

## PHÂN TÍCH CÁC ĐỘT BIẾN TRÊN PROTEIN GAI CỦA CÁC DÒNG BIẾN THỂ SARS-COV-2

ANALYSIS OF SPIKE PROTEIN MUTATIONS OF SARS-COV-2 VARIANT LINES

TRƯƠNG THẾ QUANG<sup>(\*)</sup>

**TÓM TẮT:** Ba mươi ba trình tự protein gai (S) của các biến thể SARS-CoV-2 và một trình tự S tham chiếu (QIC53213) được thu thập từ cơ sở dữ liệu GenBank, Hoa Kỳ. Kết quả sắp hàng hai trình tự S của từng biến thể SARS-CoV-2 với trình tự S tham chiếu nhận diện được bảy dòng biến thể SARS-CoV-2 Alpha, Beta, Gamma, Delta, Delta Plus, Iota, Kappa và chín dòng biến thể N1-N9 chưa có mã WHO. Các đột biến thường xảy ra trên S của các dòng biến thể SARS-CoV-2 với xác suất xuất hiện là D614G 39,29%, N501Y 14,29%, L452R 1,90%, P681R 10,71%, P681H 9,52%, E484K 8,33% và K417N 5,95%.

**Từ khóa:** các dòng biến thể SARS-CoV-2; đột biến; protein gai; sắp hàng hai trình tự.

**ABSTRACT:** Thirty-three spike protein (S) sequences of the SARS-CoV-2 variants and a reference S sequence (QIC53213) were collected from the GenBank database, USA. We identified seven variant lines Alpha, Beta, Gamma, Delta, Delta Plus, Iota, Kappa and nine variant lines N1-N9 without WHO code using pairwise S sequence alignment of each SARS-CoV-2 variant and the reference S sequence. The most common mutations on the S of the SARS-CoV-2 variant lines with the occurrence probability are D614G 39.29%, N501Y 14.29%, L452R 1.90%, P681R 10.71%, P681H 9.52%, E484K 8.33% and K417N 5.95%.

**Key words:** SARS-CoV-2 variant lines; mutation; spike protein; pairwise sequence alignment.

### 1. ĐẶT VẤN ĐỀ

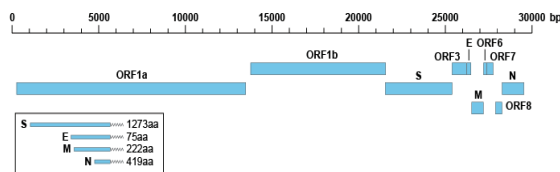
Coronavirus hội chứng hô hấp cấp tính nghiêm trọng 2 (Severe Acute Respiratory Syndrome Coronavirus 2) viết tắt SARS-CoV-2 là nhóm coronavirus mới gây bệnh COVID-19 bắt đầu đột bùng phát vào tháng 12-2019 tại Vũ Hán, Trung Quốc với quy mô và tốc độ lây nhiễm ngày càng cao trên toàn thế giới. Theo cập nhật của Worldometers về đại dịch COVID-19 trên toàn thế giới tính đến 00:22 GMT ngày 27-10-2021 đã có 223 quốc gia với 245.253.523 người bị nhiễm bệnh và số người chết lên đến 4.978.144 [13].

SARS-CoV-2 là virus có bộ gene RNA sợi đơn, dương tính (+ssRNA) dài khoảng 30.000 nt mã hóa cho 16 protein không cấu trúc (Nonstructural

Protein) NSP1-16 thuộc ORF1a và ORF1ab; 4 protein cấu trúc S, E, M, N và 6 protein phụ ORF3a, ORF6, ORF7a, ORF7b, ORF8, ORF10 [2, tr.914-921] (hình 1). Bốn protein cấu trúc chính đã được xác định trong SARS-CoV-2 bao gồm các protein gai (S), vỏ (E), màng (M) và nucleocapsid (N), chúng có chung nhận dạng đáng kể với SARS-CoV. Bốn protein cấu trúc này được mã hóa bởi ORF2-10 và cần thiết cho quá trình hình thành áo virus và đóng gói bộ gene. So với NSPs, protein cấu trúc tạo ra phản ứng miễn dịch cao hơn, cả qua trung gian dịch thể và tế bào [9]. Các protein gai (Spike Protein) của SARS-CoV-2 (1,273 aa) có hai vùng chức năng S1 (14-667 aa) và S2 (668-1255 aa), trong đó thụ thể tế bào mediate ràng buộc và phản ứng

<sup>(\*)</sup> TS. Trường Đại học Văn Lang, quangtruongthe@gmail.com, Mã số: TCKH31-02-2022

tổng hợp màng, tương ứng. Sự gắn kết và xâm nhập của virus vào tế bào chủ phụ thuộc vào sự tương tác của vùng liên kết thụ thể của chồi (RBD: 306-527 aa) với các protein tế bào cụ thể. Enzyme chuyển hóa angiotensin 2 (ACE-2) và protease TMPRSS2 đã được xác định là thụ thể tế bào và protease môi cần thiết cho sự xâm nhập tế bào của SARS-CoV-2 tương ứng. Protease TMPRSS2 hoặc cathepsin L phân cắt protein gai ở điểm nối S1-S2, cho phép hợp nhất giữa vỏ virus và màng tế bào cần thiết cho sự xâm nhập của virus. SARS-CoV-2 cũng có thể bị phân cắt bởi furin, chất nhận biết trình tự bốn amino acid tại điểm nối S1-S2. Bên cạnh sự phân cắt của protein gai tại điểm nối S1-S2, các protease tế bào còn phân cắt tiểu đơn vị S2 tại một trình tự S2' được công nhận, điều này rất quan trọng để kích hoạt protein gai và kết quả là hợp nhất màng [9].



