(RESEARCH ARTICLE)

Check for updates

SARS-CoV-2 variants in Vietnam: A comprehensive analysis of nucleotide changes in the spike gene

Quan Ke Thai [1, *] and Phuoc Huynh [2]

[1] Faculty of Natural science Education, Saigon University, 273 An Duong Vuong, Ward 3, District 5, Ho Chi Minh City, Vietnam, 700000.
[2] VNU HCMC University of Science, 227 Nguyen Van Cu, Ward 4 District 5, Ho Chi Minh city, Vietnam, 700000.

Abstract

SARS-CoV-2, causing the COVID-19 pandemic, has claimed millions of lives worldwide. SARS-CoV-2 has a high mutation rate in its genome, resulting in thousands of variants. The success of SARS-CoV-2 is attributed to a mutation in the S gene, which encodes the spike protein that interacts directly with hACE2. Mutations in this gene are known to increase the transmission rate and the ability to escape the immune system of the virus. This study focused on analyzing the nucleotide changes in the S gene of the SARS-CoV-2 variants that appeared in Vietnam. The results showed that Vietnam recorded many VOC variants, including Alpha, Beta, Delta, and Omicron, with Delta and Omicron being the most prevalent. The S1 region of the S gene had the highest mutation rate, with missense and C to T mutations being the most common. The NTD region contained all deletion and insertion mutations, with the nucleotide at 22198 to 22206 being the hotspot for insertion. The RBD region showed a positive selection during evolution, indicating that it had undergone harsh missense mutation. Overall, this study demonstrates that the S gene in Vietnam has high haplotype and nucleotide diversity. The Omicron variant in Vietnam had the highest nucleotide, haplotype diversity indexes, and average number of mutations in the S gene. These findings provide insights into the genetic diversity of SARS-CoV-2 variants in Vietnam and the impact of the S gene mutations on the evolution of the virus.

Keywords: Spike gene; SARS-CoV-2; Genetic diversity; Vietnam; COVID-19

1. Introduction

The outbreak of COVID-19 pandemic because of SARS-CoV-2, which is belong to Betacoronavirus infected in human has taken away millions of lives. SARS-CoV-2 contain the single strand RNA+ in core that is encoded for the four major structure proteins: Spike (S), Envelope (E), Membrane (M), N (Nucleocapsid) proteins. Basicly, the coronavirus or RNA virus had the high mutation [1, 2], that is the main reason why SARS-CoV-2 spawned thousands of variations. Because of high population diversity, therefore, WHO was classified the variants of SARS-CoV-2 into three group: VUM (Variant Under Minitor), VOI (Variant Of Interest) and VOC (Variant Of Concern). Therein, the VOCs variant is the most dangerous which are contained advantage phenotype when compared to wildtype. Heretofore, Alpha (B.1.1.7) emergence as the first VOC variant in global which rapid spread in United Kingdom in the end of 2020 [3]. Following, the novel VOCs were Beta (B.1.351), Gamma (P.1) appeared in South Africa and Brazil, respectively also ahead of transmission in each countries [4]. More remarkable, the emergence of the Delta variant has changed the global pandemic landscape. Delta had the transmission rate higher Alpha about 63% to 167% [5], therefore, this variant rapid became the domant variant in worldwide. Moreover, Delta successful intrusion and transmission over 200 countries/territories that is more dangerous than previous VOCs variant [6]. Interestingly, the successful of Alpha, Beta, Gamma, or Delta were considered by mutation in spike protein. The previous studies showed that mutations appearing in the RBD – Receptor Binding Domain (L452R, T478K, E484K/Q, Q498R, N501Y) enhance the affinity to the receptor [7-10].In addition, P681H/R

---

* Corresponding author: Quan Ke Thai ORCID: 0000-0002-2264-2579

mutations in FCS – Furin Cleavage Site are the primary mutation support for the success of the Alpha and Delta variants [11, 12]. Furthermore, spike protein located on the envelope, therefore, it is represents as an antigen of SARS-CoV-2. Thus, this protein is a crucial candidate for vaccine strategy and acceptance represented by antibodies [13].

In SARS-CoV-2 genomic, spike protein is encoded by the Spike (S) gene. Therefore, missense nucleotide mutations in the S gene led to amino acid substitution in this protein. Correspondingly, the S gene had the highest mutation rate and was positively selected during evolution compared to other structure genes [14, 15]. In just one year, by the end of 2020, the S gene has appeared with more than 4000 mutations [16], indicating that this gene region will contribute positively to the transmission of SARS-CoV-2. Indeed, recent research has also shown that the S gene helps SARS-CoV-2 adapt to the host cell [17, 18], especially in the appearance of a new VOC variant Omicron in December 2021 complicates the COVID-19 epidemic. Omicron appears with more than 30 amino acid substitutions on the spike protein. Hundreds of sub-Omicron variants have appeared, including dangerous new variants such as BA.2.75, BA.4, BA.5 and XBB, containing characteristic mutations on spike protein [19]. Thus, examination of the nucleotide drive in this gene is very important [20]. Understanding this situatition, PANGO Linages developed the secondary database of S gene sequences for tracking the mutations [21]. Currently, WHO also uses the mutation in the spike protein for classifying SARS-CoV-2 variants. Therefore, a detailed evaluation of nucleotide changes in the S gene as soon as possible helps to point out the evolutionary direction and response to the host cell. Information extracted from the diversity and dynamics of the S gene will inform further studies – particularly in support of epidemiology.

