

## VAI TRÒ CỦA BIẾN THỂ DI TRUYỀN TRONG BỆNH LÝ COVID-19

● Vũ Diễm My\*<sup>1</sup> ● Huỳnh Anh Phương<sup>1</sup> ● Nguyễn Nhật Quỳnh Như<sup>1</sup>  
● Hồ Quốc Chương<sup>1</sup> ● Lê Thái Khương<sup>1</sup> ● Hoàng Minh<sup>2</sup>

<sup>1</sup> Trung tâm Y Sinh học phân tử, Đại học Y Dược Thành phố Hồ Chí Minh

<sup>2</sup> Trung tâm Cúm quốc gia, Viện Pasteur Thành phố Hồ Chí Minh

### TÓM TẮT

COVID-19 là bệnh lý với phổ bệnh đa dạng, biểu hiện lâm sàng và biến chứng phức tạp. Nhiều nghiên cứu đã đề cập các yếu tố dân số như giới tính, độ tuổi, bệnh lý nền, đồng nhiễm hoặc bội nhiễm với vi khuẩn làm tăng đáng kể nguy cơ diễn tiến bệnh nặng hoặc nguy kịch trên người bệnh COVID-19. Tuy nhiên, những thông số này chưa đủ để giải thích tiên lượng bệnh khác nhau trên những bệnh nhân có đặc điểm nền tương đồng, hoặc các trường hợp nặng đến nguy kịch thuộc nhóm người khỏe mạnh trẻ tuổi. Bên cạnh đó, sự xuất hiện của các ca lây nhiễm đột phá, hiệu quả chủng ngừa thay đổi phụ thuộc chủng tộc cũng cho thấy vai trò của yếu tố di truyền trong kiểm soát và phòng ngừa dịch bệnh. Các phân tích bộ gen của bệnh nhân COVID-19 với độ bệnh khác nhau đã ghi nhận một số biến thể đặc trưng trên các gen tham gia đáp ứng miễn dịch, mã hóa thụ thể tế bào có tương quan đáng kể với sự nhạy cảm và độ bệnh nặng. Ngoài ra, các locus gen liên hệ với tiến triển bệnh, cơn bão cytokine cũng đã được xác nhận. Những kiến thức này là cơ sở cho việc xác định các cơ chế phân tử phục vụ việc phát triển các liệu pháp kháng SARS-CoV-2, sàng lọc và điều trị sớm những bệnh nhân có nguy cơ cao.

**Từ khóa:** biến thể di truyền, COVID-19.

## THE ROLE OF GENETIC VARIATION IN THE PATHOGENESIS OF COVID-19

● Vu Diem My ● Huynh Anh Phuong ● Nguyen Nhat Quynh Nhu  
● Ho Quoc Chuong ● Le Thai Khuong, Hoang Minh

### ABSTRACT

COVID-19 is a disease with a diverse spectrum of clinical manifestations and complications. Many studies have shown that demographic factors such as sex, age, underlying medical conditions, co-infection or superinfection with bacteria significantly increase the risk of severe or critical illness in COVID-19 patients. However, these parameters are not sufficient to explain different disease prognosis in patients with similar background characteristics, or severe to critical conditions in young and healthy subjects. In addition, the occurrence of breakthrough cases, the effectiveness of vaccination varies among populations that also shows the role of genetic factors in disease control and prevention. Genome analysis of COVID-19 patients at different levels of disease have showed numerous genetic variants of immune response or cellular receptor genes that are significantly associated with susceptibility and severe illness. In addition, gene loci correlated with disease progression and cytokine storm have also been confirmed. This knowledge provided insights of molecular mechanisms that can be used in the development of anti-SARS-CoV-2 therapies and early treatment of high-risk patients.

\* Tác giả liên hệ: TS. Vũ Diễm My, Email: diemmyvu@ump.edu.vn

(Ngày nhận bài: 12/10/2022; Ngày nhận bản sửa: 29/10/2022; Ngày duyệt đăng: 10/11/2022)

**Keywords:** *genetic polymorphism, COVID-19 susceptibility and severe illness.*

## 1. NGUỒN GỐC VÀ SỰ TIẾN HÓA CỦA SARS-CoV-2

COVID-19 là căn bệnh nguy hiểm có tính lây nhiễm cao gây ra bởi chủng virút hô hấp SARS-CoV-2. Theo thông tin dịch tễ học, SARS-CoV-2 ban đầu được ghi nhận tại thành phố Vũ Hán vào tháng 12/2019 tại Vũ Hán (Trung Quốc), trên một nhóm bệnh nhân với triệu chứng viêm phổi không rõ nguyên nhân làm việc tại chợ hải sản Hoa Nam. Virút được cho là có nguồn gốc từ động vật, trải qua nhiều quá trình đột biến, tái tổ hợp và thích nghi trên các vật chủ khác nhau, chúng được lây truyền sang người.