**Hình 1.** Cấu trúc bộ gene SARS-CoV-2 trong đó có gene S mã hóa cho spike protein [9]

SARS-CoV-2 liên tục biến đổi tạo ra nhiều biến thể do những đột biến (đột biến tự nhiên) trong quá trình sao chép bộ gene. Dòng là một nhóm các biến thể virus có liên quan chặt chẽ về mặt di truyền, có nguồn gốc từ cùng một cấu trúc sơ khai chung. Một biến thể SARS-CoV-2 có một hoặc nhiều đột biến tách biệt với các biến thể khác của virus. Các dòng biến thể di truyền của SARS-CoV-2 đã xuất hiện và lây truyền trên khắp thế giới kể từ khi bắt đầu đại dịch COVID-19. Tổ chức liên ngành SARS-CoV-2 Hoa Kỳ, viết tắt là SIG (The United States Government SARS-CoV-2 Interagency Group) đã phân loại các biến thể SARS-CoV-2 thành bốn loại tăng theo mức độ nghiêm trọng như sau:

1) Biến thể đang được theo dõi VBM (*Variant Being Monitored*) gồm các biến thể có dữ liệu cho thấy trong quá khứ có tác động tiềm ẩn hoặc rõ ràng đối với các biện pháp đối phó của hệ thống y tế hoặc có liên quan đến bệnh nặng hơn hoặc việc lây truyền gia tăng nhưng hiện tại không còn được phát hiện hoặc đang tồn tại trong cộng đồng ở mức độ rất thấp. Do đó, không gây ra nguy cơ đáng kể và áp lực đối với hệ thống y tế trong tương lai. Theo hệ thống phân loại SIG ngày 21-9-2021, VBM gồm các biến thể Alpha (B.1.1.7, Q.1-Q.8), Beta (B.1.351, B.1.351.2, B.1.351.3), Gamma (P.1, P.1.1, P.1.2), Epsilon (B.1.427, B.1.429), Zeta (P.2), Eta (B.1.525), Iota (B.1.526), Kappa (B.1.617.1), B.1.617.3, Mu (B.1.621, B.1.621.1) [5].

2) Biến thể đáng quan tâm VOI (*Variant of Interest*) gồm các biến thể có những thay đổi với liên kết thụ thể, giảm trung hòa bởi các kháng thể được tạo ra để chống lại tình trạng lây nhiễm hoặc tiêm chủng trước đó, giảm hiệu quả của các biện pháp điều trị, tác động chẩn đoán tiềm tàng, hoặc gia tăng dự đoán về khả năng lây truyền hoặc độ nghiêm trọng của bệnh. Theo hệ thống phân loại SIG ngày 21-9-2021, VOI không có biến thể nào [5].

3) Biến thể đáng lo ngại VOC (*Variant of Concern*) gồm các biến thể có bằng chứng về sự gia tăng khả năng lây truyền, gây bệnh nghiêm trọng hơn do tăng số lần nhập viện hoặc tử vong, giảm đáng kể khả năng trung hòa bởi các kháng thể được tạo ra trong quá trình lây nhiễm hoặc tiêm chủng trước đó, giảm hiệu quả của phương pháp điều trị, giảm hiệu quả của vaccine hoặc không phát hiện được khi chẩn đoán. Theo hệ thống phân loại SIG ngày 21-9-2021, VOC gồm các biến thể Delta (B.1.617.2) và Delta Plus (B.1.617.2.1 hoặc AY.1) [5].

4) Biến thể có hậu quả nghiêm trọng VOHC (*Variant of High Consequence*) gồm các biến thể có bằng chứng rõ ràng về hiệu quả của các biện pháp phòng ngừa hoặc biện pháp ứng phó của hệ thống y tế đã giảm đáng kể so với

các biến thể lưu hành trước đó. Theo hệ thống phân loại SIG ngày 21-9-2021, VOHC không có biến thể nào [5].

Nghiên cứu phân tích các đột biến trên protein gai của các biến thể SARS-CoV-2 giúp nhận biết sự tiến hóa của các dòng biến thể SARS-CoV-2, phân loại các biến thể theo hệ thống SIG, đánh giá mức độ nghiêm trọng về khả năng lây nhiễm, độc lực, khả năng kháng vaccine của virus.

