

CÁC DÒNG ĐỘT BIẾN CỦA SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS 2 (SARS-CoV-2) DỰA TRÊN ORF8

MUTANT FLOWS OF SEVERE ACUTE RESPIRATORY SYNDROME CORONAVIRUS 2 (SARS-CoV-2) BASED ON ORF8

TRƯƠNG THẾ QUANG^(*)

TÓM TẮT: Mười sáu trình tự protein ORF8 của SARS-CoV-2 và một trình tự ORF8 tham chiếu YP_009724396.1 Vũ Hán, Trung Quốc được thu thập từ cơ sở dữ liệu GenBank, Mỹ. Kết quả sắp hàng hai trình tự ORF8 của SARS-CoV-2 với trình tự ORF8 tham chiếu YP_009724396.1 nhận diện được bốn dòng đột biến. Trong đó, có 15 đột biến ưa nước làm tăng độc lực và khả năng lây nhiễm của virus có xác suất xuất hiện cao nhất 48,40%, 6 đột biến kỵ nước bảo toàn tính kỵ nước của ORF8 có xác suất xuất hiện 19,35%, 4 đột biến trung tính ưa nước không làm thay đổi chức năng của ORF8 có xác suất xuất hiện 12,90% và 6 đột biến trung tính kỵ nước không làm thay đổi chức năng của ORF8 có xác suất xuất hiện 19,35%. Chứng tỏ ORF8 là một loại protein tiến hóa cao.

Từ khóa: coronavirus hội chứng hô hấp cấp tính nghiêm trọng 2; các dòng đột biến; protein phụ ORF8.

ABSTRACT: Sixteen ORF8 protein sequences of SARS-CoV-2 and a reference ORF8 sequence YP_009724396.1 in Wuhan, China were collected from USA GenBank database. The results of pairwise sequence alignment of each SARS-CoV-2 ORF8 sequence and reference ORF8 sequence YP_009724396.1 identified four mutant flows. Of which, there are 15 hydrophilic mutations that increase the virulence and infectivity of the virus with the highest appearing probability 48.40%, 6 hydrophobic mutations that preserve the hydrophobicity of ORF8 have the appearing probability 19.35%, 4 hydrophobic neutral mutations that do not change the functions of ORF8 have appearing probability 12.90% and 6 hydrophilic neutral mutations that do not change the functions of ORF8 have appearing probability 19.35%. Demonstrates ORF8 is a highly evolved protein.

Key words: severe acute respiratory syndrome coronavirus 2; mutant flows; ORF8 accessory protein.

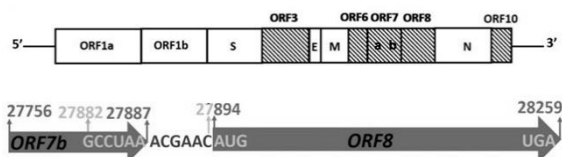
1. ĐẶT VẤN ĐỀ

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 đợt bùng phát vào tháng 12-2019 với quy mô và tốc độ lây nhiễm ngày càng cao trên toàn thế giới. Tổ chức Y tế thế giới WHO tuyên bố đợt bùng phát bệnh COVID-19 này là một đại dịch. Theo cập nhật của Worldometers về đại dịch COVID-19 trên toàn thế giới tính đến 08:07 GMT ngày 13-05-2021 đã có 222 quốc gia với 161.104.581 người bị nhiễm

bệnh và số người chết lên đến 3.345.813 [11]. 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 [3, tr.914-921] (Hình 1). Trong số các protein phụ này, ORF8 (Open Reading Frame 8) là một protein hoàn chỉnh có nhiều đột biến tự nhiên tạo ra các biến thể liên quan đến khả năng gây bệnh của SARS-CoV-2. Trình tự ORF8 của SARS-CoV-2 (ORF8) là

