A Systematic Review of SARS-CoV-2 Mutations in Relation to its Pathogenicity*

Polly T. Chua-Chan MD, Paula Kristen I. Acal, Alron Troy G. Alde, Yyona Pauline R. Alvarez, Althea C. Bajon, Sheena Mae R. Bobier, Josephine Anne G. De Jesus, Kessathea P. Dela Cruz, Jo Arrym B. Dellosa, Megan Margery A. Dulatre, Kyra Victorien V. Dimalanta, Kyla Alexa N. Ecalla, Charlotte Angela B. Fayloga, Marc Lester T. Fermin, Therese Bianca C. Flores, Krizza Mae A. Francisco, Keena Marie C. Gaditano, Maria Ana Therese D.R. Garong, Maureen Diane C. Geli, Kristine Ann B. Geronaga, Genree Ann B. Gille, Paula Marie L. Humilde

ABSTRACT

Objective: The rapid spread of COVID-19 by SARS-CoV-2 has caused an alarming and dramatic increase in cases globally. Findings from different studies suggest that one cause of this increased infectivity lies in the evolution of the virus and the emergence of different strains. Some strains might coexist with one another but have different mutation patterns. Currently, there are 8 novel mutations of SARS-CoV-2 that were observed from different countries around the world. Recently, two more have been identified this early 2021. This paper provides a systematic review of the different mutations and an understanding of its pathogenic mechanisms in relation to pathogenicity and mortality.

Method: A systematic literature search was conducted in the study using different databases and grouping findings for discussion.

Results: The gathered data have proven that certain variants have developed specific mutations, which provided the identification of the changes in its structure, thus causing its several significant pathophysiologic effects on its host. One of which is the most common and dominant variant that is responsible for the massive increase in the past and current cases, as well as the leading cause of fatality as of to date is the variant D614G. With its mutations characterized by an amino acid substitution that replaces aspartic acid with glycine on codon 614, this resulted to an increased Spike protein incorporation into the virion, resulting to an increased viral stability and flexibility, thus conferring the variant a higher rate of infectivity as well as of enhanced pathogenicity. Another is the N501Y variant which was observed to have rapidly spread throughout the UK during early 2021. This variant has been correlated to have an ability for immune escape, whereby mutations on its three amino acids on the RBD region, specifically on the N501Y single residue has produced its neutralization resistance. Therefore, the variant's acquired capability to avoid antibodies has not only decreased its sensitivity for the host's immune response, but it also has compromised the efficacy of the existing vaccines that we now have today.

Conclusion: Several variants of SARS-CoV-2 were identified and categorized in this study and results show that mutations affect the pathogenic mechanisms of the virus which result in varying effects on transmission, pathogenicity, and mortality.

Key words: SARS-CoV-2, mutation, pathogenicity

* From Group 1, Section C, Department of Community and Family Medicine, College of Medicine

Since the emergence of COVID-19 due to SARS-CoV-2, the number of cases across the world has increased dramatically. Consequently, human-to-human transmission has been observed.[1] Due to the increasing exposure, hosts are able to adapt to defeat the virus, yet viruses are able to mutate to infect its host. For SARS-CoV-2 to infect hosts, it should withstand its new mutation and remain flexible during disease progression.[2] Experts believe that the mutations resembled the original virus strain that was first identified in Wuhan, China. Dr. Benjamin Neuman discussed that RNA viruses, including SARS-CoV-2, mutate slowly. Due to the emerging mutations of this virus, it has caused a rapid increase of infected people and longer duration in finding treatment of the virus. With these, the researchers conducted a review that will help future researchers to better understand the SARS-CoV-2 mutations.

Studies that describe mutations of SARS-CoV-2 are continuously emerging such as in a study by Maria Pacheti, et al.[2] who described 8 novel recurrent mutations of SARS-CoV-2 that were observed and studied from different countries around the world. Numerous similar studies have suggested that the virus is evolving with strains coexisting with one another despite having different mutation patterns. Further studies are recommended to closely monitor the mutations since COVID-19 is constantly evolving, producing new mutations, commonly in glycoproteins. In a study conducted by Muhammad Adnan Shereen, et al., the sequence analysis of SARS-CoV revealed a mutation at Spike Glycoprotein, and thus, it could be another suitable option to develop therapeutics targeting this specific part of the virus. However, other studies suggest that the virus shows little variability which is beneficial especially for researchers working on a possible vaccine since the characteristic of this virus results in new copies that are similar to the original one. Grubaugh, Petrone, and Holmes[3] recognized in their study that mutations are inevitable consequences which can further enhance the authors' knowledge on emerging outbreaks which will allow the recognition of their functions and monitor their spread. Despite the emergence of these new mutations, studies regarding how they affect its pathogenicity in terms of the manifestations, severity, and duration of illness are lacking. Thus, there is a great need to further advance this study.

A systematic review of the different mutations and an in-depth understanding of the pathogenic mechanisms of the virus as well as its causation on mortality is provided in this study. The establishment of reference for the pathogenicity of different SARS-CoV-2 mutations could benefit both the biological study of the virus and the diagnosis, monitoring, and treatment of the infection in the future as well as the possible development of vaccines and curative medications in general. In this study, the authors determined the different mutations of SARS-CoV-2 and their effects on pathogenicity and mortality. The authors also looked into the implications of the mutations on mitigations, quarantine measures, and therapeutics in general.

MATERIALS AND METHODS

The studies that were considered eligible were case studies and ecological studies that discussed the pathogenicity and mortality of COVID-19, and its mutations. Studies such as randomized controlled trials (RCTs) were excluded. Studies were limited to the English language. There were no restrictions regarding publication status.

The authors included studies with participants who: 1) were diagnosed with COVID-19, regardless of their age, sex, and ethnicity; 2) presented with a positive RT-PCR result for SARS-CoV 2; 3) were identified to have available genome sequence data and with the description of clinical course of illness.