Vietnam is one of the few countries that successfully controlled in the early stage of the COVID-19 pandemic [22]. However, the appearance of Delta in the fourth wave has made a single worse COVID-19 situation. Our previous studies on the genetic diversity of SARS-CoV-2 in Vietnam have shown the migration of numerous variants [23, 24]. Delta and Omicron variants are the ones that show a high degree of diversity in Vietnam. Over a year since our report, Vietnam has recorded many new dangerous variants, such as BA.2.75, BA.4, BA.5, and XBB. Therefore, evaluating and investigating novel mutations in the S gene region is necessary for the present complicated situation.

2. Material and methods

2.1. Data collection

The genome and subsequent information on SARS-CoV-2 isolated in Vietnam were redeemed from the GISAID database. To extract the S gene, we used the SARS-CoV-2 full-length MSA tool of the MAFFT [25] server based on the S gene of the reference sequence (accession number NC_045512.2) in GenBank. All the S gene splits will also be aligned in MAFFT for downstream analysis. We used the in-house developed software to identify the S gene mutation. Then these genes were haplotype numbering by DnaSP software ver 6.12.03 [26].

2.2. The phylogenetic tree and haplotype network analysis

The IQtree software ver 2.2.0 was used to build the Maximum-Likelihood (ML) phylogenetic tree [27]. In order to build the ML tree, we chose the GTR+F+I+G4 method as the best model based on the BIC score of IQtree and setting the reference sequence as a root of the tree. A 1000 bootstrap values were used to provide consistency of the ML tree. The Figtree ver 1.4.4 (http://tree.bio.ed.ac.uk/software/fitree/) was used to visualize the output tree of IQtree. Then, PopART [28] software based on the Minimum-Spanning method built the haplotype network of the S gene in Vietnam.

2.3. Examination of genetic diversity

The genetic diversity measures, including the nucleotide diversity, haplotype diversity, average nucleotide difference distance of each variant, and the comparison nucleotide difference distance of each variant, were determined by Arlequin ver 3.5 [29]. Then, our study used the t-test in R packages for statistic analysis.

3. Results

3.1. Haplotype diversity

To 10th March 2023, we collected 8372 S gene sequences in Vietnam that were classified into the 1978 haplotype. Analyzing the timeline of variants in Vietnam indicated each variant displayed for each COVID-19 outbreak wave. The A – sub-A lineages and B – sub-B lineages dominate in the 1st and 2nd waves, following the Alpha variant as the primary cause of the 3rd wave. The 4th wave recorded two VOC variants that are Delta's emergence in an early stage, and Omicron variants replaced this variant in March 2022 (Figure 1A). The haplotype diversity analysis showed that Delta and

Omicron variants had a high haplotype diversity index and numerous haplotypes (Figure 1B). This result displayed that Delta and Omicron were the significant variants contributing to Vietnam's genetic diversity. Besides, we also noted the low haplotype diversity of B and sub-B Lineages (Figure 1B) displayed for the limit of nucleotide changes in this variant in Vietnam (Figure 3). A and sub-A Lineages had a high level of haplotype diversity, but they contained a low number of haplotypes. Extracting output of the nucleotide pairwise differ of each haplotype at 3.2 (Figure 3) and timeline of variants (Figure 1A) indicated these lineages entry in Vietnam twice in the 1st and 3rd wave. In addition, the metadata of Vietnam showed that two sequences (EPI_ISL_16034628 and EPI_ISL_16034629) were classified into Beta variant (B.1.351) that were not recorded in previous studies. Especially these sequences were collected on 08/05/2021, but that is not submitted to GISAID until 17/12/2022 – more than one year.
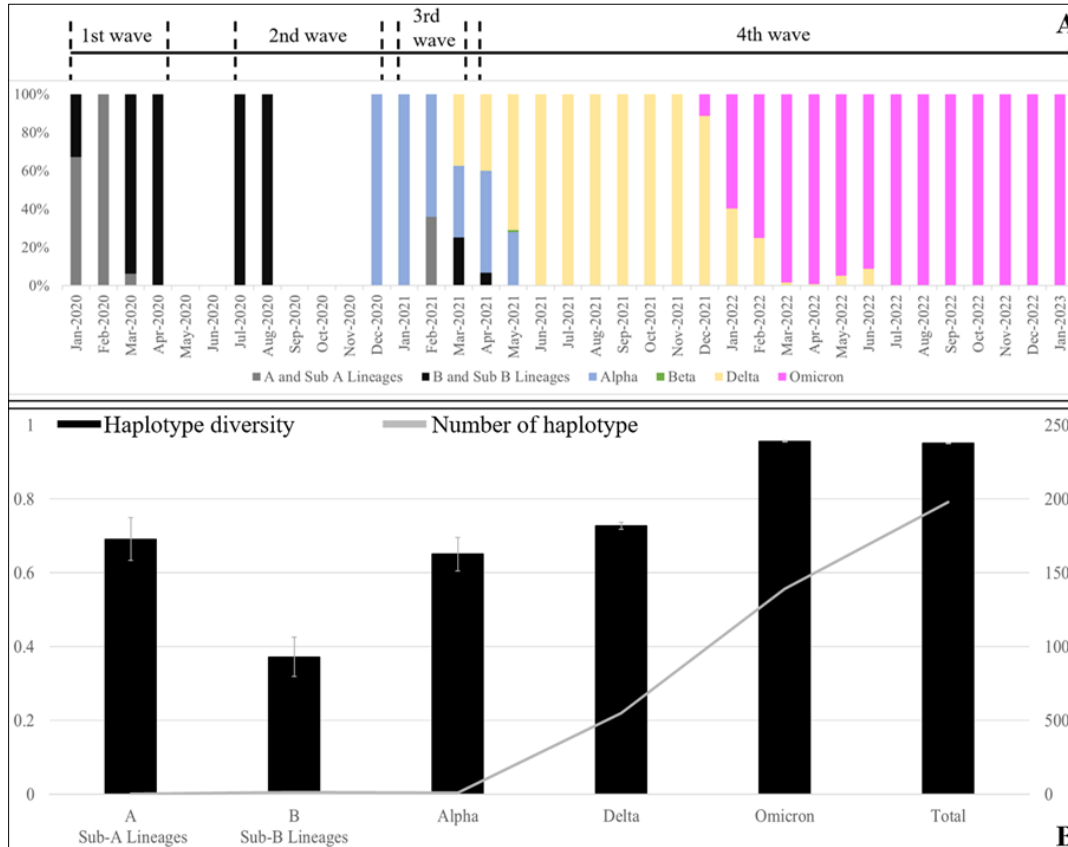