## 2. NỘI DUNG

### 2.1. Phương pháp nghiên cứu

#### 2.1.1. Thu thập các trình tự protein gai của các biến thể SARS-CoV-2

Các trình tự protein gai (trình tự S) của các biến thể SARS-CoV-2 ở người (SARS-CoV-2) và trình tự S tham chiếu (QIC53213) được thu thập từ cơ sở dữ liệu GenBank thuộc Trung tâm Thông tin Công nghệ sinh học Quốc gia NCBI, Mỹ (*The National Center for Biotechnology Information, USA*).

#### 2.1.2. Nhận diện các vị trí đột biến trên trình tự S

Mỗi trình tự S của từng biến thể SARS-CoV-2 được sắp hàng hai trình tự [1, tr.93-95] với trình tự S tham chiếu, ứng dụng công cụ BLASTp (Protein Basic Local Alignment Search Tool) [1, tr.137-141] trên cơ sở dữ liệu NCBI. Kết quả sắp hàng hai trình tự được dùng để nhận diện các vị trí đột biến thay thế amino acid (aa) hoặc mất aa, những thay đổi về cấu trúc và chức năng của S tương ứng [3, tr.9-31]. Cần lưu ý rằng đột biến thay thế một aa A1 bằng aa A2 tại vị trí p trên trình tự S của biến thể SARS-CoV-2 được ký hiệu là A1pA2, đột biến mất aa A tại vị trí q trên trình tự S của biến thể SARS-CoV-2 được ký hiệu là Aq- hoặc Aqdel.

#### 2.1.3. Tỷ lệ tương đồng

Tỷ lệ tương đồng (Identity)  $I_{xr}$  (%) của biến thể x so với biến thể tham chiếu r dựa trên trình tự S được ước lượng theo công thức (1).

$$I_{xr}(\%) = \frac{100 \cdot (L_r - M_{xr})}{L_r} \quad (1)$$

Trong đó:  $I_{xr}$  (%) là tỷ lệ tương đồng trình tự S của biến thể x so với trình tự S tham chiếu r,  $M_{xr}$  là tổng số đột biến trên trình tự S của biến thể x khi sắp hàng hai trình tự với trình tự S tham chiếu r,  $L_r = 1273$  (aa) là chiều dài của trình tự S tham chiếu r.

#### 2.1.4. Xác suất xuất hiện đột biến

Xác suất xuất hiện đột biến k trên trình tự S được ước lượng theo công thức (2).

$$P_k(\%) = \frac{100 \cdot M_k}{M} \quad (2)$$

Trong đó:  $P_k$  (%) là xác suất xuất hiện đột biến k trên trình tự S,  $M_k$  là số đột biến k xuất hiện trên trình tự S, M là tổng số đột biến thường xuất hiện trên tất cả các trình tự S của các biến thể SARS-CoV-2 nghiên cứu.

#### 2.1.5. Phân tích thống kê

Kết quả ước lượng số lượng đột biến, tỷ lệ tương đồng và xác suất xuất hiện đột biến được trình bày dưới dạng giá trị trung bình. Xử lý số liệu, so sánh các giá trị trung bình bằng phương pháp phân tích phương sai ANOVA (Analysis of Variance) với xác suất tin cậy 95%.

## 2.2. Kết quả và thảo luận

### 2.2.1. Danh mục các trình tự S

Các trình tự S của biến thể SARS-CoV-2 và trình tự S tham chiếu được thu thập từ cơ sở dữ liệu GenBank (NCBI, USA) gồm 33 trình tự nghiên cứu và một trình tự tham chiếu (bảng 1).

### 2.2.2. Các dòng biến thể SARS-CoV-2 dựa trên trình tự S

Nhận diện các đột biến trên các trình tự S của các biến thể SARS-CoV-2 bằng cách giải bài toán sắp hàng hai trình tự S của từng biến thể SARS-CoV-2 với trình tự S tham chiếu, ứng dụng công cụ BLASTp (NCBI, USA). Kết quả sắp hàng hai trình tự ước lượng được tỷ lệ tương đồng, nhận diện được các đột biến trên trình tự S trong quá trình tiến hóa và các dòng biến thể SARS-CoV-2 dựa trên trình tự S (bảng 2).

**Bảng 1.** Danh mục các trình tự S của các biến thể SARS-CoV-2

STT	Accession	Nơi lấy mẫu	Ngày lấy mẫu	STT	Accession	Nơi lấy mẫu	Ngày lấy mẫu
1	QIC53213 <sup>(*)</sup>	China	08-01-20	18	QXO92800	Indonesia	12-01-21
2	QVL62123	Armenia	26-8-20	19	BCX28609	Japan	13-02-21
3	QXP82487	Armenia	16-01-21	20	BDA76997	Japan	27-6-21
4	QXP82523	Armenia	27-01-21	21	QXI88922	Malaysia	01-4-21
5	QVQ64541	Hong Kong	21-3-21	22	QZD74968	Malaysia	01-7-21
6	QVQ64733	Hong Kong	10-4-21	23	QSV99839	Philippines	25-01-21
7	QVQ64709	Hong Kong	11-4-21	24	QUN71013	Philippines	22-3-21
8	QVX27130	Hong Kong	05-5-21	25	QUN71084	Philippines	08-4-21
9	QVX27226	Hong Kong	12-5-21	26	QUL62216	USA	16-2-21
10	QVP61498	India	12-8-20	27	QSL84998	USA	20-2-21
11	QVL01151	India	22-8-20	28	QTZ79480	USA	15-3-21
12	QXN66784	India	08-2-21	29	QUI88445	USA	15-4-21
13	QYM54880	India	08-3-21	30	QXY72640	USA	15-5-21
14	QYJ09531	India	08-4-21	31	QXV98254	USA	15-6-21
15	QWT95288	India	08-5-21	32	QYN99533	USA	15-7-21
16	QXN18161	India	08-6-21	33	QZW32714	USA	08-8-21
17	UAC16107	India	10-8-21	34	UAY16900 [11, 12]	Vietnam	15-7-21