SARS-CoV-2 thuộc họ *Coronaviridae* bao gồm các virút có bộ gen là RNA mạch dương. Dựa trên đặc điểm di truyền và huyết thanh học, họ virút này được chia thành bốn chi: *Alphacoronavirus* (*alpha-CoV*), *Betacoronavirus* (*beta-CoV*), *Gammacoronavirus* (*gamma-CoV*), và *Deltacoronavirus* (*delta-CoV*) phân kỳ ở giai đoạn 2400-3000 năm trước công nguyên và có khuynh hướng lây nhiễm trên các nhóm động vật khác nhau [1]. Các virút thuộc chi *alpha* và *beta* chủ yếu được tìm thấy ở động vật hữu nhũ, trong khi nhóm virút ở hai chi còn lại phần lớn gây bệnh trên động vật lông vũ. Nhóm virút corona gây đại dịch hô hấp ở người bao gồm SARS-CoV-1, MERS-CoV hay SARS-CoV-2 đều thuộc chi *Betacoronavirus*, phân chi *Sarbecovirus*. Một số chủng virút corona khác cũng đã được ghi nhận gây bệnh trên người gồm có HCoV-229E và HCoV-NL63 thuộc chi *alpha*; HCoV-OC43 và HCoV-HKU1 thuộc chi *beta*, phân chi *Embecovirus*.

Đến thời điểm hiện tại, các nhà nghiên cứu vẫn chưa thể khẳng định nguồn gốc chính xác của SARS-CoV-2. RaTG13 là chủng virút corona có cấu trúc bộ gen tương đồng nhất (~96%) được thu nhận trên dơi tại Trung Quốc vào năm 2013. Từ thông tin này, giả thuyết về cội nguồn của SARS-CoV-2 đã được thiết lập, xoay quanh một số luận điểm: (1) chủng SARS-CoV-2 tiền thân lưu hành trên dơi với các đột biến gen được tích lũy theo thời gian, sau đó lây nhiễm trên người cách ngẫu nhiên; (2) sự tái tổ hợp gen giữa RaTG13 và chủng CoV chưa xác định xảy ra trên vật chủ trung gian tạo thành SARS-CoV-2, và virút được lây truyền sang người; (3) bộ gen SARS-CoV-2 là dạng thể khảm kết hợp giữa RaTG13 và chủng corona có nguồn gốc từ tê tê ở Quảng Đông [1]. Bên cạnh đó, tuy các thông tin về SARS-CoV-2 là sản phẩm tạo ra bằng kỹ thuật di truyền và “trốn thoát” khỏi phòng thí nghiệm được lan truyền rộng rãi, phần lớn các bằng chứng khoa học đến thời điểm hiện tại đều cho thấy virút có nguồn gốc từ động vật và lây truyền sang người [2].

Hiện tượng tái tổ hợp diễn ra khá thường xuyên trên virút corona mang ý nghĩa quan trọng cho sự tiến hóa và thích nghi trên các nhóm vật chủ khác nhau. Nhìn chung, tỷ lệ tái tổ hợp các chủng corona lây nhiễm trên người vẫn thấp hơn tỷ lệ tiến hóa thông thường từ 10-20 lần, dao động khoảng  $1 \times 10^{-5}$  tại mỗi vị trí qua từng năm [3]. Phân tích phát sinh loài cho thấy có sự tái tổ hợp xảy ra giữa chủng vi rút MERS-CoV từ người và lạc đà, một số dữ liệu cũng ghi nhận hiện tượng này trong quá trình tiến hóa của SARS-CoV-1. Trong trường hợp của SARS-CoV-2, virút được báo cáo có sự tái tổ hợp xảy ra mức độ thấp hơn hai loại trên. Cụ thể, khi phân tích khoảng 1,6 triệu trình tự, nhóm của Lanfear và Corbett-Detig xác định 598 sự kiện tái tổ hợp tương ứng với 2.7% cấu trúc bộ gen, xảy ra chủ yếu tại vị trí mã hóa protein S [4]. Sự kết hợp của yếu tố ức chế hay ngăn trở tại các vùng gen liên quan được cho là nguyên nhân làm giảm sự xuất hiện tái tổ hợp. Trái ngược với xu hướng trên, trình tự mã hóa của SARS-CoV-2 có sự thay thế xấp xỉ  $1,1 \times 10^{-3}$  vị trí/năm, cao hơn so với MERS-CoV là  $6,9 \times 10^{-4}$  hay HCoV-OC43 là  $2,1 \times 10^{-4}$  [3]. Trong đó, tỷ lệ thay thế sai nghĩa so với đồng nghĩa là 0,028, trong khi con số này ở bộ gen người là 0,2 [1]. Dựa trên tỷ lệ thay thế, các nhà khoa học ước tính thời gian phân kỳ giữa SARS-CoV-2 và RaTG13 diễn ra trong khoảng 18 đến 71 năm; rơi vào một trong những mốc thời gian 1948, 1969 hay 1982 [5]. Những thông tin trên cho thấy tuy bộ gen SARS-

CoV-2 có hiện tượng thay thế trình tự diễn ra ở mức độ cao hơn, nhưng vẫn mang tính ổn định do các biến đổi này phần lớn là dạng đồng nghĩa, cùng với việc tái tổ hợp xảy ra ở mức độ thấp.