(A) The color grey, black, light blue, green, yellow and pink were presented for A and Sub-A Lineages, B and Sub-B Lineages, Alpha, Beta, Delta and Omicron, respectively; the timeline was exhibited column for the frequency of the variant in the month. B) The haplotype diversity of the S gene in Vietnam. All error bars in this figure show the SD: standard deviation.)

Figure 1 The timeline and haplotype diversity of SARS-CoV-2 in Vietnam

A phylogenetic tree, phylogenetic network, and haplotype nucleotide difference analysis were provided to examine each variant's impact on population diversity. Our ML tree based on the S gene sequences, built by IQtree software, indicates the clear split of each variant (Figure 2A). The Alpha variant was split from B Lineages, and the Omicron variant was cut off from Delta with bootstrap values 84 and 100, respectively. On the ML tree, the correlation between high haplotype diversity and the complexities of Delta and Omicron variants was displayed (Figure 2A). On the phylogenetic network, the Delta haplotype differed by an average of 2.1 nucleotides (Figure 3), and the haplotype had close related (Figure 2B). However, it had numerous polymorphic sites that implication for this variant, which had a long time circulating in Vietnam (Figure 1A)) that is enough to store novel mutations. The phylogenetic network also demonstrated very complicated nucleotide changes in the S gene of Omicron variants. While A Lineages, B Lineages, Alpha and Delta variants had the close related haplotype in phylogenetic network analysis (Figure 2), the Omicron variant had the high distance. Therefore, Omicron not only rapidly transmission in Vietnam but also introduces new sub-variant with high distance genetically into the country.
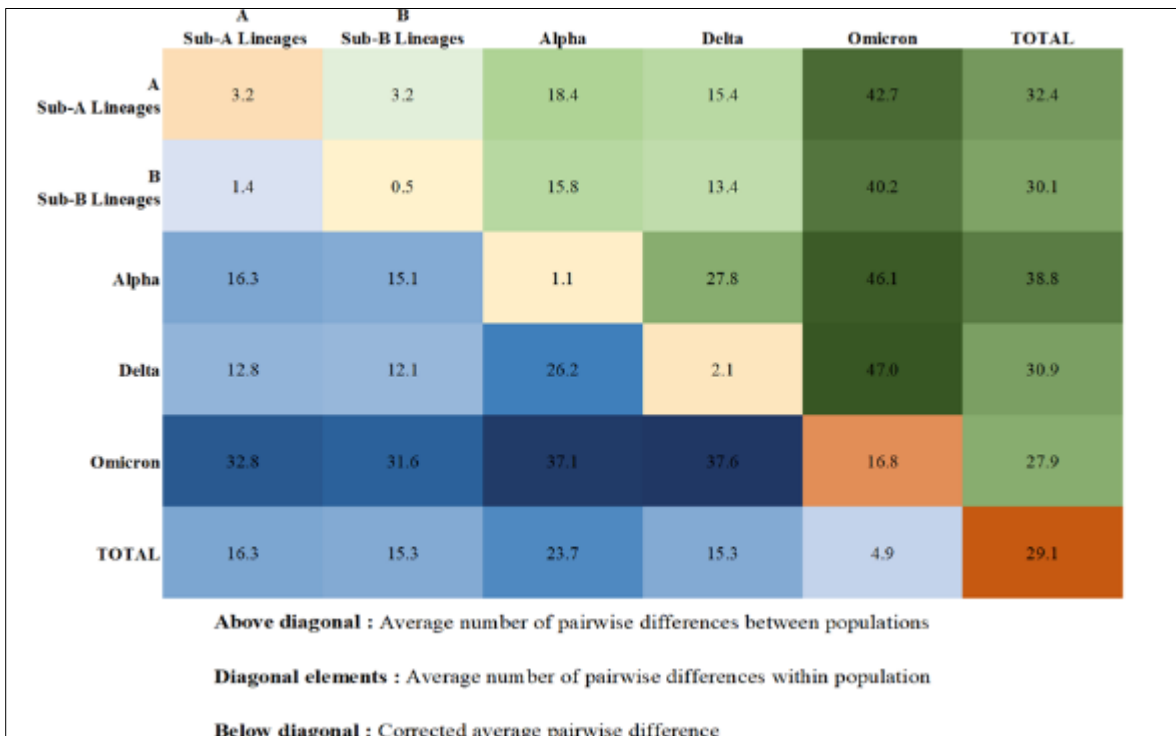
The analyzed nucleotide difference displayed the high impact of Delta and Omicron variants on population genetic diversity. The B Lineages and Alpha variants recorded the lower nucleotide distance (Figure 3) compared with the A Lineages, which had multi-time entries in Vietnam. Parallelism, the variants comparison analysis indicated that the