Chú thích: (\*) Trình tự tham chiếu

**Bảng 2.** Các dòng biến thể SARS-CoV-2 dựa trên trình tự S

STT	Accession	Mã WHO	Dòng biến thể	Phân loại SIG	Tỷ lệ tương đồng (%)	Các đột biến trên trình tự S (aa-vị trí-aa)	Số đột biến
1	QIC53213	n/r	n/r	n/r	100.00	Trình tự tham chiếu	0
2	QVQ64541	Alpha	B.1.1.7	VBM	99.21	H69-, V70-, Y145-, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H	10
3	QYM54880	Alpha	B.1.1.7	VBM	99.21	H69-, V70-, Y145-, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H	10
4	BCX28609	Alpha	B.1.1.7	VBM	99.21	H69-, V70-, Y145-, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H	10
5	BDA76997	Alpha	B.1.1.7	VBM	99.21	H69-, V70-, Y145-, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H	10
6	QSV99839	Alpha	B.1.1.7	VBM	99.21	H69-, V70-, Y145-, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H	10
7	QUL62216	Alpha	B.1.1.7	VBM	99.21	H69-, V70-, Y145-, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H	10
8	QVX27130	Beta	B.1.351	VBM	99.06	A27S, D80A, D215G, L242-, A243-, L244-, K417N, E484K, N501Y, D614G, A701V, G1124C	12
9	QVX27226	Beta	B.1.351	VBM	99.06	A27S, D80A, D215G, L242-, A243-, L244-, K417N, E484K, N501Y, D614G, A701V, G1124C	12
10	QUN71013	Beta	B.1.351	VBM	99.14	A27S, D80A, D215G, L242-, A243-, L244-, K417N, E484K, N501Y, D614G, A701V	11
11	QUN71084	Beta	B.1.351	VBM	99.14	A27S, D80A, D215G, L242-, A243-, L244-, K417N, E484K, N501Y, D614G, A701V	11
12	QYJ09531	Delta Plus	AY.1	VOC	99.06	T19R, T95I, G142D, E156G, F157-, R158-, K417N, L452R, T478K, D614G, P681R, D950N	12

STT	Accession	Mã WHO	Dòng biến thể	Phân loại SIG	Tỷ lệ tương đồng (%)	Các đột biến trên trình tự S (aa-vị trí-aa)	Số đột biến
13	QWT95288	Delta	B.1.617.2	VOC	99.14	T19R, G142D, E156G, F157-, R158-, A222V, L452R, T478K, D614G, P681R, D950N	11
14	QXN18161	Delta	B.1.617.2	VOC	99.45	T19X, G142D, L452R, T478R, D614G, P681R, D950N	7
15	UAC16107	Delta	B.1.617.2	VOC	99.53	T19R, L452R, T478K, D614G, P681R, D950N	6
16	QZD74968	Delta	B.1.617.2	VOC	99.53	T19X, L452R, T478K, D614G, P681R, D950N	6
17	QZW32714	Delta	B.1.617.2	VOC	99.21	T19R, T95I, E156G, F157-, R158-, L452R, T478K, D614G, P681R, D950N	10
18	UAY16900	Delta	B.1.617.2	VOC	99.45	T19R, G142D, L452R, T478R, D614G, P681R, D950N	7
19	QXV98254	Gamma	P.1	VBM	99.06	L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I, V1176F	12
20	QYN99533	Gamma	P.1	VBM	99.06	L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I, V1176F	12
21	QXY72640	Iota	B.1.526	VBM	99.53	L5F, T95I, D253G, E484K, D614G, A701V	6
22	QXN66784	Kappa	B.1.617.1	VBM	99.45	V16I, T95I, L452R, E484Q, D614G, P681R, Q1071H	7
23	QVL62123	N1	n/r	n/r	99.84	N460I, D614G	2
24	QXP82487	N2	n/r	n/r	99.84	D614G, P681H	2
25	QXP82523	N2	n/r	n/r	99.84	D614G, P681H	2
26	QVQ64733	N3	n/r	n/r	99.84	D614G, A1087S	2
27	QVQ64709	N3	n/r	n/r	99.84	D614G, A1087S	2
28	QVP61498	N4	n/r	n/r	99.84	Y508H, D614G	2
29	QVL01151	N5	n/r	n/r	99.45	T76I, T95I, E484D, D614G, P812R, T1066N, P1143L	7
30	QXO92800	N6	n/r	n/r	99.92	D614G	1
31	QXI88922	N7	n/r	n/r	99.76	D614G, A701V, G1223C	3
32	QSL84998	N8	n/r	n/r	99.69	S13I, W152C, L452R, D614G	4
33	QUI88445	N8	n/r	n/r	99.69	S13I, W152C, L452R, D614G	4
34	QTZ79480	N9	n/r	n/r	99.69	S494P, D614G, P681H, T716I	4