^(*) TS. Trường Đại học Văn Lang, quangtruongthe@gmail.com, Mã số: TCKH28-02-2021

một protein dài 121 amino acid (aa) gồm hai vùng D1 và D2. Vùng D1 là vùng peptide đầu N kỵ nước dài 15 aa (1-15 aa), vùng D2 dài 106 aa là trình tự protein (16-121 aa) mang tính thể dimer. Ngoài ra, ORF8 có chức năng liên quan đến khả năng lây nhiễm giữa các vật chủ là động vật hoặc người và hiệu quả sao chép của virus [7, tr.9-31]. Tuy nhiên, ORF8 chủ yếu hoạt động như một bộ điều biến miễn dịch bằng cách điều chỉnh giảm các phân tử MHC lớp I (Major Histocompatibility Complex class I), do đó bảo vệ các tế bào bị nhiễm bệnh chống lại tế bào T gây độc tế bào, giết chết các tế bào đích. Đồng thời, nó là một chất ức chế mạnh đường truyền tín hiệu IFN1 (Interferon 1), một thành phần chính trong phản ứng miễn dịch của vật chủ kháng virus [4, tr.43-53]. ORF8 cũng điều chỉnh phản ứng UPR (Unfolded Protein Responses) gây ra do stress ER (endoplasmic reticulum) bằng cách kích hoạt ATF6 (Activating Transcription Factor 6), do đó tăng cường khả năng sống sót của các tế bào bị nhiễm bệnh [6, tr.402-413].



Hình 1. Cấu trúc bộ gene SARS-CoV-2 trong đó có gene ORF8 mã hóa cho protein phụ ORF8 [5, tr.677]

SARS-CoV-2 đã gây ra mối đe dọa nghiêm trọng đối với sức khỏe con người nhưng nhiều khía cạnh sinh học cơ bản của nó vẫn chưa được biết đến. Bộ gene của nó mã hóa các gene phụ khác biệt đáng kể trong các coronavirus và góp phần vào khả năng gây bệnh của virus. Trong số các gene phụ, ORF8 nổi bật bởi tính biến đổi cao và cho thấy những thay đổi cấu trúc được nghi ngờ có liên quan đến khả năng lây lan của virus. Tuy nhiên, nhiều chức năng của ORF8 vẫn chưa được làm sáng tỏ, khiến nó ít được nghiên cứu hơn các gene SARS-CoV-2 khác. ORF8 tác động đến các quá trình khác nhau của vật chủ và thay đổi các chức năng khác nhau thông qua đột biến để né tránh các phản ứng miễn dịch của vật chủ, vì vậy

điều cần thiết là phải nghiên cứu các đột biến ORF8 để hiểu rõ hơn về khả năng lây nhiễm của virus và ứng dụng để phát triển các loại vaccine kháng virus mạnh chống lại SARS-CoV-2 [2, tr.672-685]. Nghiên cứu đã xác định được bốn dòng đột biến trên các biến thể ORF8 trong quá trình tiến hóa, một số đột biến đã làm thay đổi cấu trúc và chức năng của ORF8 theo hướng tăng độc lực của virus cũng như tăng khả năng lây nhiễm trên vật chủ là người.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Thu thập trình tự ORF8 của các biến chủng SARS-CoV-2

Các biến thể trình tự protein ORF8 (ORF8) của SARS-CoV-2 ở người (SARS-CoV-2) và SARS-CoV-2 tham chiếu YP 009724396.1 Vũ Hán, Trung Quốc được thu thập từ cơ sở dữ liệu GenBank thuộc Trung tâm Thông tin Công nghệ sinh học Quốc gia NCBI, Hoa Kỳ (The National Center for Biotechnology Information, USA).

2.2. Nhận diện các vị trí đột biến trên ORF8

Mỗi ORF8 của SARS-CoV-2 được sắp hàng hai trình tự [1, tr.93-95] với ORF8 của SARS-CoV-2 tham chiếu YP 009724396.1 Vũ Hán, Trung Quốc, 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, những thay đổi về cấu trúc và chức năng của ORF8 tương ứng [7, tr.9-31]. Cần lưu ý rằng một đột biến thay thế một amino acid A1 bằng A2 tại vị trí p trên một biến thể ORF8 của SARS-CoV-2 được ký hiệu là A1pA2.

