ng akan diencerkan sedangkan sampel dengan nilai kurang dari 20 ng akan diisolasi ulang dengan meningkatkan konsentrasi atau dapat diganti dengan sampel lain yang memenuhi kriteria. Sampel yang

memenuhi kriteria konsentrasi qubit akan dilanjutkan pada pemeriksaan QRT-PCR. Pengerjaan QRT-PCR mengikuti protokol kit mBIOCOV. Pengaturan alat pada setiap tahapannya yaitu *Pre-denaturation, Denaturation, Annealing dan Extension* dan suhu pada setiap tahapan disesuaikan dengan hasil optimasi yang dilakukan. Selama pengerjaan juga ditambahkan kontrol negatif dan kontrol positif dalam plate PCR yang sama untuk meningkatkan validitas pengerjaan. Hasil QRT-PCR kemudian dianalisa, sampel dengan Ct > 25 dieksklusikan dari penelitian dan akan digantikan dengan sampel baru yang memenuhi kriteria. Isolat yang memenuhi kriteria akan disimpan didalam kulkas -80°C untuk menjaga kualitas RNA tetap baik hingga pengerjaan *library preparation*.

*Library preparation* dilakukan sesuai protokol kit *Illumina RNA prep with enrichment (L) tagmentation*. Terdapat sepuluh tahapan dalam pengerjaan meliputi denaturasi, sintesis *first strand*, sintesis *second strand*, tagmentasi cDNA, *clean up library*, normalisasi *library*, Hibridisasi *probe*, *capture hybridized probe*, amplifikasi *enrichment library*, dan *clean up enrichment library*. Setiap step pengerjaan peneliti akan didampingi teknisi ahli dari provider Illumina. Tahap akhir setelah menyelesaikan sepuluh step *library preparation*, sampel akan dianalisa kualitas hasil pengerjaan serta konsentrasi masing masingnya dan akan dipersiapkan untuk tahap sekuensing menggunakan miseq sequencer.

## HASIL DAN PEMBAHASAN

### Isolasi RNA SARS-Cov-2

Penelitian ini dilakukan untuk menganalisa variasi RBM gen Spike pada pasien positif COVID-19 hasil koleksi laboratorium PDRPI. Sampel yang digunakan pada penelitian ini sebanyak 89 sampel. Populasi sampel yang direncanakan untuk ikut dalam penelitian akan dilakukan proses pengujian kelayakan berupa kuantifikasi dan RTqPCR ulang terlebih dahulu diikutsertakan dalam penelitian. Sebanyak 140 ul volume sampel digunakan pada tahapan ekstraksi ulang. SARS-CoV-2 memiliki materi genetik berupa RNA, materi genetik RNA bersifat lebih labil jika dibandingkan dengan DNA sehingga cenderung membutuhkan teknik khusus untuk mengkonservasi RNA. Panjang minimal nukleotida yang dapat tertahan menggunakan membran silika QIAmp adalah 200 pasang basa nukleotida sedangkan panjang RNA virus SARS-CoV-2 (MN908947.3) sepanjang 29.903 nukleotida.

Nilai akhir konsentrasi RNA kandidat sampel didapati cukup beragam, nilai terendah didapatkan pada sampel T2 dengan nilai konsentrasi 2,5 ng/ul dan total input RNA menjadi sebesar 21,3 ng/8,5 ul dan nilai konsentrasi tertinggi didapatkan pada sampel dengan kode T63 dengan konsentrasi 11,7 ng/ul sehingga total input RNA menjadi 99,5 ng/ 8,5 ul. Nilai rerata konsentrasi RNA pada sampel yang akan dilakukan pengerjaan adalah 52,6 ng/8,5ul.

Hasil kuantifikasi Qubit pada 89 sampel memenuhi kriteria untuk dilanjutkan pada tahap *library preparation*, dimana cut off batas nilai untuk masuk dalam tahap *library preparation* adalah sebesar 20-100ng/8,5 ul per sampel yang akan diuji.