population diversity was closer to that of the Omicron variant, demonstrated by nucleotide pairwise difference and nucleotide population distance between Omicron and data analyzed (Figure 3). Generally, the S gene in Vietnam had a low level of haplotype diversity and limit of nucleotide changes that were only increased by the entries Delta and Omicron variants.
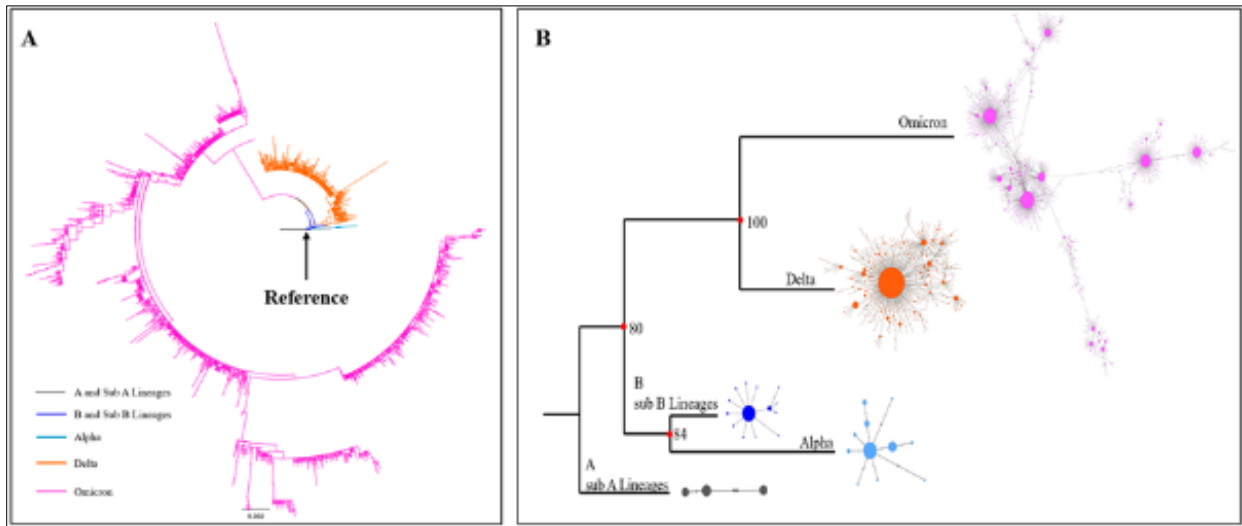
(A) The ML tree was built by IQtree software with 1000 bootstrap values. Each colour coloured each variant on this tree (grey, blue, light blue, orange and pink presentation for A and sub-A Lineages, B and Sub-B Lineages, Alpha, Delta and Omicron variants, respectively. **B)** The phylogenetic networks tree of the S gene. The bootstrap values of the middle node were attached to the phylogenetic network tree.)

Figure 2 The Maximum-Likelilhood phylogenetic tree and phylogenetic network of the S gene in Vietnam

	A Sub-A Lineages	B Sub-B Lineages	Alpha	Delta	Omicron	TOTAL
A Sub-A Lineages	3.2	3.2	18.4	15.4	42.7	32.4
B Sub-B Lineages	1.4	0.5	15.8	13.4	40.2	30.1
Alpha	16.3	15.1	1.1	27.8	46.1	38.8
Delta	12.8	12.1	26.2	2.1	47.0	30.9
Omicron	32.8	31.6	37.1	37.6	16.8	27.9
TOTAL	16.3	15.3	23.7	15.3	4.9	29.1

Above diagonal : Average number of pairwise differences between populations

Diagonal elements : Average number of pairwise differences within population

**Below diagonal :** Corrected average pairwise difference

(In this picture, the nucleotide distance of each variant was displayed in three levels (three diagonal). The explanation in detail for the mean of the diagonal was displayed in the picture. Generally, each diagonal exhibited each colour, and the scale of colour shows the high distance nucleotide.)

