



Original Article

Identifying Putative Drug Target in Methicillin Resistant *Staphylococcus aureus* Using *in silico* Proteome Subtraction Analysis

Le Anh Vu^{1,*}, Phan Thi Cam Quyen², Nguyen Thuy Huong¹

¹Faculty of Chemistry Engineering, Ho Chi Minh City University of Technology,
268 Ly Thuong Kiet, 10 District, Ho Chi Minh City, Vietnam

²Department of Biotechnology, KienGiang Seed Research Center,
731 Nguyen Trung Truc, An Haa, Rach Gia, Kien Giang, Vietnam

Received 07 December 2018

Revised 24 December 2018; Accepted 27 December 2018

Abstract: This paper analyzes the proteomes of 15 different MRSA strains using an *in silico* subtraction approach to identify putative drug targets. Various bioinformatics tools were used to screen for paralogous sequences and homologous proteins against the host (*Homo sapiens*) from the bacterial proteome. The remaining proteins were further analyzed to identify essential proteins using the Database of Essential Genes (DEG). The results show that 2,235 proteins in MRSA were not homologous with the host proteome, and 158 of these proteins were identified as the essential proteins for the viability of *S. aureus* according to DEG. Moreover, metabolic pathway analysis of essential proteins with the Kyoto Encyclopedia of Genes and Genomes (KEGG) verified that 49 proteins participated in 11 unique metabolic pathways in MRSA. The identified proteins are expected to have great potential in drug design against MRSA.

Keywords: Drug targets, essential proteins, human non-homologous, MRSA, proteome subtraction.

* Corresponding author.

Email address: lavu68@gmail.com

<https://doi.org/10.25073/2588-1140/vnunst.4813>



Xác định mục tiêu thuốc giả định ở vi khuẩn *Staphylococcus aureus* kháng methicillin bằng phương pháp phân tích loại trừ trong hệ protein *in silico*

Lê Anh Vũ¹, Phan Thị Cẩm Quyên², Nguyễn Thúy Hương¹

¹Khoa Kỹ thuật Hóa Học, Đại học Bách Khoa TP. HCM, 268 Lý Thường Kiệt, Quận 10, TP.HCM, Việt Nam

²Phòng Công nghệ Sinh học, Trung tâm Giống Kiên Giang,

731 Nguyễn Trung Trực, An Hòa, Rạch Giá, Kiên Giang, Việt Nam

Nhận ngày 07 tháng 12 năm 2018

Chỉnh sửa ngày 24 tháng 12 năm 2018; Chấp nhận đăng ngày 27 tháng 12 năm 2018

Tóm tắt: Trong nghiên cứu này, các hệ protein của 15 chủng MRSA đã được phân tích loại trừ theo hướng tiếp cận *in silico* để xác định các mục tiêu thuốc giả định. Các công cụ tin sinh học khác nhau đã được sử dụng để sàng lọc các chuỗi protein paralog và các chuỗi tương đồng với vật chủ (*Homo sapiens*) từ hệ protein của vi khuẩn. Các protein còn lại tiếp tục được phân tích để xác định các protein thiết yếu bằng cách sử dụng cơ sở dữ liệu Database of Essential Genes (DEG). Kết quả cho thấy có 2235 protein ở MRSA không tương đồng với hệ protein vật chủ, và trong đó 158 protein được xác định là cần thiết cho sự tồn tại của *S. aureus* theo cơ sở dữ liệu DEG. Ngoài ra, phân tích trao đổi chất của các protein thiết yếu với cơ sở dữ liệu Kyoto Encyclopedia of Genes and Genomes (KEGG) đã phân loại được 49 protein thiết yếu tham gia vào 11 con đường trao đổi chất đặc trưng ở MRSA. Các protein được xác định dự kiến sẽ có tiềm năng lớn trong ứng dụng làm mục tiêu thuốc ở MRSA.

Từ khóa: MRSA, mục tiêu thuốc, phân tích loại trừ trong hệ protein, protein không tương đồng, protein thiết yếu.