### Amplifikasi RNA

Amplifikasi RNA menggunakan *Realtime Quantitative Reverse Transcription Polymerase Chain Reaction* (qRT-PCR) bertujuan untuk mendapatkan RNA target dengan jumlah cukup sehingga dapat divisualisasikan secara *real time*. Pada penelitian ini kit yang digunakan adalah mBioCov. mBioCov merupakan one step qRT-PCR kit yang dikembangkan secara khusus untuk mendeteksi sekuens RNA SARS-CoV-2, kit ini menggunakan primer dan probe multiplex. Multiplex PCR mempermudah proses pengerjaan dengan menggunakan probe berfluorofor tertentu yang dapat berikatan dengan titik spesifik pada sekuens target, hal ini memungkinkan untuk menggabungkan beberapa primer dan probe dalam satu well pemeriksaan PCR. Kit mBioCoV yang digunakan pada penelitian ini menggunakan marker fluorofor FAM untuk mendeteksi gen *Helicase*, fluorofor HEX untuk mendeteksi gen RDRP serta fluorofor CY5 yang digunakan untuk mendeteksi gen RPP30 yang digunakan sebagai internal kontrol pada penelitian ini. Volume total reaksi pada setiap well dalam penelitian ini adalah 20 ul. Jumlah template akan berhubungan dengan

*viral load* dan menurunkan nilai *Ct value*, namun hubungan tersebut bersifat tidak langsung karena dalam penegerjaannya tidak membentuk kurva standar.

Hasil qRT-PCR pada 89 sampel mengkonfirmasi bahwa materi genetik virus SARS-CoV-2 dapat ditemukan pada seluruh isolat yang digunakan pada penelitian ini. Nilai *Ct* pada gen RDRP bervariasi mulai dari 14,46 dan tertinggi 26,71, nilai rerata sebesar 21,83. Nilai *Ct* pada gen *Helicase* juga bervariasi mulai dari 13,51 dan tertinggi 28,83, nilai rerata 20,75. Nilai *Ct* pada gen *IC* meski bukan bagian dari marker untuk SARS-CoV-2 namun akan tetap dipertimbangkan pada penelitian ini sebagai internal kontrol untuk memastikan bahwa koleksi spesimen yang digunakan berasal dari swab nasofaring dan sebagai kontrol pada proses ekstraksi materi genetik. Seluruh sampel yang telah melewati minimal cut off pemeriksaan qubit dan qRT-PCR akan dilanjutkan pada *library preparation*.

#### **Library Preparation dan sequencing**

*Library preparation* merupakan preparasi molekular sampel sebelum *sequencing*, pada tahap ini adapter akan dilekatkan pada kedua ujung fragmen DNA atau RNA, kemudian dilanjutkan proses amplifikasi untuk meningkatkan jumlah input ampikon per sampel<sup>3</sup>. Penelitian ini menggunakan *Illumina RNA prep with enrichment (L) tagmentation kit* untuk tahapan *library preparation*. *Illumina RNA prep with enrichment (L) tagmentation kit* menggunakan teknologi *paired-end*

*reads* serta *dual-index* pada fragmen RNA. Teknologi *paired-end reads* memungkinkan peneliti dapat meningkatkan presisi dalam sintesis konsensus urutan basa nukleotida karena pembacaan menggunakan pasang urutan yang bersesuaian<sup>4</sup>. Teknologi *Dual Index* pada kit membuat sampel ditandai dengan dengan marker yang spesifik sehingga pengerjaan dengan jumlah besar menjadi memungkinkan. *Illumina RNA prep with enrichment (L) tagmentation kit* menggunakan *biotinylated probe* yang dikombinasikan menggunakan *magnetic bead* untuk menangkap sekuens target. Panel yang digunakan ada penelitian ini adalah *Respiratory Virus Oligo Panel (RVOP)*. Protokol ini dioptimasi untuk 10-100 ng *purify total RNA* atau 20-100 ng RNA dari degradasi jaringan atau FFPE. Tahap akhir *library preparation* berupa normalisasi dan pengenceran hingga konsentrasi 4 nM yang merupakan nilai input sesuai protokol pada mesin *sequencing* Miseq<sup>5</sup>.