Figure 3 The nucleotide pairwise differences in the S gene of each variant

3.2. Nucleotide diversity

Evaluating the diversity of the S gene at the nucleotide level showed that this gene contains an average of 34 mutations, in which the region encoded for the S1 subunit had the most mutation (Figure 4A). Furthermore, in the S1 region, we recorded NTD region had a higher mutation when compared to RBD. The explanation for the NTD region had the amount of mutation because all indels appeared in this region (Figure 5A). Examination of nucleotide changes in the S gene noted 1147 polymorphic sites, including 188 indels, 809 transitions and 394 transversions (Figure 4B). All types of transition and transversion were displayed. In detail, the transition mutation C to T is the highest frequency of the S gene, following T to C (Figure 4C). The Delta and Omicron variants had high polymorphic sites correlated with the high haplotype diversity of these variants in Vietnam. Determination of the number of mutations showed that the Omicron variant had the highest mutation, at 43 ± 7, followed by Alpha (at 18 ± 1) and Delta (at 15 ± 2) (Figure 4D). Despite Alpha having higher mutation with Delta, this variant had low nucleotide diversity (0.000276 ± 0.000205) (Figure 4D) and polymorphic sites (17 sites) when compared to Delta variants (0.000553 ± 0.000341 and 500 polymorphic sites). The main reason for this is that Alpha contained more deletion mutations in NTD than Delta. Generally, the diversity of nucleotide diversity level of the S gene in Vietnam is very high, with a nucleotide diversity index of 0.00753 ± 0.003641, of which Delta and the Omicron variant are the main contributors to this diversity (0.004360 ± 0.002145) (Figure 4D). This result is consistent with diversity at the haplotype level, suggesting that the Delta variant and Omicron had compressibility nucleotide changes in the S gene.
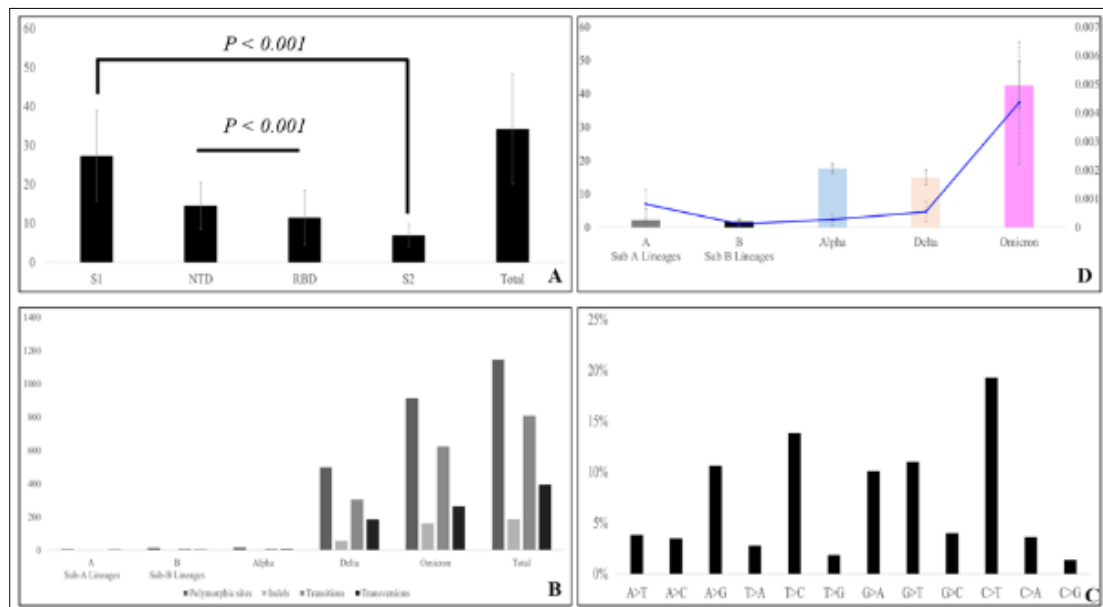
(A) The amount of mutation in the S gene, including the gene region encoded S1 subunit (NTD and RBD region) and S2 subunit. The *t*-test was used to identify the differences in mutation in each region, with a *p*-value < 0.05 displayed for the statistics significance. **B)** The examination of nucleotide changes in the S gene of each variant and Vietnam (total), including polymorphic site, indels (insertion and deletion mutation), transitions and transversions. **C)** The frequency of each type of transition and transversion was displayed in this picture. **D)** This picture displayed the amount of mutation (column) as the primary axis and nucleotide diversity index (blue line) as the secondary axis. All error bars in this figure show the SD: standard deviation)