Chú thích: Ni-9 tên các dòng biến thể chưa có mã WHO, n/r chưa có tên dòng hoặc chưa được phân loại theo hệ thống SIG

1) *Biến thể Alpha (B.1.1.7)*: có tỷ lệ tương đồng 99,21% so với chủng SARS-CoV-2 gốc (QIC53213) dựa trên trình tự S. Phân loại theo hệ thống phân loại SIG ngày 21-9-2021, biến thể này thuộc loại VBM. Các trình tự S của biến thể Alpha nghiên cứu gồm QVQ64541, QYM54880, BCX28609, BDA76997, QSV99839, QUL62216 có 10 đột biến trên mỗi trình tự, các đột biến thường xảy ra là D614G, N501Y, P681H. Biến thể Alpha được phát hiện lần đầu ở Vương Quốc Anh vào tháng 12-2020. Cho đến nay, biến thể này đã xuất hiện tại hơn 100 quốc gia, trong đó có Việt Nam. Chỉ số tái sinh cơ bản  $R_0$  là số ca lây nhiễm trung bình mà một trường hợp có thể lây sang trong suốt thời kỳ lây nhiễm trong dân số chưa có miễn dịch. Đối với chủng SARS-CoV-2 gốc,  $R_0$  chỉ khoảng 2,3 đến 2,5, tức là một người có thể lây bệnh cho

2,3 đến 2,5 người khác. Đối với biến thể Alpha chỉ số  $R_0$  trong khoảng từ 3,4 đến 4,5 [8].

2) *Biến thể Beta (B.1.351)*: có tỷ lệ tương đồng 99,06 đến 99,14% so với chủng SARS-CoV-2 gốc dựa trên trình tự S. Biến thể này thuộc loại VBM theo hệ thống phân loại SIG ngày 21-9-2021. Các trình tự S của biến thể Beta nghiên cứu gồm QVX27130, QVX27226, QUN71013, QUN71084 có từ 11 đến 12 đột biến trên mỗi trình tự, các đột biến thường xảy ra là D614G, N501Y, E484K. Biến thể Beta lần đầu tiên được xác định tại Nam Phi vào tháng 9-2020, tuy nhiên trước đó, biến thể đã có mặt tại ít nhất 20 quốc gia khác, bao gồm Vương Quốc Anh. Biến thể Beta có khả năng lây nhiễm gấp 1,5 lần so với biến thể Alpha, đồng thời có khả năng tiến hóa và thích nghi cao hơn chủng virus cũ [8].

3) *Biến thể Gamma (P.1)*: có tỷ lệ tương đồng 99,06% so với chủng SARS-CoV-2 gốc dựa trên trình tự S. Biến thể này được xếp vào loại VBM theo hệ thống phân loại SIG ngày 21-9-2021. Các trình tự S của biến thể Gamma nghiên cứu gồm QXV98254, QYN99533 có 12 đột biến trên mỗi trình tự, trong đó các đột biến thường xảy ra là D614G, N501Y, E484K. Biến thể Gamma lần đầu tiên được xác định tại Brazil vào tháng 12-2020, nhưng trước đó cũng đã lan rộng tại 10 quốc gia khác, bao gồm Vương Quốc Anh. Biến thể Gamma có khả năng lây nhiễm cao hơn gấp 2,5 lần so với chủng SARS-CoV-2 gốc và có khả năng tái lây nhiễm cho những người trước đó đã khỏi bệnh [8].

4) *Biến thể Delta (B.1.617.2)*: có tỷ lệ tương đồng từ 99,14-99,53% so với chủng SARS-CoV-2 gốc dựa trên trình tự S. Theo hệ thống phân loại SIG ngày 21-9-2021, biến thể này được xếp vào loại VOC. Các trình tự S của biến thể Delta nghiên cứu gồm QYJ09531, QWT95288, QXN18161, UAC16107, QZD74968, UAY16900 có từ 6 đến 11 đột biến trên mỗi trình tự, trong đó các đột biến thường xảy ra là D614G, L452R, P681R. Biến thể Delta xuất hiện lần đầu tiên tại Ấn Độ vào tháng 12-2020, biến thể này có khả năng lây nhiễm nhiều hơn và nhanh hơn so với các biến thể trước đó của virus gây bệnh COVID-19, nó có thể gây ra tình trạng bệnh nghiêm trọng hơn so với các biến thể SARS-CoV-2 trước đây ở người chưa được tiêm vaccine. Biến thể này có khả năng lây lan với tốc độ kinh hoàng, khó truy vết, nhanh chóng trở thành cơn ác mộng của Ấn Độ, Vương Quốc Anh, các nước thuộc khu vực Đông Nam Á, trong đó có Việt Nam và nhiều quốc gia khác trên thế giới, làm áp đảo hệ thống y tế toàn cầu. Đối với biến thể Delta mức độ lây lan nhanh hơn rất nhiều với chỉ số  $R_0$  khoảng từ 6,7 đến 7,6 [8].