## 2. ĐẶC ĐIỂM CẤU TẠO VÀ CƠ CHẾ XÂM NHIỄM CỦA SARS-CoV-2

SARS-CoV-2 có cấu trúc virion dạng hình cầu, với đường kính dao động từ 60-140 nm. Bộ gen của virus là RNA sợi đơn dương có chiều dài từ 26-32kb, bao gồm 14 khung đọc mở (Open Reading Frame - ORF) mã hóa cho 29 loại protein khác nhau [6]. Khoảng 2/3 trình tự nằm về phía đầu 5 của virus mã hóa cho hai polyprotein có cấu trúc chồng lấp là pp1a và pp1b sẽ được phân cắt để tạo thành 16 protein không cấu trúc (Non structural protein - NSP) tham gia vào quá trình sao chép và phiên mã của virus. Các NSP có chức năng khác nhau trong quá trình nhân bản, điển hình như: NSP1 - phân hủy mRNA tế bào vật chủ, ức chế dịch mã hệ thống; NSP2 -ngăn trở chu trình tế bào; NSP3, NSP4, NSP6 - hình thành khoang trung gian sao chép; NSP5 - protease chủ đạo; NSP7 - phức hợp sao chép; NSP8 - primase; NSP9 - RNA bắt cặp protein; NSP10 - đồng yếu tố của NSP14 và NSP16, NSP11 - chưa rõ chức năng, NSP12 - RNA polymerase phụ thuộc RNA; NSP13 - helicase, 5' phosphatase; NSP14 - 3'-5' exonuclease, N7-MTase; NSP15 - endonuclease; và NSP16 - gắn mũ mRNA, O-Mtase[7]. Phần trình tự còn lại về phía đầu 3' mã hóa cho các protein cấu trúc như nucleocapsid (N), protein gai (S), protein màng (M) và protein vỏ (E). Các protein này tham gia vào quá trình đóng gói của virion và ức chế phản ứng miễn dịch của vật chủ. Các gen mã hóa cho các protein phụ (ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8b, ORF9b và ORF14) nằm xen lẫn với các gen mã hoá protein cấu trúc. Ngoại trừ protein ORF3a và ORF7a, phần lớn các protein phụ chỉ đóng vai trò điều hòa sự lây nhiễm của virus trong tế bào vật chủ, chứ không tham gia vào quá trình hình thành virion. Khi so sánh với SARS-CoV-1, bộ gen SARS-CoV-2 mang hai điểm khác biệt đáng kể. Đầu tiên, sự hiện diện của ORF8 chỉ được tìm thấy ở SARS-CoV-2 do khung đọc mở này ở SARS-CoV-1 được chia thành hai tiểu phần ORF8a và ORF8b bởi sự xuất hiện của một đột biến xóa 29 nucleotide. Thứ hai, ORF10 là khung đọc mở hoàn toàn mới chỉ xuất hiện trên SARS-CoV-2 mà không hiện diện trên SARS-CoV-1[7]. Một số nghiên cứu chức năng cho thấy ORF8 hỗ trợ việc lẩn tránh hệ thống miễn dịch của vi rút, trong khi ORF10 ức chế quá trình đáp ứng miễn dịch bẩm sinh bằng cách phân hủy protein MAVS. Thông tin này phần nào giải thích sự khác nhau về hoạt tính giữa hai chủng vi rút trên.

SARS-CoV-2 xâm nhiễm vào tế bào vật chủ thông qua quá trình nhập bào (endocytosis) hay hợp nhất màng qua trung gian thụ thể. Tương tự như MERS-CoV, SARS-CoV-2 có S được phân cắt sẵn thành hai tiểu đơn vị S1 và S2 bởi enzyme chuyển đổi tiền protein như Furin trong quá trình sinh tổng hợp. Trên một số corona khác, sự phân cắt S chỉ diễn ra khi vi rút xâm nhiễm vào tế bào đích[8]. Trong sự hợp nhất màng, tương tác giữa S1 và thụ thể ACE2 làm bộc lộ vị trí S2', qua đó hoạt hóa sự phân cắt bởi protein TMPRSS2 trên bề mặt tế bào và hình thành "lỗ" nhập bào cho phép vi rút phóng thích bộ gen vào tế bào chất của vật chủ. Ở trường hợp nhập bào, quá trình này được thực hiện bởi Cathepsin L trong tế bào chất. Sau khi được phóng thích, bộ gen virus được phóng thích và quá trình dịch mã được khởi tạo tại ORF1a và ORF1b bởi ribosome của tế bào chủ hình thành hai chuỗi polypeptide pp1a và pp1ab[9]. Hai protease PL<sup>pro</sup> và 3CL<sup>pro</sup> tiếp tục phân cắt các chuỗi này thành các protein không cấu trúc hình thành phức hợp sao chép và phiên mã hoàn chỉnh. Cùng với sự xuất hiện của phức hợp phiên mã, các bào quan sao chép được hình thành nhằm tạo điều kiện thuận lợi cho sự nhân đôi của RNA virus bởi protein RdRp (RNA-dependent RNA polymerases) và sự phiên mã các RNA thứ cấp. Các RNA mới được tổng hợp sẽ trải qua các bước biến đổi sau phiên mã và gắn mũ nhằm lẩn tránh sự phân hủy bởi hệ thống miễn dịch và gia tăng hiệu suất dịch mã. Các protein cấu trúc tổng hợp ở giai đoạn này sẽ được chuyển từ lưới nội chất đến khoang trung gian của bộ máy Golgi, tại đây chúng được đóng gói cùng với RNA virút hoàn chỉnh

(nucleocapsid) tạo các virion trưởng thành chờ xuất bào.

### 3. CÁC NGHIÊN CỨU VỀ DI TRUYỀN VẬT CHỦ TRONG BỆNH LÝ COVID-19

Đã có rất nhiều nghiên cứu được tiến hành nhằm xác định vai trò của các yếu tố di truyền trong sự nhạy cảm với SARS-CoV-2 và mối tương quan của chúng đến tiên lượng nặng của bệnh nhân. Trong đó, phần lớn các công trình chủ yếu tập trung theo ba hướng nghiên cứu chính: (1) nhóm gen liên quan đến đáp ứng miễn dịch, (2) nhóm gen mã hóa cho thụ thể và (3) nhóm các gen khác được phát hiện thông qua các nghiên cứu GWAS.