1. Giới thiệu

Staphylococcus aureus (*S. aureus*) là một trong những tác nhân gây bệnh phổ biến nhất

trong môi trường bệnh viện và cộng đồng, với nhiều biểu hiện lâm sàng khác nhau, từ nhiễm khuẩn bề mặt đến các biến chứng đe dọa tính mạng. Các vị trí phổ biến nhất bị ảnh hưởng là da và mô mềm; biểu hiện của nhiễm khuẩn ở các vị trí này bao gồm viêm nang lông, mụn nhọt, bệnh chốc lở, nhiễm khuẩn vết thương và hội chứng da do tụ cầu. Các nhiễm khuẩn nghiêm trọng hơn bao gồm nhiễm khuẩn huyết,

*Tác giả liên hệ.

Địa chỉ email: lavu68@gmail.com

<https://doi.org/10.25073/2588-1140/vnunst.4813>

viêm phổi, viêm nội tâm mạc, nhiễm khuẩn xương khớp và hội chứng sốc độc tố. *S. aureus* cũng là nguyên nhân gây ra ngộ độc thực phẩm [1]. Việc điều trị nhiễm khuẩn do *S. aureus* bao gồm sử dụng các kháng sinh nhóm β -lactam, từ Benzylpenicillin thế hệ thứ nhất đến các Penicillin hiện đại (Methicillin, Oxacillin). Tuy nhiên, sau khi Methicillin được ứng dụng trong lâm sàng, các chủng *S. aureus* kháng Methicillin (MRSA) đã được phân lập [2]. Ngoài ra, việc sử dụng không kiểm soát nhiều loại kháng sinh trong những năm qua đã góp phần vào sự xuất hiện của các chủng MRSA đa kháng thuốc. Các chủng này là kết quả của những đột biến gen mã hóa cho các protein đích và thông qua việc tiếp nhận và tích lũy các gen kháng kháng sinh [3]. Các báo cáo gần đây về chủng *S. aureus* có tính kháng trung gian hoặc hoàn toàn đối với kháng sinh glycopeptide Vancomycin đã báo hiệu thời kỳ hóa trị liệu trong đó kháng sinh diệt khuẩn chống lại vi khuẩn này có thể không còn hữu hiệu [2]. Do đó, việc tìm kiếm các hợp chất chống tụ cầu mới là một vấn đề cấp bách. Một hướng tiếp cận để hạn chế vấn đề kháng thuốc là tìm kiếm các mục tiêu tác động mới, trong khi hướng khác là phát triển các tác nhân có thể làm giảm sự đề kháng của vi khuẩn với các kháng sinh hiện có [3].

Khám phá các ứng viên thuốc mới không phải là một nhiệm vụ đơn giản và đó là quá trình đòi hỏi thời gian và chi phí tốn kém. Vấn đề này đặt ra thách thức cho các nhà nghiên cứu là cần đưa ra các phương pháp thay thế nhanh chóng và tiết kiệm hơn để có thể đề xuất các ứng viên thuốc mới. Một phương pháp hữu ích để khám phá các ứng viên thuốc này là xác định các mục tiêu thuốc duy nhất ở tác nhân gây bệnh, đặc biệt là đối với những mục tiêu quan trọng. Thông tin về tính thiết yếu của gen và các sản phẩm của nó được coi là các mục tiêu tiềm năng cho việc phát triển các loại thuốc mới [4]. Những tiến bộ gần đây trong lĩnh vực sinh học tính toán và tin sinh học đã tạo ra sự khác biệt trong phân tích *in silico* và phương pháp thiết kế thuốc, giúp giảm thời gian và chi phí liên quan đến thử nghiệm và hạn chế thử nghiệm sai lệch trong quá trình phát triển thuốc