5) *Biến thể Delta Plus (B.1.617.2.1 hoặc AY.1)*: có tỷ lệ tương đồng 99,06% so với chủng SARS-CoV-2 gốc dựa trên trình tự S. Phân loại

theo hệ thống SIG ngày 21-9-2021, biến thể này thuộc loại VOC. Trình tự S của biến thể Delta Plus nghiên cứu là QYJ09531 có 12 đột biến, trong đó các đột biến thường xảy ra là D614G, L452R, P681R. Biến thể Delta Plus xuất hiện đầu tiên tại Vương Quốc Anh vào tháng 6-2021 đã lan rộng đến hơn 130 quốc gia và vùng lãnh thổ. Về mặt di truyền, Delta và Delta Plus cơ bản là giống nhau, nhưng Delta Plus có đột biến bổ sung K417N ở protein gai giúp virus có thể tiếp cận với tế bào của người [8]. Delta Plus được coi là có khả năng lây nhiễm cao và lây lan nhanh chóng. Nó cũng cho thấy khả năng bám chặt hơn vào các thụ thể của tế bào phổi của người và có khả năng làm giảm phản ứng kháng thể đơn dòng. Đối với biến thể Delta Plus mức độ lây lan nhanh hơn rất nhiều, lấn át các biến thể khác với chỉ số  $R_0$  khoảng từ 9 đến 10.

6) *Biến thể Iota (B.1.526)*: có tỷ lệ tương đồng 99,53% so với chủng SARS-CoV-2 gốc dựa trên trình tự S. Biến thể này thuộc loại VBM theo hệ thống phân loại SIG ngày 21-9-2021. Trình tự S của biến thể Iota nghiên cứu là QXY72640 có 6 đột biến, trong đó các đột biến thường xảy ra là D614G, E484K. Biến thể Iota được tìm thấy đầu tiên tại Hoa Kỳ vào tháng 11-2020, biến thể này làm giảm độ nhạy đối với biện pháp điều trị bằng kháng thể đơn dòng, tuy nhiên ý nghĩa lâm sàng của việc suy giảm này chưa được biết rõ. Khả năng trung hòa bằng huyết thanh giảm trong thời kỳ dưỡng bệnh và sau tiêm chủng [6].

7) *Biến thể Kappa (B.1.617.1)*: có tỷ lệ tương đồng 99,45% so với chủng SARS-CoV-2 gốc dựa trên trình tự S. Phân loại theo hệ thống phân loại SIG ngày 21-9-2021, biến thể này thuộc loại VBM. Trình tự S của biến thể Kappa nghiên cứu là QXN66784 có 7 đột biến, trong đó các đột biến thường xảy ra là D614G, L452R, P681R. Biến thể Kappa xuất hiện lần đầu tiên tại Ấn Độ vào tháng 10-2020, biến thể này có khả năng giảm trung hòa đối với một số biện

pháp điều trị bằng kháng thể đơn dòng hoặc huyết thanh trong thời kỳ sau tiêm chủng [6].

8) Các dòng biến thể SARS-CoV-2 chưa có mã WHO: so với chủng SARS-CoV-2 gốc (QIC53213) dựa trên trình tự S, các dòng biến thể N1 (QVL62123), N2 (QXP82487, QXP82523), N3 (QVQ64733, QVQ64709) và N4 (QVP61498) đều có tỷ lệ tương đồng 99,84%, tỷ lệ tương đồng của biến thể N5 (QVL01151) là 99,45%, biến thể N6 (QXO92800) là 99,92%, biến thể N7 (QXI88922) là 99,76% và các dòng biến thể N8 (QSL84998, QUI88445), N9 (QTZ79480) đều có tỷ lệ tương đồng 99,69%. Trình tự S của các biến thể N1-N9 có đột biến thường gặp là D614G. Trình tự S của biến thể N8 có đột biến thường gặp là L452R. Trình tự S của biến thể N2 và N9 có đột biến thường gặp là P681H.

### 2.2.3. Các đột biến thường xuất hiện trên S của các biến thể SARS-CoV-2

Các đột biến điểm thay thế hoặc mất aa được tìm thấy trên các trình tự S của các biến thể SARS-CoV-2 nghiên cứu có chiều dài 1273 aa bắt đầu từ vị trí aa 5 đến 1223. Số lượng đột biến aa đã được phát hiện trên mỗi trình tự S từ 1 đến 12 (bảng 2).