2.3. Tỷ lệ tương đồng

Tỷ lệ tương đồng (Identity) I_{xs} (%) của biến thể x so với tham chiếu s dựa trên trình tự ORF8 được ước lượng theo công thức (1).

$$I_{xs}(\%) = \frac{100.M_{xs}}{L} \quad (1)$$

Trong đó: I_{xs} (%) là tỷ lệ tương đồng ORF8 của biến thể x so với ORF8 tham chiếu s, M_{xs} là tổng số đột biến trên trình tự ORF8 của biến thể x khi sắp hàng hai trình tự với ORF8 tham

chiều s, L = 121 (aa) là chiều dài của trình tự ORF8 trên biến thể x hoặc tham chiếu s.

2.4. Xác suất xuất hiện đột biến

Xác suất xuất hiện đột biến k trên ORF8 được ước lượng theo công thức (2).

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

Trong đó: $P_k(\%)$ là xác suất xuất hiện đột biến k trên ORF8, M_k là số đột biến k xuất hiện trên ORF8, M là tổng số đột biến xuất hiện trên ORF8 của tất cả các biến thể.

2.5. 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 mẫu. Xử lý số liệu, so sánh các giá trị trung bình mẫu bằng phương pháp phân tích phương sai ANOVA (Analysis of Variance) với xác suất tin cậy 95%.

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Danh mục các trình tự ORF8

Các biến thể ORF8 của SARS-CoV-2 và ORF8 tham chiếu YP_009724396.1 được thu thập từ cơ sở dữ liệu GenBank (NCBI, USA) (Bảng 1).

Bảng 1. Danh mục các trình tự và trình tự tham chiếu ORF8

STT	Version	Nơi lấy mẫu	Các đột biến
1	YP_009724396.1	China: Wuhan	Trình tự tham chiếu
2	QUN07725.1 [9], [10]	Việt Nam	
3	QJR88780.1	Australia: Victoria	L84S
4	QLH57924.1	USA: Florida	L84S, V62L, F16L
5	QKG86865.1	USA: Massachusetts	L84S, V62L, P36S
6	QJR88936.1	Australia: Victoria	L84S, V62L
7	QMU91550.1	USA: Florida	L84S, V62L
8	QJD48694.1	USA: Washington	L84S, H112Q
9	QMU91334.1	USA: Florida	L84S, D63G
10	QKU37052.1	Saudi Arabia: Jeddah	L84S, W45L
11	QMS54342.1	USA: Washington, Quận King	L84S, E110Q
12	QMS53022.1	USA: Washington, Quận Yakima	L84S, T11I
13	QMT50804.1	USA: Washington, Quận Spokane	L84S, E19D
14	QKV06506.1	USA: Washington, Quận Mason	L84S, S67F
15	QLY90504.1	USA	S24L, S67F
16	QLH58953.1	USA: Florida	S24L, P38S
17	QMU92030.1	USA: Florida	G8R, D35Y

3.2. Nhận diện các đột biến tự nhiên trên ORF8

Nhận diện các đột biến tự nhiên trên các biến thể ORF8 của SARS-CoV-2 bằng cách giải bài toán sắp hàng hai trình tự ORF8 của SARS-CoV-2 với trình tự ORF8 tham chiếu YP_009724396.1 của SARS-CoV-2 Vũ Hán, Trung Quốc, ứng dụng công cụ BLASTp (NCBI, USA). Kết quả sắp hàng hai trình tự nhận diện được các đột biến tự nhiên trên ORF8 trong quá trình tiến hóa như sau: trình tự QUN07725.1 Việt Nam có tỷ lệ tương đồng 100%, nên không có đột biến nào trên ORF8 so với trình tự tham chiếu YP_009724396.1 Vũ Hán, Trung Quốc.

3.2.1. Trình tự ORF8 có một đột biến

Trình tự QJR88780.1 được lấy mẫu ngày 20-03-2020 tại Victoria, Australia có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu, sở hữu một đột biến L84S làm thay đổi cả cấu trúc và chức năng của protein ORF8 theo chiều hướng tăng độc lực gây bệnh cho vật chủ nặng hơn.

3.2.2. Trình tự ORF8 có hai đột biến

Trình tự QJR88936.1 được lấy mẫu ngày 20-03-2020 tại Victoria, Australia có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu. Trình tự này có hai đột biến L84S và V62L.