They have included studies that discuss and compare how the strains differ from the reference Wuhan genome NC_045512.2. Studies on different populations such as Asia (specifically in Wuhan, China), Europe, and North America were also included.

They did a systematic literature search by searching the following electronic bibliographic English databases: PubMed, EBSCO, ScienceDirect, and Frontiers. They also examined the latest updates about COVID-19 mutation from the World Health Organization.

All databases were searched from the available data from inception through 2020. The search strategy included the following terms: "Covid-19 mutation and pathogenicity," or "SARS-CoV 2 mutation and pathogenicity" and Covid-19 mutation and mortality," or "SARS-CoV 2 mutation and mortality." Records were managed through Mendeley.

The authors did a keyword search from the online databases, assessed the eligibility of studies,

Table 1. Checklist of items to include when reporting a systematic review or meta-analysis

Section/topic	Item No.	Checklist item	Reported on Page no.
Title			
Title	1	Identify the report as a systematic review, meta-analysis, or both	Page 2
Abstract			
Structured Summary	2	Provide a structured summary including, as applicable, background, objectives, data sources, study eligibility criteria, participants, interventions, study appraisal and synthesis methods, results, limitations, conclusions and implications of key findings, systematic review registration number	Page 3
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	Page 4
Objectives	4	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	Page 5
Methods			
Eligibility criteria	5	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	Page 6
Information sources	6	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	Page 6
Search strategy	7	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	Page 6
Study records: a. Data management b. Selection process c. Datal collection process	8	 Describe the mechanism(s) that will be used to manage records and data throughout the review State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis) Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	Page 7
Data items	9	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	Page 7
Risk of bias in individual studies	10	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	Page 9
Data Synthesis	11	• Describe criteria under which study data will be quantitatively synthesized • If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I2, Kendall's τ) • Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression) • If quantitative synthesis is not appropriate, describe the type of summary planned	Page 7

and removed studies in duplicates. Following the removal of duplicates, studies were screened, and the potentially eligible full-text version studies were obtained based on the inclusion criteria. Any discrepancies in the suitability of a study for inclusion were analyzed among the researchers. Subsequently, the authors organized studies according to the study period, location, and the results included in the analysis and ensured not to use the same articles more than once in the analysis.

The pathogenicity of SARS-CoV-2 encompassed the pathophysiologic severity, clinical presentations, and illness' duration of COVID-19 patients. The assessment of risk bias was performed by the researchers of the study by utilizing the Cochrane Risk Bias Tool. Three items were assessed which included: partial results data, selective reporting, and further additional sources of bias. Every item was assessed as a low risk, high risk, or unclear. The overall risk bias of studies included was evaluated. The members solved all disagreements between reviewers. The authors have scrutinized the final articles to draw out the following data from each: the author, year of publication; study design and discussions that include pathogenicity and mortality in relation to SARS-CoV-2 mutations.

Data Items:

- Mutation- any deviation from the original Wuhan strain, may be classified based on categories if available or the sequence

- Pathogenicity - is the clinical course of the disease classified as mild, moderate, severe or critical as defined by the National Institutes of Health

RESULTS

The gathered data from the searched journals were collated and examined for possible variants. These variants were classified and further discussed under the following specific categories. These categories were labeled according to the viral structures that were affected due to the virion's own mutations. The categories were Spike protein, nucleocapsid phosphoprotein, NSP, ORF and micro-RNA respectively.

1. Spike Proteins

A. D614G

The D614G variant, undetected in January to February, progressed to uncommon in March, and associated with increasing cases in April and May conveying a greater transmission of the variant. It is currently observed to be the dominant variant observed in Europe, North America, and Southeast Asia and is the leading cause of fatality. The D614G variant is characterized by an amino acid substitution mutation wherein aspartic acid is replaced by glycine on codon 614. Higher viral loads are observed in patients infected with the variant, particularly in the upper respiratory tract suggesting higher rates of viral replication. Studies performed on animal models also showed faster transmission through aerosol and droplets. Various studies stated that the mutation allows the virus to have limited shedding of the spike S1 domain, increasing the incorporation of the S protein into the virion, thereby enhancing its transmission. Accompanying cysteine to threonine substitutions are also observed in some studies. The