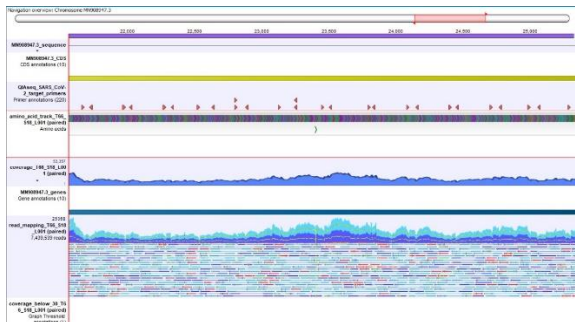
Pengerjaan *library preparation* menghabiskan waktu selama dua hari dan terpisah pada dua *batch*. Output konsentrasi *library* rata-rata adalah 86,07 konsentrasi tersebut kemudian diencerkan hingga menjadi 4nM. Proses *sequencing* menggunakan mesin Miseq membutuhkan waktu lebih kurang 12 jam untuk setiap *running* mesin.

#### **Analisa bioinformatika**

Output berupa data FASTQ dari mesin sequencer akan di proses dengan aplikasi *CLC genomic workbench*. Proses

yang dilakukan meliputi *quality control*, *trim*, *mapping read*, *coverage mapping*, sintesis konsensus, *variant calling*.

Gambar 1 menunjukkan hasil analisa bioinformatika pada sampel T66 tampak pada gambar *coverage* pada gen S tidak memiliki gap. Celah pada *coverage* akan mengakibatkan konsensus urutan basa nukleotida menjadi tidak lengkap. Kelengkapan konsensus urutan basa sangat penting bagi penelitian yang ingin mengungkap variasi variasi nukleotida pada suatu organisme<sup>6</sup>.



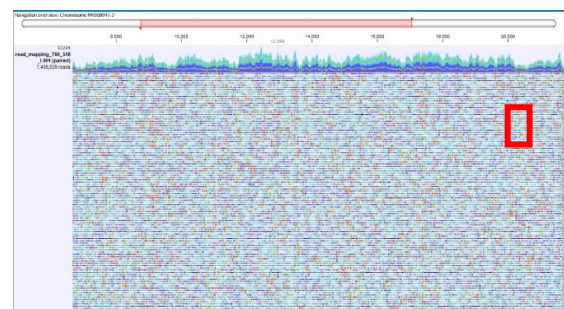
**Gambar 1.** Analisa bioinformatika CLC Genomic Workbench

Nilai *phred score* Q telah lama dijadikan sebagai indikator untuk menilai kualitas output hasil *high throughput sequencing*. Semakin tinggi standar *phred score* yang digunakan akan menurunkan nilai kemungkinan kesalahan pada *base calling* ketika dilakukan penyusunan konsensus urutan basa<sup>7</sup>. Nilai *phred score* dapat dilihat pada Tabel 1. Standar *phred score* minimal yang digunakan pada penelitian ini adalah Q30 yang memiliki akurasi hingga 99,9 persen.

**Tabel 1. Phred score**

Phred score	probabilitas kesalahan base call	akurasi base call
10	1 dari 10	90%
20	1 dari 100	99%
30	1 dari 1.000	99,9%
40	1 dari 10.000	99,99%
50	1 dari 100.000	99,999%
60	1 dari 1.000.000	99,9999%

*Coverage depth* menunjukkan rerata jumlah *read* yang telah disejajarkan dengan *reference* yang digunakan, semakin tinggi nilai *coverage depth* maka akan menurunkan kemungkinan kesalahan dalam penyusunan konsensus. Gambar 2 dan 5.5 menunjukkan contoh tampilan *coverage depth* pada sampel T66 dan Tabel 3 menunjukkan rerata *coverage depth* dari 89 sampel yang digunakan pada penelitian 3980,8.