[5]. Những phương pháp này nhằm sàng lọc các mục tiêu thuốc tiềm năng, sau đó sử dụng chúng để thử nghiệm trong phòng thí nghiệm. Bên cạnh đó, các nghiên cứu *in silico* docking giữa các mục tiêu thuốc được xác định và các loại thuốc hiện có với sự thay đổi nhỏ trong cấu trúc có thể dẫn đến việc phát hiện ra các loại thuốc mới có hiệu quả điều trị. Kết quả là, một loạt các mục tiêu thuốc và các hợp chất khởi nguồn có thể được xác định trước khi thử nghiệm trong phòng thí nghiệm. Phân tích loại trừ trong hệ protein là một trong những cách tiếp cận *in silico* được sử dụng để xác định mục tiêu thuốc dựa trên việc xác định các protein thiết yếu và không tương đồng với vật chủ ở sinh vật gây bệnh [5]. Hơn nữa, các protein tương đồng với vật chủ được sàng lọc còn có thể giúp tránh các phản ứng có hại của thuốc trong các giai đoạn phát triển thuốc tiếp theo. Bằng cách lựa chọn các protein thiết yếu đặc trưng cần cho sự tồn tại và phát triển của mầm bệnh, phương pháp loại trừ protein cho phép xác định các mục tiêu thuốc tiềm năng ở mầm bệnh [5]. Các tiến bộ gần đây trong lĩnh vực genomics, transcriptomics và metabolomics đã cho phép dự đoán cấu trúc, chức năng của các mục tiêu thuốc với độ chính xác cao. Kết hợp thành tựu của các ngành khoa học trên, phân tích loại trừ protein là hướng tiếp cận mới trong xác định các mục tiêu thuốc, bằng cách sử dụng các công cụ tin học và cơ sở dữ liệu (CSDL) sinh học, và đã góp phần tăng tốc quá trình khám phá thuốc hiện đại [4].

Trong nghiên cứu này, phương pháp phân tích loại trừ protein đã được áp dụng để xác định các protein thiết yếu cần cho sự tồn tại của MRSA và không tương đồng với hệ protein vật chủ (*Homo sapiens*). Những protein này có thể được sử dụng như là mục tiêu thuốc tiềm năng cho các giai đoạn phát triển thuốc trong tương lai. Đã có nhiều nghiên cứu cho thấy khả năng áp dụng các mục tiêu dự đoán thành công khi sử dụng phương pháp này [4]. Ngoài ra, việc tìm kiếm các ứng viên thuốc mới chống lại các mục tiêu điều trị như trên còn giúp giảm thiểu khả năng mầm bệnh chống lại các phương pháp điều trị hiện có và do đó, có thể làm giảm tác dụng phụ đối với vật chủ. Mặc dù chiến lược

tương tự với MRSA đã được báo cáo trong một số nghiên cứu [4, 6-7] nhưng ở đây chúng tôi áp dụng quy trình phân tích trên 15 chủng MRSA cùng lúc. Việc phân tích như vậy cho phép xác định các protein tiềm năng để có thể sử dụng làm mục tiêu cho sự phát triển của các tác nhân trị liệu trên một loạt các chủng MRSA.

2. Phương tiện và phương pháp

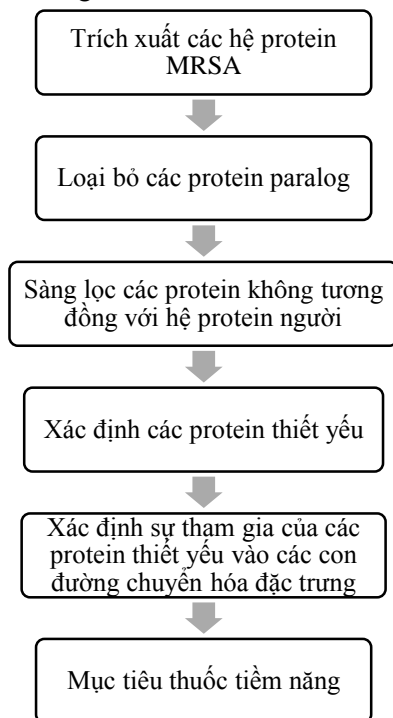
2.1. Phương tiện

Tất cả các CSDL, công cụ và chương trình sử dụng trong nghiên cứu này đều có sẵn trên Internet và được truy cập/tải về theo như mô tả trong mục phương pháp.

2.2. Phương pháp nghiên cứu

Truy xuất hệ protein của MRSA

Các trình tự protein hoàn chỉnh (hệ protein) của MRSA được trích xuất từ CSDL Universal Protein Resource (<https://www.uniprot.org>) và được lưu dưới định dạng FASTA [8]. Các bước phân tích trong nghiên cứu hiện tại được thể hiện như trong Hình 1.



Hình 1. Sơ đồ xác định các mục tiêu thuốc tiềm năng ở MRSA.