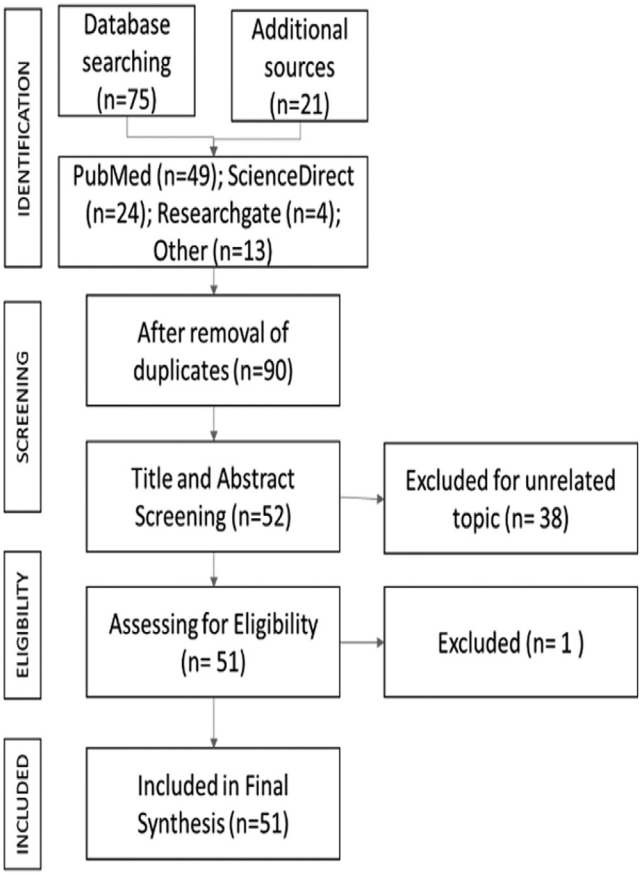

Figure 1. PRISMA flow diagram

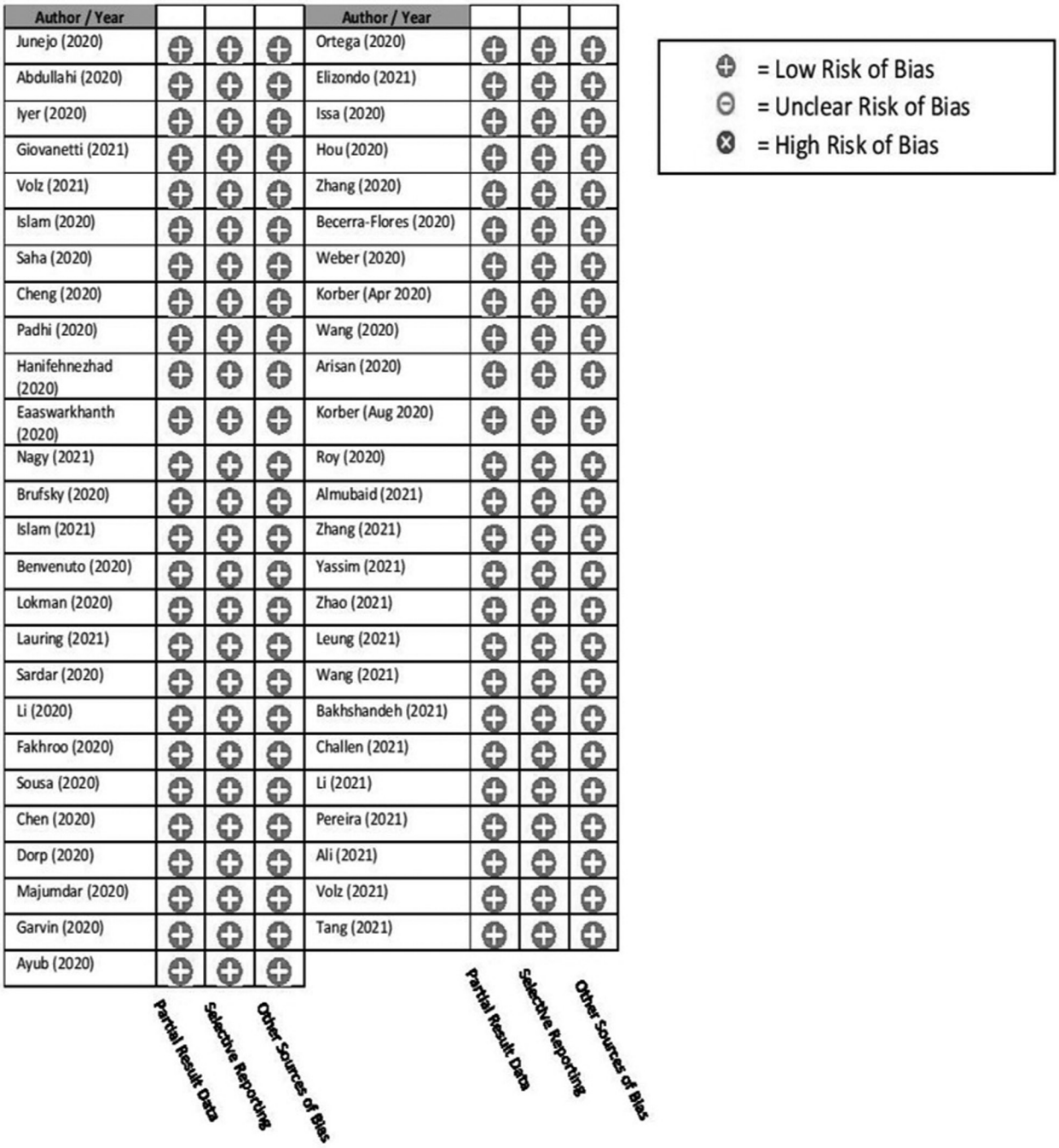

Figure 2. Risk bias assessment

high number of deaths due to the variant is a significant finding for all studies reviewed and is noted to be increased in association with comorbidities such as hypertension, renal failure and ICU admission. Although it is a common variant observed in deaths, no observed increase in pathogenesis of the disease is observed in animal models.

It is a common observation that mutations occurring in the surface proteins either allows the virus to escape immune response and enhance pathogenicity. D614G mutations increase viral stability and flexibility resulting in higher infectivity accounting for the massive increase in cases. Its dominance suggests that infectivity is increased in varying temperatures due to enhanced stability. No evidence of resistance to neutralizing human antibodies have been observed.