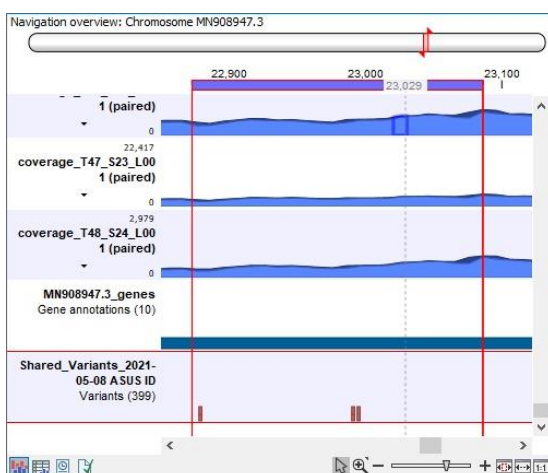


**Gambar 2.** Coverage depth sampel T66

Delapan puluh sembilan sampel dilanjutkan pada analisis *variant calling*. Peneliti menggunakan *minimum variety above 50* sebagai batas *variant base call* untuk setiap sampelnya. Hal ini ditujukan untuk menyeleksi kemungkinan *error calling* setiap sampel yang dianalisa perbedaan urutan basanya dengan *reference*. Gambar 4 menunjukkan keseluruhan *variant calling above 50* yang terdeteksi oleh CLC genomic workbench.



**Gambar 3.** magnifikasi Coverage depth sampel T66



**Gambar 4.** Variant calling analysis 89 sampel penelitian

### Variasi RBM genetik gen Spike

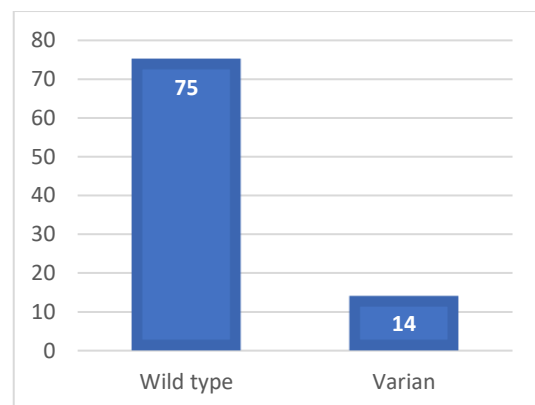
Setiap nota bar berwarna merah pada gambar 5.6 menunjukkan posisi mutasi yang terdeteksi disepanjang RBM gen S, total terdapat 3 bar yang menunjukkan telah terjadi mutasi *non-synonymous* pada 3 titik disepanjang 89 sekuens RBM yang dianalisa dengan menjadikan MN9089947.3 sebagai *reference*. Tampak tiga bar yang menunjukkan pada titik tersebut terjadi mutasi dengan output masing masing yang berbeda.

Setelah dilakukan pengujian dengan *variant calling* disimpulkan tidak terdapat 14 (15,7%) sampel dengan

mutasi pada RBM SARS-CoV-2 dari 89 sampel yang dianalisis (Gambar 5.7). Rerata jumlah mutasi per sampel didapatkan sebesar 0,01%. Beberapa penelitian telah melaporkan jumlah rerata mutasi persampel pada keseluruhan sekuens SARS-CoV-2 di Jepang (4,55), Hongkong (5,0), Kenya (5,38)<sup>8</sup>.