Loại bỏ các protein paralog

Các hệ protein MRSA đã tải xuống được phân tích bởi USEARCH

(<https://www.drive5.com/usearch>) bằng cách sử dụng các dòng lệnh tuân tự [9]. Đầu tiên, lệnh `-cluster_fast` được sử dụng để tìm/phân cụm các protein paralog với “sequence identity cut off value” ở mức 0,6 (tức 60% mức độ giống nhau). Protein có hơn 60% mức độ giống nhau được xem là protein paralog [10] và được loại bỏ. Sau đó, các nhóm protein còn lại được sàng lọc thêm để xác định các chuỗi trùng lặp (còn được gọi là quá trình dereplication) sử dụng lệnh `-fastx_uniques`. Các bản sao được loại bỏ và lệnh `-sortbylength` được sử dụng để sàng lọc các protein có hơn 100 acid amin. Protein có ít hơn 100 acid amin trong chuỗi trình tự hầu như không thể đại diện cho các protein thiết yếu [10].

Xác định các protein không tương đồng

Các protein thu được sau khi phân tích với USEARCH được so sánh với hệ protein *H. sapiens* sử dụng chương trình BLASTp (<https://blast.ncbi.nlm.nih.gov>) với giá trị ngưỡng kỳ vọng (E-value) 10^{-4} . Refseq được chọn làm CSDL để tìm ra các protein MRSA không tương đồng với protein của người. Các trình tự cho thấy sự tương đồng đáng kể với protein người được loại bỏ và các chuỗi không tương đồng được giữ lại cho những phân tích tiếp theo [10].

Xác định các protein thiết yếu ở MRSA

Các protein không tương đồng với protein của *H. sapiens* được tiếp tục phân tích sử dụng BLASTp dựa vào CSDL Database of Essential Genes (DEG) (<http://www.essentialgene.org>), trong đó các protein không thể thiếu cho sự tồn tại của *S. aureus* được chọn làm CSDL tham chiếu. Các giá trị truy vấn “E value cut off score less than 10^{-10} ” và “minimum bit score cut-off of 100” được sử dụng để thu được một tập hợp các protein thiết yếu. Do đó, các chuỗi protein thu được không tương đồng với hệ protein của *H. sapiens* và không thể thiếu cho các hoạt động của mầm bệnh [5].

Phân tích các con đường trao đổi chất và xác định mục tiêu thuốc tiềm năng

CSDL KEGG (Kyoto Encyclopedia of Genes and Genomes)

(<https://www.genome.jp/kegg>) được sử dụng để so sánh các con đường trao đổi chất giữa MRSA và *H. sapiens*, trong đó các con đường chỉ có ở MRSA được lọc ra. Sau đó, các protein thiết yếu của MRSA xác định ở các bước trên được đưa vào phân tích trao đổi chất bằng cách sử dụng CSDL KAAS (KEGG Automatic Annotation Server) để xác định các protein thiết yếu tham gia vào những con đường chuyển hóa đặc trưng cho MRSA. Để lựa chọn các protein trao đổi chất từ các protein truy vấn được cung cấp, KAAS sử dụng phân tích BLASTp đối với hệ protein của các sinh vật nhân sơ [10]. Kết quả thu được có chứa các mã gán KO (KEGG Orthology) và các đường dẫn KEGG được tạo tự động [5].

3. Kết quả và thảo luận

Nghiên cứu này được thực hiện nhằm xác định các mục tiêu thuốc tiềm năng ở MRSA bằng cách sử dụng phương pháp tiếp cận loại trừ trong hệ protein giúp phân loại và xác định các protein mục tiêu theo các tiêu chí đã lựa chọn. Bảng 1 mô tả số lượng các protein đầu ra được sàng lọc ở mỗi bước.

Hệ protein của MRSA

Hệ protein hoàn chỉnh (tổng cộng 40304 protein) của MRSA được tải về từ CSDL UniProt bao gồm các chủng và phân loài khác nhau (Bảng 2).