Table 2. Summary of results

Variants	Mutations	Effects on Pathogenicity and Mortality
Spike Proteins		
D614G	Amino acid substitution mutation wherein aspartic acid is replaced by glycine on codon 614	• Higher viral load • Higher infectivity
N501Y. 501Y.V2	Substitution mutation of asparagine to tyrosine at amino acid position 501	• Higher transmissibility • Better viral entry and infection • No significant enhancement in infectivity
D839Y/N/E and B.1.1.7 (VOC-202012/01)	Substitution of glutamate in the mutant D839E	• Increased strength of molecular association (increases pathogenicity) • Potential for increased infectivity
Others (S477N, N439K, P479S, D364Y, T478I, E484D, L452R, Q414E, N439K, V367F, V483A)		• Strains were directly proportional to the binding free energy (BFE) between RBD of individual strains and host cell expression of ACE2 • Positive BFE value indicates the mutation increases the free energy of the binding complex as a whole, therefore making the virus more infectious
MicroRNA		
miR8066	Point mutations	• Associated with the cytokine storm
miR-1307, miR-8066, and miR-129-2-3	Point mutations	• Immune evasion
ORF		
ORF3a a. Domain III b. Domain IV c. Domain V	 H93Y mutation in this domain was important due to the association to the loss of potassium channel and diminished proapoptotic activity Regulate the uptake and trafficking of 3a protein to the plasma or endomembrane Significant role in the transport of 3a protein and mutations in YXX motif were linked to aggregations of the protein (role in intracellular trafficking and surface transport.)	 • Increased mortality • Possible immune evasion and immoderate cytokine storm
ORF7a	Gene deletion	• Immune evasion and killing of infected cells once viral replication have been completed
ORF 8	Gene deletion	• Milder or less severe clinical course • Patients experience symptoms for longer period and may therefore increase possibility of transmission
Non-Structural Proteins (NSP)		
NSP1	285 transitions, 148 transversions, 271 missense, and 155 synonymous mutations	• Inhibit translation by binding to the host's 40S ribosome, and also inhibit IFN signaling
NSP2	Involved amino acid substitutions	• Inhibits two host proteins proinhibitin1 and proinhibitin2 to disrupt the cellular environment • Immune evasion
NSP 6	Missense mutation of threonine to isoleucine & valine to isoleucine	• Reduces stability of protein structures thus affecting pathogenicity • Compromises the ability of the autophagosomes to deliver viral components to lysosomes for degradation, favoring infection
NSP 12 (also called SARS-CoV-2-RdRp)	Mutation involving the substitution of aspartic acid to glycine in position 614	• Forms a complex with other NSP such as NSP7 and NSP8 • NSP14 or also known as ExoN helps in the fidelity synthesis by proofreading errors made by NSP12
NSP13	Relatively conserved	• Known to be an interferon antagonist which means that it suppresses the ability

		of the cell to recognize which it is a foreign invader
Nucleocapsid Phosphoprotein		
D614G variant	7 modifications in nucleic acid	• Significant mutation causing impact on B cell epitope enhancing binding and transmissibility
ORF8, NSP6, OFR3a, NSP4, NSP3, ORF6, ORF3A NSP7 variants		• 5 mutations have been identified to cause mild clinical presentation, while 15 produced severe manifestations • Increased pathogenicity
Others		
S motif	Six out of nine recombinants in the S motif	• Accounted to be vital for viral pathogenicity
ACE2, TMPRSS2, CD209, IFITM3 and MUC5B genes	Genetic polymorphisms	• ACE2, being the major viral receptor for COVID-19, was demonstrated to have polymorphisms that could have an effect on host susceptibility to COVID-19 • Variants with high TMPRSS2 expression were demonstrated to have a greater risk of severe conditions such as H1N1 and H7N9 influenza, and other respiratory disorders • Polymorphisms in CD209, IFTM3 and MUC5B genes can also affect the immune function of the host as well as repair mechanisms of the lungs
Hotspot Mutations in nucleotide sequences	Out of the 37 selected sequences, 12 of them were detected to have the GGG to AAC mutation and C-T point mutations	• Hypothesized to gain advantages during active viral replication

B. N501Y, 501Y.V2

The variant N501Y is the most recent variant detected in SARS-CoV-2. This was first found in South Africa when the second wave hit around mid-October 2020 and is observed to be the variant that rapidly spread throughout the UK early 2021. The N501Y variant is characterized by substitution mutation of asparagine to tyrosine at amino acid position 501 found in the Spike protein. Due to its rapid spread, most studies consider this variant with higher transmissibility as compared to other prevalent variants like D614G. One hypothesis of the variant's origin is that it may have been due to intra-host viral evolution from a prolonged infection in an immunocompromised host. It was found out that this variant had a higher binding affinity to human ACE2 due to the increased salt bridge electrostatic interaction between ACE2 and the RBD which results in a better viral entry and infection.

Despite numerous researches showing the increase in infectivity, one study mentioned that there is no significant enhancement in infectivity for variant N501Y when compared to D614G but they do have immune escape. It was found that the number of mutation sites is correlated with immune escape from an increasing number of monoclonal bodies. The RBD region has three single-residue variants, namely K417N, E484K and N501Y. These variants are responsible for the neutralization resistance seen in SARS-CoV-2. As with other variants, this discovery in mutation is a point of interest in the production of vaccines. The predominant variant found in the UK, specifically the N501Y.V2, also has the mutations of the three amino acids that decrease the sensitivity to neutralization, thus compromising the efficacy of existing vaccines.

C. D839Y/N/E and B.1.1.7 (VOC-202012/01)

Three other strains in Europe were found to have spike protein mutations in three residues, namely the D614G, A8314 and D839Y/N/E. The substitution of glutamate in the mutant D839E was said to increase the strength of the intermolecular association which increases its pathogenicity. Another mutation in the UK called Variant of Concern 2020 12/01 (VOC 202012/01) was found to have a phenotypic change to the cell binding mechanism that has the potential for increased infectivity but more importantly, the variant was said to affect the performance of polymerase chain reaction tests with an S-gene target. Out of 14 genetic mutations found in VOC 202012/01, 8 of them occurred in the RBD of spike protein which impairs

the detection of the S-gene. This also has a strong association with the B.1.1.7 variant lineage.

D. Others (S477N, N439K, P479S, D364Y, T478I, E484D, L452R, Q414E, N439K, V367F, V483A)

Epidemiologic and biochemical studies have shown that the pathogenicity of other SARS-CoV 2 strains were directly proportional to the binding free energy (BFE) between the RBD of individual strains and the host cell expression of ACE2. A positive value on the BFE indicates the mutation increases the free energy of the binding complex as a whole, therefore making the virus more infectious. In a study, among the 725 mutations on SARS-CoV-2 Spike protein, 89 were located on the RBD suggesting this area was more susceptible to mutations. A total of 194 residues were confirmed from the 89, and further studies from January early in the year 2020 have confirmed many mutations on the RBD.