Figure 4 The diversity at nucleotide level of the S gene in Vietnam

Examination of the mutation located in the S gene, we noted NTD conserved the deletion mutation (del21633_41ACCCCCTG – del24-26LPP, del22028_22033 – del157-58FR) and the nucleotide position at 22198 to 22206 as the hot region for insertion mutation when it appeared multiple time nucleotide inserting, including, ins22198GAGAGCCCAGAA (example sample: EPI_ISL_15897003), ins22198AATGGTGAG (example sample: EPI_ISL_16073578), ins22204GAGCCAGAA (many sequences, including one sample of Delta variant) and ins22206CGCAGTGGCAGT (example sample: EPI_ISL_16201111). We recorded a unique insert ins21620CCAGAGCGT (sample: EPI_ISL_13358783). Significantly, the RBD of S gene high conserved missense mutations (more than 10% frequency), including C22995A (T478K), G22578A (G339D), T22679C (S373P), G22813T (K417N), C22686T (S375F), A22688G (T376A), G22775A (D405N), A22786C (R408S), T22917G (L452R), A23063T (N501Y), T23075C (Y505H), A23055G (Q498R), G22992A (S477N), A23013C (E484A), T22882G (N440K), A23040G (Q493R), T23018G (F486V) and G22898A (G446S). In addition, the FCS region also noted some high-frequency mutations, such as A23403G (D614G), C23604A (P681H) and C23604G (P681R). Generally, the S gene had significant nucleotide changes and a high frequency of missense mutations, implying that this gene is a positive selection with dN/dS > 1 (Figure 5B).
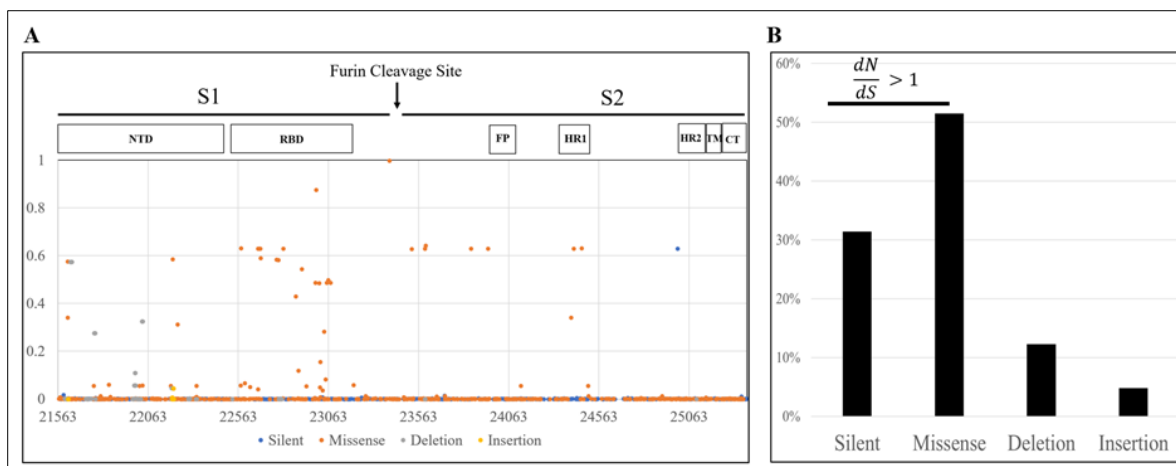
(A). This figure displayed the position of the mutation in the S gene. In which the S gene encoded two subunits (S1 subunit includes two major domains: NTD (N-terminal Domain) and RBD (Receptor Binding Domain); S2 subunit include FP (Fusion Peptide); HR1 and HR2 (Heptad repeat 1 and 2); TM (Transmembrane Domain); CT (Cytoplasmic Tail) and FCS (Furin Cleavage Site). The y-Axis is the frequency of mutation, and the x-Axis is the position of the S gene. The colour, including blue, orange, grey and yellow dots in the chart, displayed silent, missense, deletion and insertion mutation, respectively. **B).** The frequency of silent, missense, deletion and insertion mutation in the S gene and the rate of evolution (dN/dS) were discovered

Figure 5 The distribution and examination of mutation in the S gene

4. Discussion

In this study, we provide the whole landcapse genetic diversity of the S gene in Vietnam. In the metadata, we noticed a delay in the announcement and recognition of new variants in Vietnam. It is evident that in the dataset updated in July 2021, we did not notice the occurrence of the Beta variant [23]. However, in this study, we identified the occurrence of Beta due to the uncertainty in the data release. This is extremely dangerous because Beta can spread faster and more dangerous than the Wuhan variant [30, 31]. Thankfully, we have not seen any popularity of Beta, possibly due to Delta's rapid dominance since it first appeared in March 2021. In haplotype diversity, Vietnam noted the high genetic diversity of the S gene of SARS-CoV-2. Delta and Omicron variants are the ones with the most nucleotide changes. The rapid change of Delta and Omicron is also noted around the world. The Delta variant caused the worst outbreaks in India and the United States [6, 32]. This variant also quickly developed into many subvariant, such as Delta Plus in the world [33] or famous in Vietnam, AY.57. In Vietnam, the Omicron variant is more miscellaneous than Delta, particularly in the world general. We have recorded more than 130 Omicron subvariants that are more infectious than the ancestral variant (B.1.1.529), such as BA.2, BA.2.75 BA. 4, BA.5,... [19] In addition, from August 2022, Omicron becomes the foremost variant circulating in Vietnam. Given the high diversity of Omicron in Vietnam as well as around the world, it is imperative to monitor this variation continuously.