#### 3.1. Nhóm gen liên quan đến đáp ứng miễn dịch

Hệ thống miễn dịch đóng vai trò chủ đạo trong việc bảo vệ cơ thể chống lại các tác nhân gây bệnh như virus và vi khuẩn qua trung gian tế bào và các cytokine tham gia vào đáp ứng miễn dịch bẩm sinh và miễn dịch thích ứng.

**Human Leukocyte Antigen (HLA):** là nhóm các glycoprotein được tìm thấy trên bề mặt của hầu hết các tế bào trong cơ thể, song được biểu hiện nhiều nhất trên bề mặt bạch cầu. HLA đóng vai trò quan trọng trong việc nhận diện các peptide có nguồn gốc từ sinh vật ngoại lai, qua đó cảm ứng hệ thống miễn dịch loại bỏ các tế bào bị xâm nhiễm. Các nghiên cứu về cấu trúc protein S và N của SARS-CoV-2 cho thấy chúng có mang các epitope được nhận diện đặc hiệu bởi một số phân nhóm HLA ở người. Vì vậy, các khác biệt trong biểu hiện phân nhóm HLA có ảnh hưởng đến sự nhạy cảm và đáp ứng miễn dịch của bệnh nhân. Một số HLA phổ biến trên dân số Ý như HLA-A\*01:01, B\*08:01, C\*07:01, DRB1\*03:01 được chứng minh là có tương quan cao với tỷ lệ mắc bệnh và tử vong do COVID-19. Trong khi alen HLA-A\*11:01, B\*51:01 và C\*14:02 có liên quan chặt chẽ đến tiên lượng xấu trên dân số Trung Quốc[10]. HLA-DRB1\*08 cũng được báo cáo có liên quan đến tỷ lệ tử vong do COVID-19, tuy nhiên các phân tích lực liên kết cho thấy chúng không tương tác với bất kỳ peptide nào của SARS-CoV-2; do đó cần tiến hành làm rõ mối liên kết này.

**Cytokine:** là thuật ngữ dùng để miêu tả phổ rộng hoạt chất protein được tiết ra từ một vài các tế bào thuộc hệ miễn dịch với mục đích dẫn truyền các tín hiệu nội bào điều hòa sự tương tác giữa các tế bào trong cơ thể. Sự tăng đột ngột lượng lớn các cytokine, hay còn gọi là cơn bão cytokine, phản ánh sự mất kiểm soát của hệ thống miễn dịch dẫn đến phản ứng viêm quá mức kéo theo hiện tượng đông máu, giảm bạch cầu, thâm nhiễm tế bào đơn nhân các cơ quan... Bão cytokine là một trong những nguyên nhân chính gây nên tổn thương phổi, suy chức năng đa cơ quan dẫn đến tiên lượng xấu trên những trường hợp mắc COVID-19 trở nặng và nguy kịch. So với người bình thường, bệnh nhân nhiễm SARS-CoV-2 có sự gia tăng đáng kể nồng độ các cytokine như IL-1 $\alpha$ , IL-6, IL-7, IL-8, IL-9, IL-10, TNF- $\alpha$ ... Trong đó, IL-6 cao vượt ngưỡng được xem là “dấu ấn” đặc trưng trên những ca bệnh từ độ nặng trở lên. Các biến thể trên gen mã hóa IL-6 được biết là có ảnh hưởng đến nồng độ biểu hiện của cytokine này, và đã được gắn với các hội chứng rối loạn khác nhau. Strafella và cộng sự xác nhận có khoảng 7 SNPs trên gen IL-6 và 5 SNPs trên thụ thể IL-6R có liên quan đến cơ chế gây bệnh và biến chứng trên bệnh nhân COVID-19[11]. Đáng chú ý, tần suất các biến thể gen mã hoá cytokine có sự khác biệt phụ thuộc chủng tộc. Cụ thể, các alen quy định sự biểu hiện cao của IL-6 và IL-10 thường được tìm thấy ở người châu Phi, người gốc Tây Ban Nha và người châu Á hơn là người da trắng.

**Interferon (IFN):** bên cạnh nhóm gen cytokine, các gen mã hóa IFN cũng là mục tiêu tiềm năng cho các liệu pháp điều trị giảm nhẹ bệnh nhân nhiễm SARS-CoV-2. IFN liên kết với các thụ thể của chúng như IFNAR1 và IFNAR2, điều hòa chuyên biệt các con đường tín hiệu tham gia vào hoạt động kháng virus của tế bào. Trong một nghiên cứu trên 2.244 bệnh nhân COVID-19 nguy kịch, Pairo và cộng sự ghi nhận sự tương quan đáng kể giữa các trường hợp này với biến thể rs2236757 tại vùng intron của gen *IFNAR2*[12]. Đồng hợp alen C của rs12252 trên gen *IFITM3* đã

được báo cáo là có liên quan đến tình trạng COVID-19 nguy kịch[13].