Đột biến ư nước L84S làm tăng độc lực của virus và đột biến trung tính V62L bảo toàn chức năng kỵ nước của ORF8.

Trình tự QMU91550.1 được lấy mẫu ngày 24-03-2020 tại Florida, USA có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu. Trình tự này sở hữu hai đột biến L84S và V62L giống như trình tự QJR88936.1 làm thay đổi chức năng của ORF8 theo chiều hướng tăng độc lực của virus và bảo toàn chức năng kỵ nước.

Tương tự, trình tự QKU37052.1 được lấy mẫu ngày 15-03-2020 tại Jeddah, Saudi Arabia có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu. Trình tự này có hai đột biến L84S và W45L làm thay đổi cả cấu trúc và chức năng của protein ORF8. Đột biến L84S làm tăng độc lực gây bệnh cho vật chủ nặng hơn, đột biến trung tính W45L làm ORF8 giữ nguyên tính kỵ nước.

Trình tự QMS53022.1 được lấy mẫu ngày 16-05-2020 tại Quận Yakima, Washington, USA có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu. Trình tự này có hai đột biến L84S và T11I. Đột biến T11I thay thế threonine (T) ư nước bằng isoleucine (I) kỵ nước, do đó làm thay đổi chức năng của ORF8 theo hướng kỵ nước hơn, còn đột biến L84S làm tăng độc lực của virus.

Trình tự QMT50804.1 được lấy mẫu ngày 09-06-2020 tại Quận Spokane, Washington, USA có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu. Trình tự này có hai đột biến L84S và E19D. Đột biến L84S làm thay đổi chức năng của ORF8 theo hướng tăng độc lực của virus, còn E19D là đột biến trung tính ư nước.

Trình tự QMU91334.1 được lấy mẫu ngày 13-03-2020 tại Florida, USA có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu, có hai đột biến D63G làm thay đổi chức năng của ORF8 từ ư nước sang kỵ nước và L84S làm tăng độc lực của virus trong quá trình tiến hóa.

Trình tự QKV06506.1 được lấy mẫu ngày 02-04-2020 tại Quận Mason, Washington, USA có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu. Trình tự này sở hữu hai đột biến L84S làm

tăng độc lực của virus và S67F làm thay đổi chức năng của ORF8 theo hướng kỵ nước hơn.

Trình tự QMS54342.1 được lấy mẫu ngày 04-5-2020 tại Quận King, Washington, USA có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu. Trình tự này sở hữu hai đột biến L84S làm tăng độc lực của virus và đột biến trung tính E110Q làm thay đổi chức năng của ORF8 theo hướng ư nước hơn.

Trình tự QJD48694.1 Washington, USA có tỷ lệ tương đồng 98,35% với trình tự tham chiếu, sở hữu hai đột biến L84S và H112Q. Đột biến L84S làm tăng độc lực virus gây bệnh trên người nặng hơn, đột biến trung tính ư nước H112Q không làm thay đổi chức năng của ORF8.

Trình tự QLH58953.1 được lấy mẫu ngày 02-06-2020 tại Florida, USA có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu, sở hữu hai đột biến kỵ nước S24L và ư nước P38S làm thay đổi chức năng của ORF8 theo hướng tăng khả năng lây nhiễm virus.

Trình tự QLY90504.1 được lấy mẫu ngày 14-05-2020 tại USA có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu. Trình tự này có hai đột biến kỵ nước S24L và S67F làm thay đổi chức năng của ORF8 theo hướng kỵ nước hơn.

Trình tự QMU92030.1 được lấy mẫu ngày 31-03-2020 tại Florida, USA có tỷ lệ tương đồng 98,35% so với trình tự tham chiếu. Trình tự này có hai đột biến ư nước G8R và trung tính D35Y làm thay đổi cấu trúc và chức năng của protein ORF8 theo hướng ư nước.