Đột biến kỵ nước D614G tại vị trí 614 aspartic acid (D) ưa nước được thay thế bằng glycine (G) kỵ nước có xác suất xuất hiện cao nhất trong nghiên cứu 39,29%. Đột biến này có trên S của các biến thể Alpha, Beta, Gamma, Delta, Delta Plus, Iota, Kappa và các biến thể N1-N9 chưa có nhãn WHO. Đột biến D614G làm virus có khả năng lây nhiễm cao hơn chủng SARS-CoV-2 gốc và không kháng vaccine [7].

**Bảng 3.** Các đột biến đặc trưng thường xuất hiện trên S của các dòng biến thể SARS-CoV-2 nghiên cứu

Tên đột biến	Loại đột biến	Đột biến xuất hiện trên các dòng biến thể SARS-CoV-2 (mã WHO)	Đột biến làm thay đổi hiệu quả hoạt động của SARS-CoV-2	Xác suất xuất hiện (%)
D614G	Kỵ nước	Alpha, Beta, Gamma, Delta, Delta Plus, Iota, Kappa, N1-N9	Khả năng lây nhiễm cao, không kháng vaccine	39,29
N501Y	Trung tính ưa nước	Alpha, Beta, Gamma	Khả năng lây nhiễm tăng, ít kháng vaccine	14,29
L452R	Ưa nước	Delta, Delta Plus, Kappa, N8	Khả năng lây nhiễm cao, ít kháng vaccine	11,90
P681R	Ưa nước	Delta, Delta Plus, Kappa	Khả năng lây nhiễm cao, gây bệnh nặng hơn, ít kháng vaccine	10,71
P681H	Ưa nước	Alpha, N2, N9	Khả năng lây nhiễm tăng, gây bệnh nặng hơn, không kháng vaccine	9,52
E484K	Trung tính ưa nước	Beta, Gamma, Iota	Khả năng lây nhiễm tăng, ít kháng vaccine	8,33
K417N	Trung tính ưa nước	Delta Plus	Khả năng lây nhiễm rất cao, ít kháng vaccine	5,95

Đột biến trung tính ưa nước N501Y tại vị trí 501 asparagine (N) ưa nước được thay thế bằng tyrosine (Y) ưa nước có xác suất xuất hiện 14,29%. Đột biến này có trên S của các biến thể Alpha, Beta, Gamma. Biến thể mang đột biến này làm tăng khả năng lây nhiễm và ít kháng vaccine [7].

Đột biến ưa nước L452R tại vị trí 452 leucine (L) kỵ nước được thay thế bằng arginine (R) ưa nước có xác suất xuất hiện 11,90%. Đột biến này có trên S của các biến thể Delta, Delta Plus, Kappa và biến thể N8 chưa có mã WHO.

Biến thể mang đột biến này có khả năng lây nhiễm cao và ít kháng vaccine [7].

Đột biến ưa nước P681R tại vị trí 681 proline (P) kỵ nước được thay thế bằng arginine (R) ưa nước có xác suất xuất hiện 10,71%. Đột biến này có trên S của các biến thể Delta, Delta Plus, Kappa. Biến thể mang đột biến này có khả năng lây nhiễm cao, gây bệnh nặng hơn và ít kháng vaccine [10].

Đột biến ưa nước P681H tại vị trí 681 proline (P) kỵ nước được thay thế bằng histidine (H) ưa nước có xác suất xuất hiện 9,52%. Đột

biến này có trên S của các biến thể Alpha và các biến thể N2, N9 chưa có mã WHO. Biến thể mang đột biến P681H có khả năng lây nhiễm tăng, gây bệnh nặng hơn và không kháng vaccine [7].

Đột biến trung tính uracil E484K tại vị trí 484 glutamic acid (E) uracil được thay thế bằng lysine (K) uracil có xác suất xuất hiện 8,33%. Đột biến này có trên S của các biến thể Beta, Gamma, Iota. Đột biến E484K tạo ra biến thể có khả năng lây nhiễm tăng so với SARS-CoV-2 gốc, ít kháng vaccine [7].

Đột biến trung tính uracil K417N tại vị trí 417 lysine (K) uracil được thay thế bằng histidine (N) uracil có xác suất xuất hiện 5,95%. Đột biến này có trên S của biến thể Delta Plus. Biến thể Delta Plus có một đột biến bổ sung K417N, giúp phân biệt nó với biến thể Delta thông thường. Biến thể mang đột biến này có khả năng lây nhiễm rất cao, ít kháng vaccine [4].

### 3. KẾT LUẬN

Các trình tự protein gai (S) của các biến thể SARS-CoV-2 và một trình tự S tham chiếu QIC53213 được thu thập từ cơ sở dữ liệu GenBank, Hoa Kỳ. Kết quả sắp hàng hai trình tự S của từng biến thể SARS-CoV-2 với trình tự S tham chiếu nhận diện được bảy dòng biến thể SARS-CoV-2 Alpha, Beta, Gamma, Delta, Delta Plus, Iota, Kappa và chín dòng biến thể N1-N9 chưa có mã WHO. Các đột biến đặc trưng thường xuất hiện trên S của các dòng