Spike proteins play an important role in host cell attachment and entry. They also mediate invasion of the host cell by binding to angiotensin-converting enzyme receptors (ACE2) serving as a target for the development of antibodies, vaccines, and pharmaceuticals, making the spike protein a crucial factor in determining the future outcomes of the virus. Mutations, particularly in the spike protein, play a significant role in determining virulence and differentiation mechanisms of the virus; this could change the tropism of SARS-CoV-2, including new hosts and or increasing the viral pathogenesis. Yao, et al. report a direct link between the degree of genomic mutation to the variation in pathogenicity of SARS-CoV-2.

2. MicroRNA

Point mutation, particularly in the miRNAs in SARS-CoV-2 genomes has been found to affect the virus' pathogenicity. According to a study by Arisan, et al., some ssRNA may interact with miR processing pathways that can influence the host cell which can increase overall viral activity. The study aimed to clarify the role of potential miR-mimic sequences in the SARS-CoV-2 genome with their host target genes, which may propose a new perspective for antiviral strategies. In this study, they first identified human miRs that show sequence similarities to the SARS-CoV-2 genome, and their conservation ratio in SARS-CoV-2 isolates obtained from different geographical regions.

With the identified seven key miRs (miRs 8066, 3611, 5197, 3691-3p, 1307-3p, 1468-5p, and 3943-3p), researchers claim that viral miRs are critical in the immune evasion mechanisms which can affect host immunity-related gene regulation. The seven key miRs highlighted considerable differences between SARS-CoV-2 sequences, compared with the other viruses. As for instance, the miR8066 is found to be present in four different SARS-CoV genomes which consequently shows a strong association with target pathways that mediate clinical outcomes for SARS-CoV2 patients. More importantly, it is found to be associated with the cytokine storm in COVID19. Furthermore, three of the miRs identified (miR-1307, miR-8066, and miR-129-2-3) are found to be involved in several pathways that are subsequently associated with the host response against the virus, and its pathogenesis in the host's cells.

3. ORF

Genetic factors including the HLA type of patients are thought to play a role in relation to the disease severity and the clinical outcome of patients with COVID-19. ORF mutations are identified in 5 studies which include the variants ORF8 and ORF3a.

In one study, the 6 domains of the 3a proteins of the ORF3a variant were analyzed, and it was found that Domain III is cysteine-rich and has a potassium ion channel. Among the mutations associated with this domain, it was found that H93Y was particularly important due to previous association to the loss of potassium channel and diminished proapoptotic activity. Domain IV was found to have a caveolin-binding motif and potential interactions with caveolin-1 may regulate the uptake and trafficking of 3a protein to the plasma or endomembrane. Domain V was found to have a significant role in the transport of 3a protein and mutations in YXX motif were linked to aggregations of the protein. Hence, maintaining the said motif in all of the strains confirms its role in 3a intracellular trafficking and surface transport.

Moreover, another study showed that ORF3a mutations are associated with an increase in mortality rate in several countries. The study suggested that the functional involvement of signal transduction pathways are due to changes in phosphorylation sites and an indication of intact viral immune pathogenicity is the preservation of the N-myristoylation site. It was also mentioned that the N terminal peptide of

the ORF3a protein brought about protective humoral response involving several pathways including JAK STAT, chemokine and cytokine related pathways in patients infected with the disease.

A substantial reduction in the intensity of B cell epitope is suggestive of a possible immune evasion linked to the amplified virulence of SARS-CoV-2 having the ORF3a variant. Immoderate cytokine storms are also thought to have a potential association in the increase in mortality rate.

For the ORF8 gene, it was found that large deletions of this accessory gene increase the opportunity for transmission. This particular accessory gene is lacking in several variants reported worldwide indicating that it is not essential for the viral replication of the virus. The relevance of the ORF8 gene is associated with its involvement in the modulation of host infected cell metabolism and innate immunity evasion. It was also said that gene deletion results in milder infections or less severe clinical course due to the less systemic release of cytokines. However, patients experience symptoms for a longer period and may therefore increase the possibility of transmission.

Through deletions and targeting of the ORF10 to the heme protein, disease manifestations and clinical outcomes have been associated, suggesting further the role of deletions in the virus' virulence and pathogenicity. ORF7a on the other hand was explained to be able to assist the virus in host immune evasion and killing of infected cells once viral replication has been completed.

4. NSP

A total of 9 studies have discussed mutations regarding the non-structural proteins or NSP. Structures that were affected include NSP 1, NSP 2, NSP 3, NSP 6, NSP7, NSP 8, NSP 12 and NSP 13. It was stated that the NSP 1 mutations include 285 transitions, 148 transversions, 271 missense, and 155 synonymous mutations while NSP mutations include 981 transitions, 503 transversions, 996 missense, and 477 synonymous mutations. NSP 2 includes the deletion of ORF8, ORF7a and spike protein. Together with NSP 3, NSP 2 also had mutations that involved amino acid substitutions. There was a missense mutation of threonine to isoleucine & valine to isoleucine involving NSP 2 and NSP 6. There was also lower stability due to Leucine/ Phenylalanine residues at position 3691 and 9659 in the mutation of NSP. On the other hand, for NSP 12 there was a mutation involving the substitution of aspartic acid to glycine in position 614.

Globally, NSP1 and NSP2 were the most known mutations among the non-structural genes. NSP 1 is known to inhibit translation by binding to the host's 40S ribosome, and also inhibits IFN signaling while NSP 2 inhibits two host proteins proinhibitin1 and proinhibitin2 to disrupt the cellular environment. These mutations help in the virus in innovating novel molecular routes in order to evade their host's immunogenic response. NSP 2 and NSP 3 should also be noted because it has an important role in virulence and differentiation mechanism of the virus. NSP2 has mutations in the amino acid level of RNA primase and S protein because of the emergence of other critical mutations. Since the deletion of NSP2, it was found out that there was a probable role in viral pathogenesis in the aspartic acid that is situated on the protein surface.