**Toll-like receptor (TLR):** TLR là một họ các thụ thể nhận dạng đặc hiệu hàng loạt các sản phẩm được tạo ra bởi các vi sinh vật khi xâm nhiễm vào tế bào chủ. Đột biến mất chức năng trong gen TLR7 đã được phát hiện trên các bệnh nhân COVID-19 trẻ tuổi, không có bệnh lý nền nhưng diễn tiến nguy kịch. Khi tiến hành nghiên cứu vai trò của TLR3 và IRF7 trên những trường hợp nhiễm SARS-CoV-2 có nguy cơ tử vong cao, nhóm nghiên cứu tại đại học Rockefeller (Mỹ) phát hiện có khoảng 8 locus có sự liên quan mật thiết đến tiên lượng xấu của những bệnh nhân này[14].

### 3.2. Nhóm gen mã hóa thụ thể

SARS-CoV-2 xâm nhiễm vào tế bào qua trung gian thụ thể, đột biến xảy ra trên các gen mã hóa thụ thể có thể gây ảnh hưởng đến sự nhạy cảm và khả năng tăng sinh của virus, qua đó gây tác động đến các quá trình sinh lý của vật chủ.

**Angiotensin-converting enzyme 2 (ACE2):** là đích đến đầu tiên trong quá trình tiếp cận và bám dính của virus vào tế bào vật chủ. Protein này đóng vai trò khá quan trọng trong cơ chế điều hòa của hệ thống Renin-Angiotensin (RAS) tham gia kiểm soát huyết áp, tái hấp thu natri, phản ứng viêm và xơ hóa trong cơ thể người. Virus sau khi xâm nhiễm vào tế bào sẽ điều hoà giảm biểu hiện protein ACE2, kéo theo sự giảm hoạt động của con đường RAS làm tăng nguy cơ xảy ra các đáp ứng tiền viêm kiểm soát bởi con đường này. Một vài nghiên cứu gần đây cho thấy các biến thể di truyền mã hóa sự biểu hiện vượt mức của ACE2 khá phổ biến trên dân số Đông Á hoặc Nam Á, trong khi biến thể giảm biểu hiện thường được phát hiện ở người châu Phi. Tính đến nay, có ít nhất 8 SNPs tại bề mặt tương tác của ACE2 đã được phát hiện, tuy nhiên tác động của chúng đến sự tương tác giữa SARS-CoV-2 và tế bào vật chủ cần phải được nghiên cứu thêm[15]. Mặt khác, do gen mã hóa ACE2 nằm trên nhiễm sắc thể X, vì vậy sự giảm biểu hiện của protein này ở nam giới được cho là có liên quan đến tỷ lệ tử vong cao hơn.

**Angiotensin-converting enzyme (ACE):** protein tương đồng (homolog protein) với ACE2, cùng tham gia điều hòa con đường RAS. Sự biểu hiện của ACE và ACE2 chịu ảnh hưởng qua lại lẫn nhau thông qua một số con đường tín hiệu như ERK, p38, MAPK. Tuy các nghiên cứu đề cập đến vai trò của ACE trong bệnh lý COVID-19 còn rất ít, Delanghe và cộng sự ghi nhận tần suất alen D của ACE có tương quan nghịch với tỷ lệ nhiễm SARS-CoV-2 trên 25 quốc gia châu Âu[16]. Sự tăng biểu hiện của ACE cũng làm tăng đáng kể nguy cơ tử vong trên bệnh nhân COVID-19 tại châu Á.

**Transmembrane Serine Protease 2 (TMPRSS2):** enzyme thiết yếu cho quá trình phân cắt protein S của SARS-CoV-2, làm rút ngắn khoảng cách giữa virion và màng tế bào, qua đó thúc đẩy sự dung hợp màng giúp virus xâm nhiễm vào tế bào vật chủ dễ dàng hơn. Một số các điểm đa hình như rs12329760, rs8134378, rs2070788, rs9974589 và rs7364083 được dự đoán làm tăng sự biểu hiện của TMPRSS2, khiến tế bào của người bệnh nhạy cảm hơn với sự tấn công của virus[15]. Khi tiến hành giải trình tự toàn bộ exon của TMPRSS2 trên 120 ca bệnh có biểu hiện xơ hóa, tác giả nhận thấy mối liên hệ đáng kể giữa rs75603675, rs61735792, và rs61735794 với sự lây nhiễm của SARS-CoV-2[17]. TMPRSS2 được cho là có biểu hiện cao hơn ở các tế bào phế nang của nam giới, thông tin này phần nào giải thích tần suất viêm phổi xảy ra cao hơn trên nhóm đối tượng này.

**Furin:** protein này đóng vai trò tương tự như TMPRSS2, giúp phân cắt protein S của virus. Tuy nhiên, sự phân cắt bởi Furin diễn ra ở vùng tiếp giáp giữa 2 tiểu đơn vị S1 và S2, trong khi TMPRSS2 thực hiện quá trình này tại vị trí S2'. Một vài các nghiên cứu đề cập đến sự tăng biểu hiện của Furin trên bệnh nhân tiểu đường, giải thích nguy cơ cao hơn khi mắc COVID-19 của những ca bệnh này. Biến thể rs4702 đã được chứng minh làm giảm sự biểu hiện của Furin trên tế bào phế nang, gây ảnh hưởng đáng kể đến sự xâm nhiễm của SARS-CoV-2 [6].