3.2.3. Trình tự ORF8 có ba đột biến

Trình tự QLH57924.1 được lấy mẫu ngày 13-04-2020 tại Florida, USA có tỷ lệ tương đồng 97,52% so với trình tự tham chiếu. Trình tự này có ba đột biến L84S, V62L và F16L. Đột biến L84S làm tăng độc lực và thúc đẩy sự tồn tại của virus. V62L và F16L là các đột biến trung tính (neutral mutation) nên không làm thay đổi chức năng kỵ nước của ORF8, vì V, F và L đều là các amino acid kỵ nước).

Tương tự, trình tự QKG86865.1 được lấy mẫu ngày 26-03-2020 tại Massachusetts, USA có tỷ lệ tương đồng 97,52% so với trình tự tham chiếu. Trình tự này có ba đột biến L84S, V62L và P36S. Trong đó, L84S và P36S là đột biến ưa nước làm tăng độc lực của virus, V62L

là đột biến trung tính nên chức năng kỵ nước của ORF8 được bảo toàn.

3.3. Các dòng đột biến tự nhiên trong quá trình tiến hóa ORF8

3.3.1. Dòng đột biến I

Bảng 2. Các biến thể SARS-CoV-2 thuộc dòng đột biến I

Tham chiếu	Đột biến 1	Đột biến 2	Biến thể 2 đột biến	Đột biến 3	Biến thể 3 đột biến
YP_009724396.1 TQ QUN07725.1 VN	→ L84S	→ V62L	→ QJR88936.1 Australia	→ F16L	→ QLH57924.1 USA
			→ QMU91550.1 USA	→ P36S	→ QKG86865.1 USA
		→ W45L	→ QKU37052.1 Saudi Arabia		
		→ T11I	→ QMS53022.1 USA		
		→ E19D	→ QMT50804.1 USA		
		→ D63G	→ QMU91334.1 USA		
		→ S67F	→ QKV06506.1 USA		
		→ E110Q	→ QMS54342.1 USA		
		→ H112Q	→ QJD48694.1 USA		
			→ QJR88780.1 Australia		

Dòng đột biến I gồm SARS-CoV-2 tham chiếu YP_009724396.1 Trung Quốc (TQ), QUN07725.1 Việt Nam (VN), trình tự ORF8 của hai biến thể này có tỷ lệ tương đồng 100% và 12 biến thể SARS-CoV-2 có đột biến L84S [8]. Trong đó, có một biến thể QJR88780.1 Australia mang một đột biến (L84S); 9 biến thể mang hai đột biến là QJR889036.1 Australia (L84S, V62L), QMU91550.1 USA (L84S, V62L), QJD48694.1 USA (L84S, H112Q), QMS54342.1 USA (L84S, E110Q), QMU91334.1 USA (L84S, D63G), QKU37052.1 Saudi Arabia (L84S, W45L), QMT50804.1 USA (L84S, E19D), QMS53022.1

USA (L84S, T11I), QKV06506.1 USA (L84S, S67F) và hai biến thể mang ba đột biến là QLH57924.1 USA (L84S, V62L, F16L), QKG86865.1 USA (L84S, V62L, P36S) (Bảng 2).

3.3.2. Dòng đột biến II

Dòng đột biến II gồm SARS-CoV-2 tham chiếu YP_009724396.1 Trung Quốc, QUN07725.1 Việt Nam và hai biến thể đều mang đột biến S24L là QLH58953.1 USA có hai đột biến (S24L, P38S) và QLY90504.1 USA có hai đột biến (S24L, S67F) (Bảng 3).

Bảng 3. Các biến thể SARS-CoV-2 thuộc dòng đột biến II

Tham chiếu	Đột biến 1	Đột biến 2	Biến thể 2 đột biến
YP_009724396 Trung Quốc QUN07725.1 Việt Nam	→ S24L	→ P38S	→ QLH58953.1 USA
		→ S67F	→ QLY90504.1 USA

3.3.3. Dòng đột biến III

Dòng đột biến III gồm SARS-CoV-2 tham chiếu YP_009724396.1 Trung Quốc, QUN07725.1

Việt Nam và biến thể QMU92030.1 USA có hai đột biến (G8R, D35Y) (Bảng 4).