In nucleotide diversity, detailed analysis showed that the S1 region had the highest variability. This is consistent with the function of the S1 subunit, which serves as both an antigenic factor and a key player for direct interaction with the hACE2 receptor. Therefore, mutations that give SARS-CoV-2 the ability to evade the immune system and increase affinity for the receptor will be preferred. Deleting mutations has been shown to help mutants escape neutralizing antibodies [34, 35]. This also explains why Omicron is the variant with the most deletions and insertions recorded in our study. Notably, we identified a sample of the Delta variant (EPI_ISL_11221203) that had the insertion (ins2204GAGCCAGAA,) which has yet to be recorded in our previous study as in the world [36]. The insertion at nucleotide 22204 (ins22204GAGCCAGAA) at the S gene is most commonly of the Omicron variant in Vietnam and the world [37]. The Delta sequence (EPI_ISL_11221203) has a sampling period of February 2022, the parallel circulation of the Delta and Omicron variants in Vietnam. Therefore, we turn to the hypothesis that the most likely explanation is the co-infection and recombination of Delta and Omicron, which has been reported worldwide [38] . Based on our results, we highlight that the ongoing updating of the data and the assessment of nucleotide variations, including missense as well as indels mutation in the S gene, is very important.

5. Conclusion

In conclusion, our study shows that continuously updating nucleotide changes in the S gene region is imperative. We found that the S1 subunit coding region of the S gene has the fastest mutation rate. Furthermore, the S gene region of SARS-CoV-2 in Vietnam was selected for positivity during circulation here. Furthermore, in vaccine research, the S

glycoprotein plays an excellent target. Therefore, further studies to evaluate mutations affecting the structure and function of protein S are essential. To date, there are more than 15 million whole genome sequences of SARS-CoV-2 worldwide. Evaluating and providing information on mutations in the S gene will provide paramount data on the emergence of possible novel variants and sub-variants. It will be possible to evaluate the effectiveness of vaccine studies to be developed and to monitor changes related to the pathogenesis of the disease. Determining mutation rates and rates is also helpful because they play an important role in the virus escaping the host immune response and thus developing drug resistance.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare that there is no conflict of interest.

References

[1] Drake JW, Holland JJ. Mutation rates among RNA viruses. Proceedings of the National Academy of Sciences. 1999, 96(24):13910-3.

[2] Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell. 2020, 182(4):812-27.e19.

[3] Walker AS, Vihta K-D, Gethings O, Pritchard E, Jones J, House T, et al. Tracking the Emergence of SARS-CoV-2 Alpha Variant in the United Kingdom. New England Journal of Medicine. 2021, 385(27):2582-5.

[4] Duong D. Alpha, Beta, Delta, Gamma: What's important to know about SARS-CoV-2 variants of concern? CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2021, 193(27):E1059-e60.

[5] Earnest R, Uddin R, Matluk N, Renzette N, Turbett SE, Siddle KJ, et al. Comparative transmissibility of SARS-CoV-2 variants Delta and Alpha in New England, USA. Cell Reports Medicine. 2022, 3(4):100583.

[6] Yang W, Shaman J. COVID-19 pandemic dynamics in India, the SARS-CoV-2 Delta variant and implications for vaccination. J R Soc Interface. 2022, 19(191):20210900.

[7] Han P, Li L, Liu S, Wang Q, Zhang D, Xu Z, et al. Receptor binding and complex structures of human ACE2 to spike RBD from omicron and delta SARS-CoV-2. Cell. 2022, 185(4):630-40.e10.

[8] Di Giacomo S, Mercatelli D, Rakhimov A, Giorgi FM. Preliminary report on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Spike mutation T478K. Journal of medical virology. 2021, 93(9):5638-43.

[9] Cherian S, Potdar V, Jadhav S, Yadav P. SARS-CoV-2 Spike Mutations, L452R, T478K, E484Q and P681R, in the Second Wave of COVID-19 in Maharashtra, India. 2021, 9(7).

[10] Liu Y, Liu J, Johnson BA, Xia H, Ku Z, Schindewolf C, et al. Delta spike P681R mutation enhances SARS-CoV-2 fitness over Alpha variant. Cell Reports. 2022, 39(7):110829.

[11] Kumar S, Thambiraja TS, Karuppanan K, Subramaniam G. Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein. Journal of medical virology. 2022, 94(4):1641-9.

[12] Fan L-Q, Hu X-Y, Chen Y-Y, Peng X-L, Fu Y-H, Zheng Y-P, et al. Biological Significance of the Genomic Variation and Structural Dynamics of SARS-CoV-2 B.1.617. Frontiers in microbiology. 2021, 12:750725-.