**Cathepsin B/Cathepsin L:** ở các tế bào không biểu hiện TMPRSS2, SARS-CoV-2 sử dụng

cathepsin B hoặc cathepsin L cho quá trình phân cắt và dung hợp của virus. Có khoảng 3 biến thể tại vị trí hoạt động (active site) của cathepsin B và một đột biến sai nghĩa trên cathepsin L được biết là có liên kết với diễn tiến nặng trên bệnh nhân COVID-19[13].

### 3.3. Các gen khác

Ngoài các gen thuộc hệ miễn dịch và các gen mã hóa cho thụ thể tế bào, thông qua các nghiên cứu liên kết toàn hệ gen, một số biến thể gen hoặc các locus gen cũng đã được xác định là có liên quan với sự nhạy cảm và đáp ứng bệnh.

**3p21.31:** là một trong những locus đầu tiên được báo cáo là có tương quan đáng kể đến diễn tiến nặng của bệnh nhân COVID-19 [18]. Vùng locus này bao phủ 6 gen, gồm có: *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, và *XCRI*. Trong đó gen *SLC6A20* được xem là có vai trò chủ đạo, điều hòa yếu tố SIT1 tham gia vào sự tương tác giữa S và ACE2.

**9q34.2 (ABO):** trong cùng phân tích tổng hợp (meta-analysis) trên khoảng 8,5 triệu SNPs từ nhóm bệnh nhân COVID-19 độ nặng ở châu Âu, nhóm nghiên cứu phát hiện biến thể rs657152 tại locus 9q34.2 có ý nghĩa quan trọng trong sự nhạy cảm với virus của người bệnh[15]. Biến thể này gắn liền với vùng gen ABO mã hóa cho nhóm máu ở người. Một số các nghiên cứu tiếp theo trên nhóm lớn bệnh nhân cũng xác định các cá thể nhóm máu A có nguy cơ nhiễm SARS-CoV-2 cao hơn, trong khi người mang nhóm máu O được xem là nhóm đối tượng có nguy cơ thấp nhất. Một SNP khác (rs9411378) của ABO cũng đã được xác định là có tương quan với tiên lượng nặng và kết cục chính của bệnh nhân COVID-19[13]. Tuy cơ chế phân tử vẫn chưa được hiểu rõ, đã có giả thuyết cho rằng các yếu tố nhóm máu đóng vai trò như thụ thể hoặc đồng thụ thể được sử dụng bởi sinh vật ngoại lai trong điều hòa sự hấp thu nội bào, sự truyền tín hiệu hay tương tác với tế bào vật chủ.

**TMEM189-UBE2V1:** là cụm gen đáng chú ý nhất trong đáp ứng miễn dịch qua con đường tín hiệu IL-1. Biến thể trên cụm gen này được phát hiện qua phân tích GWAS trên 22 triệu điểm đa hình từ các bệnh nhân với độ bệnh khác nhau tại Trung Quốc[13]. Điểm đa hình tại *TMEM189-UBE2V1* gắn liền với hội chứng giảm số lượng tế bào lympho và tiểu cầu, tăng CRP, D-dimer, IL-6 trên bệnh nhân COVID-19.

**Oligoadenylate synthase (OAS):** các gen thuộc họ OAS được cảm ứng bởi các IFN ở giai đoạn đầu của quá trình nhiễm virus và chịu trách nhiệm cho sự phân hủy RNA của virus. Cụm gen *OAS1-OAS3* được xác định tham gia bảo vệ tế bào kháng lại sự nhiễm SARS-CoV-2 qua cơ chế miễn dịch bẩm sinh[12]. Do đó, biến thể trên cụm gen này được biết là có ảnh hưởng đến diễn tiến nguy kịch của bệnh nhân.

**Apolipoprotein E (ApoE):** là protein tham gia vào sự chuyển hóa lipid ở tế bào động vật. Các isoform chính của ApoE là APOE2, APOE3, và APOE4 được mã hóa lần lượt bởi các alen  $\epsilon 2 / \epsilon 3 /$

4. Quá trình xâm nhiễm của SARS[CoV]2 được biết là có sự tham gia của cholesterol, chủ yếu ở giai đoạn phân cắt bởi Furin. Một số các nghiên cứu mô tả bệnh nhân có kiểu gen đồng hợp ApoE3 có nguy cơ tử vong do COVID-19 cao gấp 4,3 lần, trong khi bệnh nhân mang kiểu gen dị hợp  $\epsilon 3 / \epsilon 4$  có nguy cơ thấp hơn[15]. Được biết các trường hợp đồng hợp tử tuy chỉ chiếm 13% tổng dân số Hoa Kỳ, có khoảng 34% các ca tử vong do COVID-19 là bệnh nhân mang kiểu gen này.

**PIEZO1:** gen này mã hóa một kênh ion không chọn lọc làm trung gian cho các đáp ứng nội mô trong sự tuần hoàn máu. Các phân tích trình tự toàn bộ gen cho thấy có 3 SNPs của *PIEZO1* có ý nghĩa đáng kể đến tỷ lệ tử vong của bệnh nhân COVID-19, độc lập với các yếu tố nguy cơ [13].