Mutation in NSP6 reduces the stability of protein structures thus affecting the pathogenicity. There was more stable binding in the endoplasmic reticulum of the protein because of the availability of multiple phenylalanine residues in the outer membrane. It compromises the ability of the autophagosomes to deliver viral components to lysosomes for degradation. Because of this, it favors coronavirus infection. NSP6 is also known to have a role in viral autophagy. However, its role to SARS-CoV-2 infections needs to be further studied in order to assess this NSP mutation.

NSP12 is also called SARS-CoV-2-RdRp. It forms a complex with other NSP such as NSP7 and NSP8. NSP14 or also known as ExoN helps in the fidelity synthesis by proofreading errors made by NSP12. Lastly, compared with other protein regions in the ORF1ab, NSP 13 is relatively conserved. This property makes it a good candidate for targets of medication and treatment of COVID-19 infection. NSP13 is also known to be an interferon antagonist which means that it suppresses the ability of the cell to recognize which is a foreign invader.

In summary, mutations in the NSP had been focusing on its increasing viral pathogenicity. Due to this, the NSP mutation has proven that it decreases the immune system of its host therefore increasing the manifestations. There is still need for further studies in order to specify which manifestation it produces or if NSP adds manifestation compared to other mutations.

5. Nucleocapsid Phosphoprotein

Mutations in the nucleocapsid phosphoprotein were associated with both mild and severe COVID-19 cases. To facilitate which variants harbor these life-threatening mutations, 2 studies have identified a total of 8 variants and these were D614G, ORF8, NSP6, OFR3a, NSP4, NSP3, ORF6, ORF3A NSP7, respectively.

For the D614G variant, its sequence possessed an identical similarity of 99.9% to that of the Wuhan-Hu-1 virus, for which it exhibited modification in 7 nucleic acid locations (T17432C, T18870C, T25556G, T26728C, G28874A, G28875A, G28876C) in its genome. In light of classification under this category, there were two amino acid replacements located in the nucleocapsid protein (RG203KR). It also carried a significant mutation in (A23403G), wherein it caused an impact on B cell epitopes, thereby enhancing its binding affinity, and thus increasing its transmissibility.

As for the ORF8, NSP6, OFR3a, NSP4, NSP3, ORF6, ORF3A NSP7 variants, these strains have acquired 5 mutations, all of which involved certain genetic changes in the ORF8 protein which is on L84S, in NSP6 protein located at L37F, in the ORF3a protein placed on G196V, in NSP4 protein positioned at F308Y and the in nucleocapsid phosphoprotein situated on S197L, and it was noted that these nucleotide alterations caused only mild clinical presentations. However, 15 mutations were recognized to produce severe manifestations, and these were in the surface (S) glycoprotein which was set on sequences L54F, D614G and V1176F, in the RNA dependent RNA polymerase specifically at sequences A97V and P323L, in ORF3a protein which occupied sequences Q57H and G251V, in ORF6 protein lodged at sequences P13L, S194L, R203K, G204R and I33T, and in nucleocapsid phosphoprotein stationed at sequences S1197R and T1198K mutations in the NSP3 protein and I292T, therefore such modifications contributed to the increased pathogenicity of the virus.

6. Others

In addition to the previous categories, there were a total of three studies which we categorized under "others" because each study has discussed structures in the SARS-CoV-2 that were not discussed in the previous categories.

A. S motif (Replacement and recombination of nucleotides)

Association between SARS-CoV-2's pathogenicity and its structural replacement and recombination of nucleotides, particularly in the S motif, has been identified in a study by Junejo, et al.[7] The researchers identified nine putative recombinants in SARSCoV-2, where they hypothesized that SARS-CoV-2 is the recombinant of SARS and SARS-like CoVs. The strains of Betacoronavirus, MERS-CoV, HKU1, OC43, and SARS-CoV have been regarded with severe human infections. One hypothesized mechanism of viral development is the replacement and recombination of nucleotides.

With the researchers' account to give a broader perspective of SARS-CoV-2's origins, genes and its variations, pathogenesis and clinical manifestations, they discussed the comparison between SARS-CoV-2 membrane, spike protein, envelope, and nucleocapsid. Accordingly, six out of nine recombinants in the S motif have been accounted to be vital for viral pathogenicity.

B. Genetic Polymorphisms (ACE2, TMPRSS2, CD209, IFITM3 and MUC5B genes)

Several host factors have demonstrated to affect the pathogenicity of the COVID-19. According to Iyer, et al.[8], there are polymorphisms in the host genes that might have some effects on the pathogenicity of the virus, that is ACE2, TMPRSS2, CD209, IFITM3, and MUC5B. ACE2, being the major viral receptor for COVID-19, was demonstrated to have polymorphisms that could have an effect on host susceptibility to COVID-19. ACE2 functions in the renin-angiotensin pathway and polymorphism with such genes could alter the expression of ACE2 that would have several effects such as hypertension, diabetes, heart disease and stroke. Another gene that might be involved in increasing pathogenicity is TMPRSS2 involved in the viral fusion with the cell membrane. Variants with high TMPRSS2 expression were demonstrated to have a greater risk of severe conditions such as H1N1 and H7N9 influenza, and other respiratory disorders that may also confer risk of infection with COVID-19. Another gene mentioned in the study is the CD209 gene that encodes DC-SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin) and is involved in innate immunity and antiviral defenses. Polymorphism in this gene might

have an implication on susceptibility or resistance to infectious diseases. IFITM3 and associated interferon-induced membrane protein that inhibits the entry of viruses into the host by preventing fusion, could also be affected by several polymorphisms. Lastly, MUC5B that are highly glycosylated macromolecular components of mucus secretions, is a gene that encodes a member of the mucin family of proteins. Several polymorphisms in the MUC5B gene had been speculated to alter the pathogenesis of idiopathic pulmonary fibrosis, whereas a promoter variant was found to be associated with diffuse panbronchiolitis. Higher expression of MUC5B also enhances damage by impairing mucus clearance and repair mechanisms of the lungs.