**Tabel 2 Coverage depth pada 89 sampel penelitian**

coverage depth	jumlah sampel
0-50	5
51-1000	23
1000-2000	9
2000-3000	7
3000-4000	12
4000-5000	15
>5000	23



**Gambar 5.** Proporsi wildtype dan varian

Variasi RBM gen S menjadi fokus utama dalam kajian molekular. Gen S menjadi pintu masuk bagi proses pengenalan virus dengan reseptor pada sel inang, sehingga merupakan target penting dalam pengembangan vaksin dan terapi<sup>9</sup>. Laju mutasi yang tinggi utamanya pada gen S menjadi salah satu masalah yang menghambat perkembangan vaksin dan terapi serta

variasi pada kondisi tertentu gen S juga berpotensi meningkatkan transmisi, keparahan penyakit, efikasi vaksin dan netralisasi antibodi<sup>10-12</sup>.

**Tabel 3 Jenis mutasi titik pada 89 sampel penelitian**

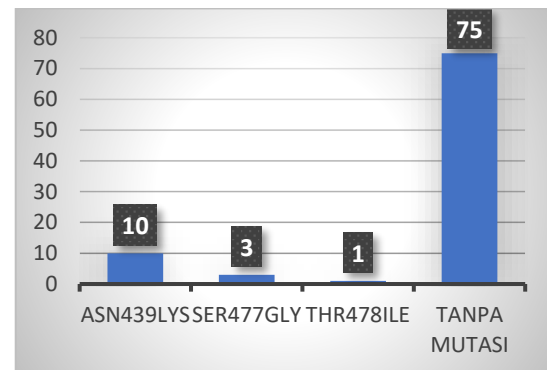
Jenis Mutasi	Jumlah Titik Mutasi Gen S SARS-CoV-2
Inseri	-
Substitusi	3
Delesi	-

Tabel 3 menunjukkan jenis mutasi titik yang ditemukan dari 89 sampel pada penelitian ini. Total peneliti mendapatkan 3 mutasi *non-synonymous* disepanjang RBM gen S. Substitusi merupakan satu satunya tipe mutasi yang ditemukan pada penelitian ini. Penelitian yang dilakukan Koyama dengan menggunakan 10.022 sampel menunjukkan hasil searah dimana substitusi terjadi pada 95% sampel dan sisanya 5% berupa delesi<sup>13</sup>. Penelitian lain juga menunjukkan bahwa substitusi merupakan tipe mutasi tersering dengan 91% dan diikuti delesi 9%<sup>14</sup>.

Mutasi titik yang terjadi pada gen S dapat berupa *synonymous* ataupun *nonsynonymous*. Mutasi *nonsynonymous* akan mengakibatkan perubahan pada asam amino yang menyusun protein S. Mutasi pada virus RNA umumnya berdampak buruk, namun dengan rentang hidup yang pendek, akumulasi mutasi dapat terkumpul dan menyediakan varian-varian yang menyesuaikan dengan seleksi alam<sup>15</sup>.

Gambar 6 menunjukkan frekuensi ketiga mutasi substitusi pada RBM gen S yang ditemukan pada 89

sampel dalam penelitian. Tampak mutasi Asn439Lys merupakan titik mutasi tersering dibandingkan mutasi lainnya (71,4%) dan terdapat pada 10 sampel dari total 89 sampel (11,2%). Ser477Gly merupakan titik mutasi tersering kedua (21%). Thr478Ile pada urutan ketiga dengan frekuensi (7%).



**Gambar 6.** distribusi frekuensi 3 mutasi titik yang ditemukan pada 89 sampel penelitian

Interaksi Region Binding Motif pada sub unit 1 protein S pada permukaan virus dengan reseptor ACE2 pada permukaan sel merupakan tahapan penting pada *viral entry* perubahan pada RBM dapat berdampak pada penurunan kemampuan virus untuk menginfeksi sel inang atau disisi lain dapat meningkatkan kemampuannya dalam menginfeksi sel inang. Pergantian asam amino Asparagin menjadi lisin pada urutan asam amino ke 439 (Asn439Lys) merupakan perubahan yang tepat terjadi pada RBM. Simulasi molekular dinamik Asn439Lys menunjukkan peningkatan *binding affinity* RBM dengan reseptor ACE2, sehingga berpotensi menjadi lebih infeksius<sup>16</sup>. Beberapa penelitian juga menunjukkan aktifitas netralisasi antibodi pada berbagai antibodi monoklonal dan poliklonal pasien pasca terinfeksi, termasuk salah satu antibodi monoklonal yang telah disetujui *Food Drug Administration* (FDA)<sup>16-18</sup>