Các protein non-paralog

Để loại bỏ các protein paralog, các thuật toán USEARCH [9] đã được sử dụng và kết quả cho thấy một tập hợp tổng số 3815 cụm protein duy nhất. Mỗi đại diện của các cụm được đánh dấu và từ phần còn lại, 972 protein với độ dài chuỗi nhỏ hơn 100 acid amin đã được loại bỏ. 2843 protein non-paralog còn lại sau đó được phân tích BLASTp đối với hệ protein của người (*H. sapiens*).

Các protein không tương đồng với hệ protein người

Từ kết quả NCBI BLASTp tổng cộng 630 protein tương đồng với hệ protein người được

tìm thấy với giá trị ngưỡng kỳ vọng E-value là 10^{-4} . Các protein tương đồng giữa người và vi khuẩn diễn tiến trong quá trình tiến hóa và tham gia vào nhiều hệ thống sinh học phổ biến giữa vật chủ và tác nhân gây bệnh [10]. Các protein tương đồng với hệ protein người được loại trừ khỏi các phân tích tiếp theo vì chúng có thể cho thấy phản ứng bất lợi đối với tế bào người nếu được chọn làm mục tiêu thuốc chống lại MRSA.

Các protein thiết yếu ở MRSA

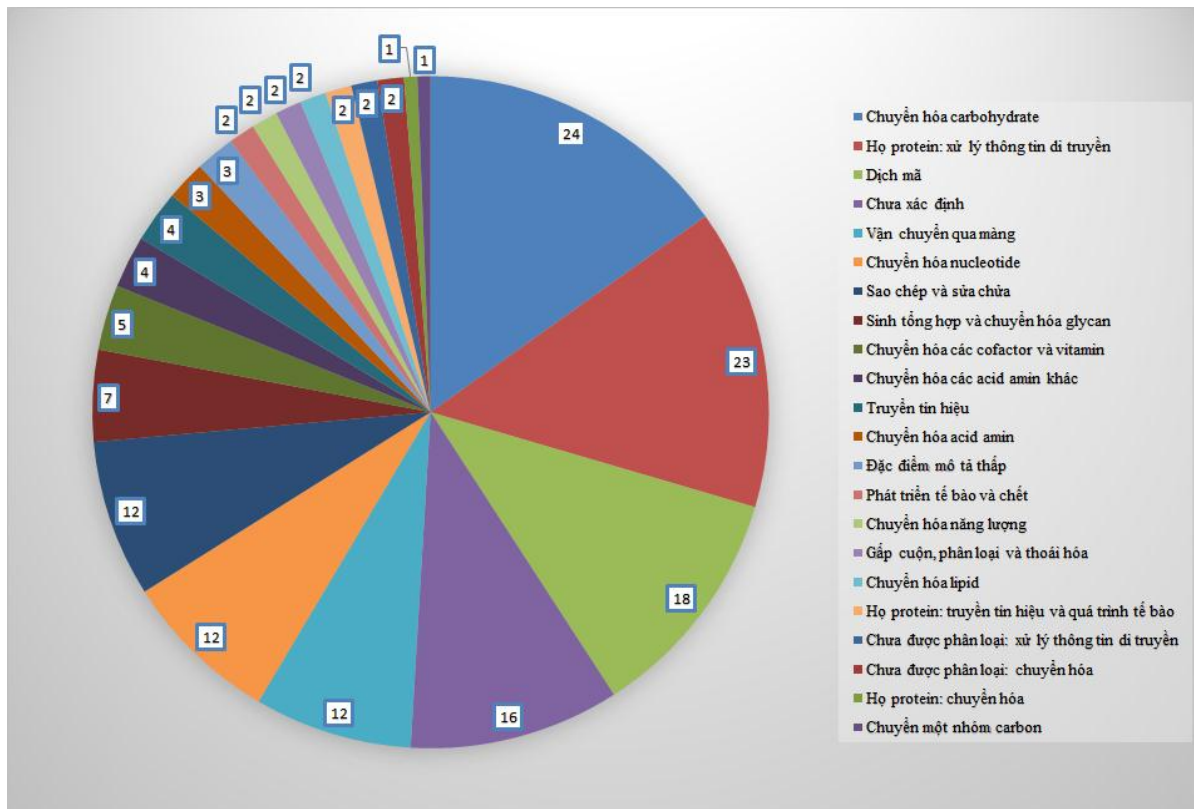
Protein thiết yếu là những protein tối thiểu cần thiết cho sự tồn tại của mầm bệnh trong điều kiện thuận lợi và đóng vai trò quan trọng trong các quá trình tế bào của chúng. Vì vậy, chúng thường là mục tiêu phù hợp và đầy hứa hẹn cho việc thiết kế thuốc [10]. Bằng phân tích NCBI BLASTp, 2213 protein có ở MRSA đã được tìm thấy không tương đồng với hệ protein người. Các protein này sau đó được phân tích với CSDL DEG. Quá trình sàng lọc thông qua DEG với các giá trị truy vấn đã chọn cho thấy có 158 protein không tương đồng với hệ protein người là các protein thiết yếu (Hình 2; Phụ lục 1).