biến thể SARS-CoV-2 là D614G, N501Y, L452R, P681R, P681H, E484K và K417N. Đột biến ký nước D614G phổ biến xuất hiện trên S của các dòng biến thể Alpha, Beta, Gamma, Delta, Delta Plus, Iota, Kappa và các dòng biến thể N1-N9 chưa có mã WHO với xác suất 39,29% làm virus có khả năng lây nhiễm cao và không kháng vaccine. Đột biến trung tính uracil N501Y xuất hiện trên S của các dòng biến thể Alpha, Beta, Gamma với xác suất 14,29% làm virus có khả năng lây nhiễm tăng và ít kháng vaccine. Đột biến uracil L452R xuất hiện trên S của các dòng biến thể Delta, Delta Plus, Kappa và N8 với xác suất 11,90% làm virus có khả năng lây nhiễm cao và ít kháng vaccine. Đột biến uracil P681R xuất hiện trên S của các dòng biến thể Delta, Delta Plus, Kappa với xác suất 10,71% làm virus có khả năng lây nhiễm cao, gây bệnh nặng hơn và ít kháng vaccine. Đột biến uracil P681H xuất hiện trên S của các dòng biến thể Alpha, N2 và N9 với xác suất 9,52% làm virus có khả năng lây nhiễm tăng, gây bệnh nặng hơn và không kháng vaccine. Đột biến trung tính uracil E484K xuất hiện trên S của các dòng biến thể Beta, Gamma, Iota với xác suất 8,33% làm virus có khả năng lây nhiễm tăng và ít kháng vaccine. Đột biến trung tính uracil K417N xuất hiện trên S của dòng biến thể Delta Plus với xác suất 5,95% làm virus có khả năng lây nhiễm rất cao và ít kháng vaccine.

### TÀI LIỆU THAM KHẢO

- [1] Trương Thế Quang (2018), *Tin sinh học (Bioinformatics)*, Nxb Đại học Quốc gia Thành phố Hồ Chí Minh.
- [2] Kim D., Lee J.Y., Yang J.S., Kim J.W., Kim V.N., Chang H. (2020), *The architecture of SARS-CoV-2 transcriptome*, Cell, 181(4).
- [3] Parlikar A., Kalia K., Sinha S., Patnaik S., Sharma N., Vemuri S.G., Sharma G.J.P. (2020), *Understanding genomic diversity, pan-genome and evolution of SARS-CoV-2*, PeerJ., 8.
- [4] Báo Tin tức Thông tấn xã Việt Nam (2021), *Biến thể virus Delta Plus nguy hiểm như thế nào?*, <http://bvdkt.vn/news/view/Bien-the-virus-Delta-Plus-nguy-hiem-nhu-the-nao/>, ngày truy cập: 26-9-2021.
- [5] Centers for Disease Control and Prevention (2021), *SARS-CoV-2 Variant Classifications and Definitions 21-9-2021*, <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html>, ngày truy cập: 25-9-2021.

- [6] Centers for Disease Control and Prevention (2021), *SARS-CoV-2 Variant Classifications and Definitions 11-9-2021*, <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html>, ngày truy cập 25-9-2021.
- [7] Hồng Lê (2021), *Những điều cần biết về các biến thể của SARS-CoV-2*, <https://nhandan.com.vn/khoa-hoc/nhung-dieu-can-biet-ve-cac-bien-the-cua-sars-cov-2-635321/>, ngày truy cập 26-9-2021.
- [8] IsofHcare (2021), *Thông tin về các loại biến thể SARS-CoV-2 hiện nay*, <https://isofhcare.com/thong-tin-ve-cac-loai-bien-the-cua-covid-hien-nay--0>, ngày truy cập 25-9-2021.
- [9] Novus Biologicals (2021), *SARS-CoV-2 Research Resources*, <https://www.novusbio.com/support/sars-cov-research-resources>, ngày truy cập: 26-9-2021.
- [10] Phúc Long (2021), *Nghiên cứu Nhật xác định tính “quá nhanh, quá nguy hiểm” của chủng Delta*, <https://tuoitre.vn/nghien-cuu-nhat-xac-dinh-tinh-qua-nhanh-qua-nguy-hiem-cua-chung-delta-20210803113901238.htm>, ngày truy cập: 26-9-2021.
- [11] Trương Thế Quang (2021), *Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/VNM/VLU-07/2021 surface glycoprotein (S) gene, complete cds*, Accession OK179081, GenBank, National Center for Biotechnology Information (NCBI), USA, <https://www.ncbi.nlm.nih.gov/nucleotide/OK179081>, ngày truy cập: 25-9-2021.
- [12] Trương Thế Quang (2021), *Surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]*, Accession UAY16900, GenBank, National Center for Biotechnology Information (NCBI), USA, <https://www.ncbi.nlm.nih.gov/protein/2096275948>, ngày truy cập: 24-9-2021.
- [13] Worldometer (2021), *COVID-19 Coronavirus pandemic last updated: Oct 27, 2021, 00:22 GMT*, <https://www.worldometers.info/coronavirus/>, ngày truy cập: 27-10-2021.

Ngày nhận bài: 27-10-2021. Ngày biên tập xong: 17-12-2021. Duyệt đăng: 12-01-2022