C. Hotspot Mutations in Nucleotide Sequences

The emergence of de novo hotspot mutations has occurred after the transmission of SARS-CoV-2 outside of China which permitted its replication worldwide. A study by Weber, et al.[4] was conducted through analyzing the nucleotide sequence data found between December 2019 up to the end of May. This allowed researchers to compare 570 nucleotide sequences of SARS-CoV-2 genomes from China, Europe, the US, and India in reference to the sequence of the Wuhan isolate. Out of the 37 selected sequences, 12 of them were detected to have the GGG to AAC mutation and C-T point mutations were noted in positions 28,854 and 28,863. It was hypothesized that there is a significance among the mutations especially those noted in different populations and have advantages during active viral replication.

During the worldwide spread of COVID-19, ten mutations were found in approximately 80% of the viral genomes, leading to amino acid exchanges in areas where replication is said to be relevant. The mutation AGG GGA to AAA CGA, which is located at the sequence position 28,881, has changed the amino acid sequence of nucleoprotein N from Arginine-Glycine to Lysine-Arginine. The frequency of this mutation was determined from the isolates. However, during that time it was noted that the possible impact of these mutations on pathogenicity will require further study.

Overall, these mutations have increased the viral effects on its human host- whether in mild or in severe form. Limiting their spread however generally terminates future genetic alterations, thus stopping further development from which causes the host's heightened pathognomonic disease processes and progression, of which all are relatively inherent from the virus itself.

DISCUSSION

1. Effects of Mutation in the General Population

In the analysis of different studies, the authors were able to categorize the mutations of SARS-CoV-2 by their location. Significant variations studied fell either into mutations found in the spike protein, non-structural protein (NSP), ORF protein (open reading frame), nucleocapsid phosphoprotein and miRNA.

The spike protein found on SARS-CoV-2 mediates host cell attachment and entry of the virus by binding to the angiotensin-converting enzyme (ACE2) receptors which is why it is a site of importance. Among all the mutations, the receptor binding domain (RBD) in the spike protein is the most commonly studied site for variations. A number of collected studies were on D614G mutation. It was characterized with a higher viral load, enhanced pathogenicity and increased transmissibility. Two more recent variations found late in 2020 are N501Y in South Africa and B.1.1.7 or VOC-202012/01 lineage found in the UK. Both these variants have similar mutations and have already been detected in different countries early 2021 due to its increased ability to infect others. For the VOC-202012/01, however, it has a novel characteristic of affecting PCR results by impairing the detection of the S-gene. Studies in the spike protein have big implications in the target for vaccines since it modulates viral entry. Aside from increased transmissibility, there is proof that some of these mutations enhance immune escape which makes them resistant to neutralizing antibodies. Such mechanism is found in mutations such as D614G and N501Y.V2.

Mutations on the NSP proteins also display an escape of the human immune system, but mostly studies focused on the increased viral pathogenicity in NSP mutations. NSP1 and NSP2 were identified as the two most likely NSPs to mutate in SARS-CoV-2 which manifests with a capability to evade host immunogenic response. Additionally, when it comes to virulence and differentiation mechanism, NSP 2 and NSP 3 play a significant role for SARS-CoV-2. NSP6 was found to have a more stable binding of the protein between this region and the ER that favors CoVid-19 infection. Other NSP proteins mentioned are NSP12, NSP7, NSP8 and NSP 14 were associated with faulty

proofreading ability and therefore, result in a higher rate of mutations. NSP13 was seen as a possible target for drugs as it has been found that its mechanism is to suppress primary interferon production and signaling increasing the virus' virulence and transmission.

There were 5 studies found on the ORF mutations and they all discussed ORF8 and ORF3a specifically. ORF3a in particular was found to be related to a higher mortality rate in patients due to possible ability to evade the immune system and enhanced virulence of SARS-CoV-2. It has been shown that the ORF3a protein induces a protective humoral response involving the JAK-STAT, chemokine and cytokine-related pathways in infected patients. On the other hand, large deletions of the ORF8 gene were found to have an increased opportunity for transmission. This gene is involved in regulating cell metabolism of infected hosts and immunity evasion. Upon deletion of the gene, cases resulted in milder and less severe clinical course.

The last two categories are nucleocapsid phosphoprotein and miRNA. Mutations in the nucleocapsid phosphoprotein were both linked to mild and severe cases. The S197L in nucleocapsid phosphoprotein was noted to cause only mild clinical manifestations while I292T was found to cause severe manifestations and contributed to the increased pathogenicity of the virus. Other nucleocapsid proteins noted were RG203KR and A23403G that has an effect on B cell epitope which increases the transmissibility of the virus. For miRNA, there were seven key miRs (miRs 8066, 3611, 5197, 3691-3p, 1307-3p, 1468-5p, and 3943-3p) that the study claimed to be critical in the immune evasion of the virus. Additionally, miR8066 shows a strong association with the cytokine storm in COVID-19 and is involved in pathways that mediate the clinical outcome of patients. Three of the miRs identified miR-1307, miR-8066, and miR-129-2-3 are found to be involved in several pathways that are related to the pathogenesis of the virus in host cells.

Overall, the effects of the mutations that the studies have pointed out is that each variation has an impact on the pathogenic mechanism of the virus which increases its transmission, pathogenicity and mortality. Multiple studies have also found that most of the mutations contribute to the immune evasion of the virus which can exacerbate the course of the virus and can provide significant information to help in the development of an efficient vaccine or drug.