**Dipeptidyl peptidase (DPP):** một số biến thể của các gen thuộc họ DPP cũng đã được miêu tả là có liên quan đến quá trình đáp ứng miễn dịch của bệnh nhân nhiễm SARS[CoV]2. Trong đó nổi

bật là DPP4 (hay còn được gọi là CD26), là một glycoprotein xuyên màng tham gia vào sự hoạt hóa tế bào T, giúp tế bào vật chủ chống lại sự nhiễm virus. Biến thể tại vùng 5'UTR của DDP4 (rs13015258) đã được báo cáo là có ảnh hưởng đến sự nhập bào của vi rút [15]. Đột biến mất chức năng của DPP7 được cho là một trong nguyên nhân dẫn đến diễn tiến bệnh khác nhau trên hai chị em ruột trong cùng một gia đình. Trong khi đó, mối liên hệ giữa biến thể của DPP9 và độ nặng của bệnh nhân mắc COVID-19 cũng đã được phát hiện qua nghiên cứu GWAS.

#### 4. CẬP NHẬT TÌNH HÌNH NGHIÊN CỨU TRONG NƯỚC

Ngày càng có nhiều các dữ liệu cho thấy kiểu gen của vật chủ có vai trò quan trọng trong khuynh hướng và mức độ đáp ứng với sự lây nhiễm. Ví dụ, hội chứng mất khứu giác hoặc vị giác do COVID-19 thường được báo cáo ở dân số châu Âu và Hoa Kỳ, nhưng ít gặp hơn trên người châu Á. Khi so sánh với các ca đơn lẻ cùng cấp độ bệnh, Ikitimur và cộng sự quan sát thấy nhóm bệnh nhân COVID-19 thuộc cùng gia đình có diễn tiến bệnh tương tự nhau, qua đó củng cố thêm vai trò của di truyền trong đáp ứng bệnh. Tuy nhiên, cần lưu ý là phần lớn các thông tin về biến thể di truyền có liên hệ với COVID-19 được thu nhận từ dân số châu Âu hoặc Hoa Kỳ, do đó có thể không mang tính đặc trưng cho người Việt Nam. Vì vậy, việc triển khai các nghiên cứu về tương quan giữa di truyền vật chủ và diễn tiến bệnh COVID-19 trên dân số Việt Nam cũng là một trong những hướng đi đang được các nhà khoa học trong nước quan tâm.

Có thể kể đến một số các kết quả ban đầu bao gồm công trình của Tôn và cộng sự đề cập đến tính đa hình của thụ thể *TMPRSS2* trên bệnh nhân COVID-19 Việt Nam. Trong đó, khi giải trình tự toàn bộ exon trên 270 bệnh nhân, nhóm nghiên cứu ghi nhận 29 biến thể tại vùng không mã hóa và 14 biến thể tại vùng mã hóa của *TMPRSS2* với rs12329760 là biến thể có tần suất xuất hiện cao nhất [19]. Nghiên cứu về mối liên quan giữa bộ gen và nhóm máu của vật chủ đến tính nhạy cảm, độ bệnh nặng COVID-19; từ phân tích toàn bộ trình tự exon của 200 bệnh nhân và 100 ca đối chứng, nhóm của Vũ Phương Nhung báo cáo nhóm máu O là yếu tố bảo vệ giúp cải thiện kết cục của người bệnh [20]. Về sự nhạy cảm với vi rút, hai biến thể rs2229207 và rs17860118 thuộc gen *IFNAR2* làm tăng nguy cơ lây nhiễm, trong khi rs139940581 của *SLC6A20* làm giảm khuynh hướng này. Trong sự tiến triển bệnh, nhóm ghi nhận tần suất cao của rs4622692 và rs1048610 trên bệnh nhân có độ bệnh trung bình, đồng thời rs12329760 có tương quan đáng kể với kiểu hình bệnh nhẹ hoặc không biểu hiện triệu chứng. Đáng chú ý, rs2304255 gen *TYK2* và rs2277735 của *DPP9* được báo cáo là có liên hệ đến nguy cơ bệnh nặng hoặc tử vong [20].

#### 5. KẾT LUẬN

Nhìn chung, đã có nhiều công trình được thực hiện nhằm tìm kiếm các gen hoặc nhóm gen đóng vai trò quan trọng trong quá trình xâm nhiễm và tăng sinh của virus, từ đó giúp đưa ra các thông tin cần thiết cho việc phát triển các liệu pháp ngăn chặn sự lây nhiễm cũng như điều trị bệnh COVID-19. Đồng thời, các nghiên cứu này cũng tập trung xác định biến thể gen có tác động đến đáp ứng miễn dịch của vật chủ - hay các “dấu ấn” di truyền - phục vụ cho sàng lọc và điều trị sớm những bệnh nhân có nguy cơ cao. Đây được xem là bước đi thiết yếu trong phòng - chống dịch COVID-19 ở giai đoạn hậu vắc-xin. Tuy nhiên, cần nhấn mạnh là tần suất của các biến thể đã được báo cáo có sự thay đổi phụ thuộc từng chủng tộc, cho thấy các nghiên cứu khảo sát biến thể di truyền trên nhiều nhóm dân số là rất cần thiết, qua đó thẩm định và xác nhận dấu ấn di truyền đặc trưng có liên quan đến nhạy cảm và tiên lượng nặng đặc trưng cho từng chủng tộc.

#### TÀI LIỆU THAM KHẢO

- [1] D. Singh and S. V. Yi, 'On the origin and evolution of SARS-CoV-2', *Exp. Mol. Med.*, vol. 53, no. 4, Art. no. 4, Apr. 2021, doi: 10.1038/s12276-021-00604-z.