Bảng 4. Các biến thể SARS-CoV-2 thuộc dòng đột biến III

Tham chiếu	Đột biến 1	Đột biến 2	Biến thể 2 đột biến
YP_009724396 Trung Quốc QUN07725.1 Việt Nam	→ G8R	→ D35Y	→ QMU92030.1 USA

3.3.4. Dòng đột biến IV

Dòng đột biến IV được tách ra từ dòng đột biến I và dòng đột biến II gồm SARS-CoV-2 tham chiếu YP_009724396.1 Trung Quốc, QUN07725.1

Việt Nam và hai biến thể đều mang đột biến S67F là QKV06506.1 USA có hai đột biến (S67F, L84S) và QLY90504.1 USA có hai đột biến (S67F, S24L) (Bảng 5).

Bảng 5. Các biến thể SARS-CoV-2 thuộc dòng đột biến IV

Tham chiếu	Đột biến 1	Đột biến 2	Biến thể 2 đột biến
YP_009724396 Trung Quốc QUN07725.1 Việt Nam	→ S67F	→ L84S	→ QKV06506.1 USA
		→ S24L	→ QLY90504.1 USA

3.4. Nhận xét về các dòng đột biến ORF8

Các đột biến thay thế được tìm thấy trên trình tự ORF8 bắt đầu từ vị trí amino acid 8 đến 112. Số lượng đột biến điểm thay thế amino acid đã được phát hiện trên ORF8 từ 1 đến 3. Đột biến L84S làm tăng độc lực của virus thuộc dòng đột biến I có xác suất xuất hiện cao nhất $12/31 = 38,71\%$. Đột biến trung tính V62L tại vị trí 62 valine (V) kỵ nước được thay thế bằng leucine (L) kỵ nước do đó hầu như giữ nguyên tính kỵ nước của ORF8, đột biến V62L thuộc dòng đột biến I có xác suất xuất hiện $4/31 = 12,90\%$. Đột biến S24L tại vị trí 24 leucine (L) kỵ nước thay thế serine (S) ưa nước, làm thay đổi cấu trúc và chức năng ORF8 theo hướng kỵ nước thuộc dòng đột

biến II, đột biến S24L có xác suất xuất hiện $2/31 = 6,45\%$. Đột biến S67F tại vị trí 67 serine (S) ưa nước được thay thế bằng phenylalanine (F) kỵ nước, làm thay đổi cấu trúc và chức năng ORF8 theo hướng kỵ nước thuộc dòng đột biến IV, đột biến S67F có xác suất xuất hiện $2/31 = 6,45\%$. Các đột biến còn lại đều có xác suất xuất hiện $1/31 = 3,22\%$. Các amino acid của ORF8 dễ bị đột biến thay thế là leucine (L) với xác suất xuất hiện 38,71%, valine (V) 12,90%, serine (S) 12,90%; còn proline (P), glutamic acid (E) và aspartic acid (D) đều có xác suất xuất hiện 6,45%. Điều này chứng tỏ ORF8 đang nhanh chóng trải qua những thay đổi đột biến, cho thấy ORF8 là một loại protein có đột biến tiến hóa cao.

Bảng 6. Phân loại các đột biến trên ORF8

Loại đột biến	Chức năng của ORF8	Tên đột biến	Xác suất xuất hiện
Ưa nước	Tăng độc lực và lây nhiễm	12×L84S, P38S, P36S, G8R	48,40%
Kỵ nước	Bảo toàn tính kỵ nước	T11I, D63G, 2×S24L, 2×S67F	19,35%
Trung tính ưa nước	Không thay đổi	D35Y, E19D, E110Q, H112Q	12,90%
Trung tính kỵ nước	Không thay đổi	F16L, 4×V62L, W45L	19,35%

Kết quả cho thấy, có 15 đột biến ưa nước làm tăng độc lực và khả năng lây nhiễm của virus có xác suất xuất hiện cao nhất 48,40%, 6 đột biến kỵ nước bảo toàn tính kỵ nước của ORF8 có xác suất xuất hiện 19,35%, 4 đột biến trung tính ưa nước không làm thay đổi chức năng của ORF8 có xác suất xuất hiện 12,90% và 6 đột biến trung tính ưa nước không làm thay đổi chức năng của ORF8 có xác suất xuất hiện 19,35% (Bảng 6).