## SIMPULAN

Terdapat tiga titik mutasi yang ditemukan disepanjang *region binding motif* (RBM) koleksi spesimen Pusat Diagnostik dan Riset Penyakit Infeksi. Salah satu dari ketiga mutasi yaitu Asn 439 Lys berpotensi memicu transmisi yang lebih kuat dan memiliki aktifitas netralisasi antibodi.

## DAFTAR PUSTAKA

1. World Health Organization. WHO Corona Virus (COVID19) Dashboard [Internet]. 2021. Available from: <https://covid19.who.int/table>
2. Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance*. 2020;25(10):2000180.
3. Van Dijk EL, Jaszczyszyn Y, Thermes C. Library preparation methods for next-generation sequencing: tone down the bias. *Exp Cell Res*. 2014;322(1):12–20.
4. Hu X, Yuan J, Shi Y, Lu J, Liu B, Li Z, et al. pIRS: Profile-based Illumina pair-end reads simulator. *Bioinformatics*. 2012;28(11):1533–5.
5. Proprietary I. Illumina RNA Prep with Enrichment, (L) Tagmentation Checklist. 2020;(August):1–5.
6. on NJ. The use of real-time PCR methods in DNA sequence variation analysis. *Clin Chim Acta*. 2006;363(1–2):32–47.
7. Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods*. 2013;10(1):57–9.
8. Mercatelli D, Giorgi FM. Geographic and genomic distribution of SARS-CoV-2 mutations. *Front Microbiol*. 2020;11:1800.
9. Du L, He Y, Zhou Y, Liu S, Zheng B-J, Jiang S. The spike protein of SARS-CoV— a target for vaccine and therapeutic development. *Nat Rev Microbiol*. 2009;7(3):226–36.
10. Focosi D, Maggi F. Neutralising antibody escape of SARS-CoV-2 spike protein: Risk assessment for antibody-based Covid-19 therapeutics and vaccines. *Rev Med Virol*. 2021;
11. Acevedo ML, Alonso-Palomares L, Bustamante A, Gaggero A, Paredes F, Cortés CP, et al. Infectivity and immune escape of the new SARS-CoV-2 variant of interest Lambda. *medRxiv*. 2021;
12. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol*. 2021;19(7):409–24.
13. Koyama T, Platt D, Parida L. Variant analysis of SARS-CoV-2 genomes. *Bull World Health Organ*. 2020;98(7):495.
14. Zekri ARN, Easa Amer K, Hafez MM, Hassan ZK, Ahmed OS, Soliman HK, et al. Genomic characterization of SARS-CoV-2 in Egypt. *J Adv Res [Internet]*. 2021;30(xxxx):123–32. Available from: <https://doi.org/10.1016/j.jare.2020.11.012>
15. Duffy S. Why are RNA virus mutation rates so damn high? *PLoS Biol*. 2018;16(8):e3000003.
16. Zhou W, Xu C, Wang P, Luo M, Xu Z, Cheng R, et al. N439K variant in spike protein may alter the infection efficiency

and antigenicity of SARS-CoV-2 based on molecular dynamics simulation. *BioRxiv*. 2020;

17. Thomson EC, Rosen LE, Shepherd JG, Spreafico R, da Silva Filipe A, Wojcechowskyj JA, et al. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-

mediated immunity. *Cell*. 2021;184(5):1171–87.

18. Wu J, Zhang L, Zhang Y, Wang H, Ding R, Nie J, et al. The Antigenicity of Epidemic SARS-CoV-2 Variants in the United Kingdom. *Front Immunol*. 2021;12:2205.