[13] Candido KL, Eich CR, de Farina LO, Kadowaki MK, da Conceicao Silva JL, Maller A, et al. Spike protein of SARS-CoV-2 variants: a brief review and practical implications. Braz J Microbiol. 2022.

[14] Berrio A, Gartner V, Wray GA. Positive selection within the genomes of SARS-CoV-2 and other Coronaviruses independent of impact on protein function. PeerJ. 2020, 8:e10234.

[15] Chaw SM, Tai JH, Chen SL, Hsieh CH, Chang SY, Yeh SH, et al. The origin and underlying driving forces of the SARS-CoV-2 outbreak. J Biomed Sci. 2020, 27(1):73.

[16] Wise J. Covid-19: New coronavirus variant is identified in UK. BMJ. 2020, 371:m4857.

[17] Ghosh M, Basak S, Dutta S. Underlying selection for the diversity of spike protein sequences of SARS-CoV-2. IUBMB Life. 2022, 74(3):213-20.

[18] Barut GT, Halwe NJ, Taddeo A, Kelly JN, Schön J, Ebert N, et al. The spike gene is a major determinant for the SARS-CoV-2 Omicron-BA.1 phenotype. Nat Commun. 2022, 13(1):5929.

[19] Chavda VP, Balar P, Vaghela D, Solanki HK, Vaishnav A, Hala V, et al. Omicron Variant of SARS-CoV-2: An Indian Perspective of Vaccination and Management. Vaccines [Internet]. 2023, 11(1).

[20] Durmaz B, Abdulmajed O, Durmaz R. Mutations Observed in the SARS-CoV-2 Spike Glycoprotein and Their Effects in the Interaction of Virus with ACE-2 Receptor. Medeni Med J. 2020, 35(3):253-60.

[21] O'Toole A, Pybus OG, Abram ME, Kelly EJ, Rambaut A. Pango lineage designation and assignment using SARS-CoV-2 spike gene nucleotide sequences. BMC Genomics. 2022, 23(1):121.

[22] Thai PQ, Rabaa MA, Luong DH, Tan DQ, Quang TD, Quach HL, et al. The First 100 Days of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Control in Vietnam. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2021, 72(9):e334-e42.

[23] Quan Ke Thai, Phuoc Huynh, Duc Minh Le. Variations in spike gene of SARS-CoV-2 isolated in Vietnam. International Journal of Biosciences. 2022, 21:83-92.

[24] Quan Ke Thai, Phuoc Huynh, Yen Le Thi, Huyen Nguyen Thi Thuong. Genetic diversity of SARS-CoV-2 Omicron variants' spike gene in Vietnam. International Journal of Biosciences. 2022, 21(3):166-76.

[25] Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution. 2013, 30(4):772-80.

[26] Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Molecular Biology and Evolution. 2017, 34(12):3299-302.

[27] Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution. 2020, 37(5):1530-4.

[28] Leigh JW, Bryant D. popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution. 2015, 6(9):1110-6.

[29] Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010, 10(3):564-7.

[30] Planas D, Bruel T, Grzelak L, Guivel-Benhassine F, Staropoli I, Porrot F, et al. Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. Nature Medicine. 2021, 27(5):917-24.

[31] Ahmad A, Fawaz MAM, Aisha A. A comparative overview of SARS-CoV-2 and its variants of concern. Infez Med. 2022, 30(3):328-43.

[32] Dougherty K MM, Naqvi O, Matson D, Stone J. SARS-CoV-2 B.1.617.2 (Delta) Variant COVID-19 Outbreak Associated with a Gymnastics Facility — Oklahoma, April–May 2021. MMWR Morbidity and mortality weekly report. 2021, 2021(70):1004–7.

[33] Kannan SR, Spratt AN, Cohen AR, Naqvi SH, Chand HS, Quinn TP, et al. Evolutionary analysis of the Delta and Delta Plus variants of the SARS-CoV-2 viruses. J Autoimmun. 2021, 124:102715.

[34] Meng B, Kemp SA, Papa G, Datir R, Ferreira I, Marelli S, et al. Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. Cell Rep. 2021, 35(13):109292.

[35] Mishra T, Dalavi R, Joshi G, Kumar A, Pandey P, Shukla S, et al. SARS-CoV-2 spike E156G/Δ157-158 mutations contribute to increased infectivity and immune escape. Life Sci Alliance. 2022, 5(7).

[36] Venkatakrishnan AJ, Anand P, Lenehan PJ, Suratekar R, Raghunathan B, Niesen MJM, et al. On the Origins of Omicron's Unique Spike Gene Insertion. Vaccines (Basel). 2022, 10(9).

[37] Ou J, Lan W, Wu X, Zhao T, Duan B, Yang P, et al. Tracking SARS-CoV-2 Omicron diverse spike gene mutations identifies multiple inter-variant recombination events. Signal transduction and targeted therapy. 2022, 7(1):138.

[38] Combes P, Bisseux M, Bal A, Marin P, Latour J, Archimbaud C, et al. Evidence of co-infections during Delta and Omicron SARS-CoV-2 variants co-circulation through prospective screening and sequencing. Clin Microbiol Infect. 2022, 28(11):1503 e5- e8.