2. Implications on Mitigation Efforts and Quarantine Measures

The widespread pandemic can be attributed to the exponential growth and mutations of the virus. Quarantine measures implemented worldwide are targeted in controlling the rate of which the virus spreads, wherein speed and scale of mitigation efforts are key to controlling the growth of the virus given the various mutations it has already undergone. These mutations have caused varying symptoms in the general public, from being asymptomatic to patients experiencing respiratory distress, all of which are dependent on the integrity of the host's immune system. Given that these mutations are the core of the virus' infectivity, specifically in the spike proteins affecting the Receptor Binding Domain (RBD), researchers have directed their efforts in developing treatment targeting these key features; using the virus' pathogenic features to create vaccines that will address the concerned parts of the virus themselves.

The World Health Organization and the Centers for Disease Control and Prevention have indicated that the more a person interacts with different people, and the longer and closer the interaction, the higher the risk is of COVID-19 in spreading; with this, efforts that attempt to prevent the spread of the virus have been implemented globally: Mass screening, Social Distancing Protocols with Contact Tracing, Limited closure of leisure facilities such as malls, Limited closure of educational institutions and the Isolation of Confirmed Cases. Attention should be given to individuals with the highest risk of contracting infection when adjusting or determining strategies of mitigating SARS-CoV-2. Additionally, these mitigation strategies can be conducted in different scales depending on the local situations since every community is unique in itself; what works for one community may or may not work for another.

3. Potential Implication on Interventions and Vaccines

Due to the detection of new variants of SARS COV 2, concerns about potential virus escape from the immunity given by vaccines have been raised. According to the World Health Organization, changes in the virus should not make vaccines completely ineffective.[5] Protection against the new variants from the currently approved vaccines is still expected as these can elicit a broad immune response.

4. Limitation

The accelerated expression of SARS-CoV-2 stipulates numerous questions for scientists and researchers worldwide. Viral mutations by other mechanisms are still unclear. It is important to take note that our study is focused on the effect of mutation in terms of its pathogenicity. Another limitation includes how these new variants affect vaccine development in terms of efficacy. Additionally, as the proportion of cases in various segments across populations vary, as well as both testing priorities and vaccination protocols differ in each country, the emergence of signals affected by operational changes rather than a biological shift in the virus itself has to be taken into account.

5. Recommendation for Future Researchers

After one year of having SARS-CoV-19, there were already several mutations that had been found in different countries. This study provides a broad knowledge of the current mutations and its correlation to pathogenicity up to date. This study can help future researchers in categorizing different mutations easily. However, due to lack of time and resources, there were gaps noted. They recommend future researchers to have more time in gathering studies. Due to lack of time, only a maximum of three reads per article were done. They suggest that every researcher should read each article in order to maximize the content of each. For a more focused study, they recommend future researchers to set ample time to review the journals.

CONCLUSION

Several SARS-CoV-2 variants were identified and categorized in this study and results show that mutations affect the pathogenic mechanisms of the virus which results in varying effects on transmission, pathogenicity and mortality. Quarantine measures are implemented to control the rapid spread of the virus across the globe. Mitigation strategies such as mass screening, social distancing protocols with contact tracing, limited closure of leisure facilities and educational institutions and the isolation of confirmed cases are efforts aimed at limiting close contact among individuals as an attempt to limit the viral spread. Numerous researchers have directed their efforts in interventions and vaccine development that target key features involved in viral mutation. According to the World Health Organization, vaccines are able to elicit a broad immune response, therefore currently approved vaccines are expected to provide immunity against new variants of SARS-CoV-2. The present study focused on the effects of mutation in terms of the pathogenicity of the virus. However, the exact mechanism of viral mutation and the effects of variants to vaccine development and efficacy still remains unclear. The authors recommend future researchers to allocate more time and have a more extensive review of available studies.

REFERENCES

1. Phan T. Genetic diversity and evolution of SARS-CoV-2. Infection, Genetics and Evolution 2020; 81: 104260. https://doi.org/10.1016/j.meegid.2020.104260
2. Pachetti M, Marini B, Benedetti F, Giudici F, Mauro E, Storici P, Masciovecchio C, Angeletti S, Ciccozzi M, Gallo RC, Zella D & Ippodrino R. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. J Translational Med 2020; 18(1): 1–4. https://doi.org/10.1186/s12967-020-02344-6
3. Grubaugh ND, Petrone ME & Holmes EC. We shouldn't worry when a virus mutates during disease outbreaks. Nat Microbiol 2020; 5: 529–30. https://doi.org/10.1038/s41564-020-0690-4
4. Weber S, Ramirez C & Doerfler W. Signal hotspot mutations in SARS-CoV-2 genomes evolve as the virus spreads and actively replicates in different parts of the world. Virus Research 2020; 289: 198170. https://doi.org/10.1016/j.virusres.2020.198170
5. WHO. (2020, August). Coronavirus disease (COVID-19) (Situation Report – 209). https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200816-covid-19-sitrep-209.pdf?sfvrsn=5dde1ca2_2
6. Arisan ED, Dart A, Grant GH, Arisan S, Cuhadaroglu S, Lange S & Uysal-Onganer P. The prediction of miRNAs in SARS-CoV-2 genomes: hsa-miR databases identify 7 key miRs linked to host responses and virus pathogenicity-related KEGG pathways significant for comorbidities. Viruses 2020; 12(6): 614. https://doi.org/10.3390/v12060614
7. Junejo Y, Ozaslan M, Safdar M, Khailany RA, Rehman S, Yousaf W & Khan MA. Novel SARS-CoV-2/COVID-19: Origin, pathogenesis, genes and genetic variations, immune responses and phylogenetic analysis. Gene Reports 2020; 20: 100752. https://doi.org/10.1016/j.genrep.2020.100752
8. Iyer GR, Samajder S, Zubeda SS, Mali V, Sharma A, Abbas NZ, Bora NS, Narravula A & Hasan Q. Infectivity and progression of COVID-19 based on selected host candidate gene variants. Frontiers in Genetics 2020; 11. https://doi.org/10.3389/fgene.2020.00861