- [2] E. C. Holmes *et al.*, ‘The origins of SARS-CoV-2: A critical review’, *Cell*, vol. 184, no. 19, pp. 4848–4856, Sep. 2021, doi: 10.1016/j.cell.2021.08.017.
- [3] N. F. Müller, K. E. Kistler, and T. Bedford, ‘Recombination patterns in coronaviruses’, *BioRxiv Prepr. Serv. Biol.*, p. 2021.04.28.441806, Feb. 2022, doi: 10.1101/2021.04.28.441806.
- [4] Y. Turakhia *et al.*, ‘Pandemic-scale phylogenomics reveals the SARS-CoV-2 recombination landscape’, *Nature*, vol. 609, no. 7929, Art. no. 7929, Sep. 2022, doi: 10.1038/s41586-022-05189-9.
- [5] M. F. Boni *et al.*, ‘Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic’, *Nat. Microbiol.*, vol. 5, no. 11, Art. no. 11, Nov. 2020, doi: 10.1038/s41564-020-0771-4.
- [6] M. R. Gitman, M. V. Shaban, A. E. Paniz-Mondolfi, and E. M. Sordillo, ‘Laboratory Diagnosis of SARS-CoV-2 Pneumonia’, *Diagnostics*, vol. 11, no. 7, Art. no. 7, Jul. 2021, doi: 10.3390/diagnostics11071270.
- [7] D. A. Jamison *et al.*, ‘A comprehensive SARS-CoV-2 and COVID-19 review, Part 1: Intracellular overdrive for SARS-CoV-2 infection’, *Eur. J. Hum. Genet.*, vol. 30, no. 8, Art. no. 8, Aug. 2022, doi: 10.1038/s41431-022-01108-8.
- [8] C. B. Jackson, M. Farzan, B. Chen, and H. Choe, ‘Mechanisms of SARS-CoV-2 entry into cells’, *Nat. Rev. Mol. Cell Biol.*, vol. 23, no. 1, Art. no. 1, Jan. 2022, doi: 10.1038/s41580-021-00418-x.
- [9] P. V’kovski, A. Kratzel, S. Steiner, H. Stalder, and V. Thiel, ‘Coronavirus biology and replication: implications for SARS-CoV-2’, *Nat. Rev. Microbiol.*, vol. 19, no. 3, pp. 155–170, Mar. 2021, doi: 10.1038/s41579-020-00468-6.
- [10] F. Wang *et al.*, ‘Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility’, *Cell Discov.*, vol. 6, no. 1, pp. 1–16, Nov. 2020, doi: 10.1038/s41421-020-00231-4.
- [11] C. Strafella *et al.*, ‘Investigation of Genetic Variations of IL6 and IL6R as Potential Prognostic and Pharmacogenetics Biomarkers: Implications for COVID-19 and Neuroinflammatory Disorders’, *Life Basel Switz.*, vol. 10, no. 12, p. E351, Dec. 2020, doi: 10.3390/life10120351.
- [12] E. Pairo-Castineira *et al.*, ‘Genetic mechanisms of critical illness in COVID-19’, *Nature*, vol. 591, no. 7848, pp. 92–98, Mar. 2021, doi: 10.1038/s41586-020-03065-y.
- [13] S. V. Kotsev *et al.*, ‘Hypotheses and facts for genetic factors related to severe COVID-19’, *World J. Virol.*, vol. 10, no. 4, pp. 137–155, Jul. 2021, doi: 10.5501/wjv.v10.i4.137.
- [14] Q. Zhang *et al.*, ‘Inborn errors of type I IFN immunity in patients with life-threatening COVID-19’, *Science*, vol. 370, no. 6515, p. eabd4570, Oct. 2020, doi: 10.1126/science.abd4570.
- [15] Z. Yildirim, O. S. Sahin, S. Yazar, and V. Bozok Cetintas, ‘Genetic and epigenetic factors associated with increased severity of Covid-19’, *Cell Biol. Int.*, vol. 45, no. 6, pp. 1158–1174, 2021, doi: 10.1002/cbin.11572.
- [16] J. R. Delanghe, M. M. Speeckaert, and M. L. De Buyzere, ‘The host’s angiotensin-converting enzyme polymorphism may explain epidemiological findings in COVID-19 infections’, *Clin. Chim. Acta Int. J. Clin. Chem.*, vol. 505, pp. 192–193, Jun. 2020, doi: 10.1016/j.cca.2020.03.031.
- [17] L. Torre-Fuentes *et al.*, ‘ACE2, TMPRSS2, and Furin variants and SARS-CoV-2 infection in Madrid, Spain’, *J. Med. Virol.*, vol. 93, no. 2, pp. 863–869, Feb. 2021, doi: 10.1002/jmv.26319.
- [18] The Severe Covid-19 GWAS Group, ‘Genomewide Association Study of Severe Covid-19 with Respiratory Failure’, *N. Engl. J. Med.*, vol. 383, no. 16, pp. 1522–1534, Oct. 2020, doi: 10.1056/NEJMoa2020283.
- [19] N. D. Ton *et al.*, ‘Polymorphism of the TMPRSS2 gene relating to COVID-19 subceptibility in Vietnamese population’, *Acad. J. Biol.*, vol. 43, no. 1, Art. no. 1, Mar. 2021, doi: 10.15625/2615-9023/15829.
- [20] V. P. Nhung *et al.*, ‘Host Genetic Risk Factors Associated with COVID-19 Susceptibility and Severity in Vietnamese’, *Genes*, vol. 13, no. 10, Art. no. 10, Oct. 2022, doi: 10.3390/genes13101884.