4. KẾT LUẬN VÀ KIẾN NGHỊ

4.1. Kết luận

Các trình tự protein ORF8 của SARS-CoV-2 và một trình tự ORF8 tham chiếu YP_009724396.1 Vũ Hán, Trung Quốc được thu thập từ cơ sở dữ liệu GenBank, Mỹ.

Kết quả sắp hàng hai trình tự ORF8 của từng biến thể SARS-CoV-2 với trình tự ORF8 tham chiếu YP_009724396.1 nhận diện được bốn dòng đột biến. Có 15 đột biến ưa nước làm tăng độc lực và khả năng lây nhiễm của virus có xác suất xuất hiện cao nhất 48,40%, 6 đột biến kỵ nước bảo toàn tính kỵ nước của ORF8 có xác suất xuất hiện 19,35%, 4 đột biến trung tính kỵ

nước không làm thay đổi chức năng của ORF8 có xác suất xuất hiện 12,90% và 6 đột biến trung tính ưa nước không làm thay đổi chức năng của ORF8 có xác suất xuất hiện 19,35%.

4.2. Kiến nghị

Tiếp tục nghiên cứu về các đột biến tại các vùng D1 và D2 của protein phụ SARS-CoV-2

ORF8 trong quá trình tiến hóa làm thay đổi cấu trúc và chức năng hoạt động của ORF8. Kết quả nghiên cứu sẽ giúp hiểu biết thêm về chức năng và cơ chế hoạt động điều biến miễn dịch của ORF8 trong quá trình gây bệnh của SARS-CoV-2 trên người.

TÀI LIỆU THAM KHẢO

- [1] Trương Thế Quang (2018), *Tin sinh học (Bioinformatics)*, Nxb Đại học Quốc gia Thành phố Hồ Chí Minh.
- [2] Frieman M., Baric R. (2008), *Mechanisms of severe acute respiratory syndrome pathogenesis and innate immunomodulation*, Microbiol. Mol. Biol. Rev., 72.
- [3] 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).
- [4] Kumar S., Nyodu R., Maurya V.K., Saxena S.K. (2020), *Host immune response and immunobiology of human SARS-CoV-2 infection*, Coronavirus Disease 2019 (COVID-19) Springer.
- [5] Sameer M., Abderrezak B., Bothina M.A., Mamoon R., Tanveer S.K., Nusaibah S. G. and Shuja S.M. (2020), *SARS-CoV-2 ORF8 and SARS-CoV ORF8ab: Genomic Divergence and Functional Convergence*, Pathogens, 9(9).
- [6] Sung S.C., Chao C.Y., Jeng K.S., Yang J.Y., Lai M.M. (2009), *The 8ab protein of SARS-CoV is a luminal ER membrane-associated protein and induces the activation of ATF6*, Virology, 387.
- [7] 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.
- [8] Hosseini Rad Sm. A., McLellan A.D. (2020), *Implications of SARS-CoV-2 mutations for genomic RNA structure and host microRNA targeting*, Int. J. Mol. Sci., 21, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7370282/>, ngày truy cập: 02-5-2021.
- [9] Trương Thế Quang (2021), *Single nucleotide polymorphism of Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strains based on ORF8 gene*, Accession MZ067003, GenBank, National Center for Biotechnology Information (NCBI), USA, <https://www.ncbi.nlm.nih.gov/nuccore/MZ067003>, ngày truy cập: 03-5-2021.
- [10] Trương Thế Quang (2021), *Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/VNM/VLU/2020 ORF7b (ORF7b) gene, partial cds; and ORF8 protein (ORF8) gene, complete cds*, European Nucleotide Archive, London, England, <https://www.ebi.ac.uk/ena/browser/view/MZ067003>, ngày truy cập: 05-5-2021.
- [11] Worldometer (2021), *COVID-19 Coronavirus pandemic*, <https://www.worldometers.info/coronavirus/>, ngày truy cập: 13-5-2021.

Ngày nhận bài: 13-5-2021. Ngày biên tập xong: 20-6-2021. Duyệt đăng: 24-7-2021