Bảng 1. Kết quả phân tích hệ protein của vi khuẩn MRSA

Các bước phân tích	Số lượng protein tìm được
Protein từ CSDL UniProt	40304
Protein paralog (>60% mức độ giống nhau) tìm thấy bởi USEARCH	26171
Protein non-paralog	14133
Protein non-paralog/không lặp lại tìm thấy bởi USEARCH	3815
Protein non-paralog/không lặp lại có độ dài chuỗi lớn hơn 100 amino acid	2843
Protein không tương đồng với các protein người (E-value 10^{-4})	2213
Protein không tương đồng/Protein thiết yếu tìm thấy trên DEG (E-value 10^{-10})	158
Các con đường chuyển hóa đặc trưng được tìm thấy trên máy chủ KAAS	27
Protein thiết yếu trong các con đường chuyển hóa đặc trưng tìm được trên máy chủ KEGG	49

Bảng 2. Các chủng MRSA được sử dụng trong nghiên cứu và số lượng protein tương ứng

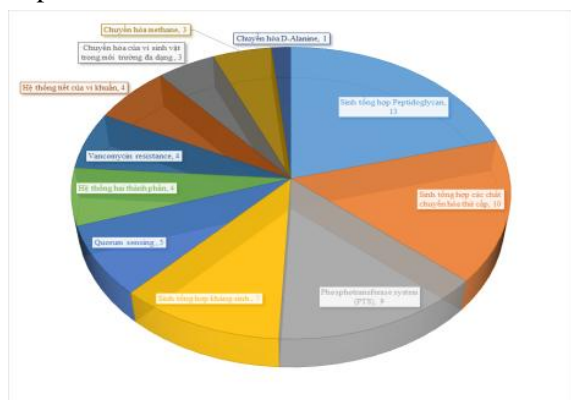
Proteome ID	Tên chủng	ID Chủng	Số lượng protein
UP000024517	Staphylococcus aureus MRSA-136	1413497	2688
UP000000596	Staphylococcus aureus (strain MRSA252)	282458	2635
UP000051781	Staphylococcus aureus MRSA_CVMN25720PS	1387326	2546
UP000003093	Staphylococcus aureus subsp. aureus DR10	1155079	2684
UP000031632	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> ST772-MRSA-V (Strain: DAR4145)	1343064	2630
UP000025090	Staphylococcus aureus MRSA-118	1413495	2661
UP000003572	Staphylococcus aureus subsp. aureus MRSA177	754026	2773
UP000005003	Staphylococcus aureus subsp. aureus 71193	1155084	2620
UP000008696	<i>Staphylococcus aureus</i> (strain MRSA ST398 / isolate S0385)	523796	2648
UP000003548	Staphylococcus aureus subsp. aureus MRSA131	754025	2770
UP000050970	Staphylococcus aureus MRSA_CVMN27231PS	1387328	2743
UP000050900	Staphylococcus aureus MRSA_CVMN26035PS	1387327	2719
UP000054253	Staphylococcus aureus subsp. aureus MRSA_S4	1337995	2886
UP000186671	<i>Staphylococcus aureus</i> (strain MRSA ST398 / isolate S0385) (Strain: 55488)	523796	2650
UP000031939	Staphylococcus aureus MRSA_CVM43477	1448137	2651



Hình 2. Các protein thiết yếu và sự tham gia vào các quá trình tế bào của chúng.

Các con đường chuyển hóa đặc trưng của MRSA

Phân tích con đường chuyển hóa có một ý nghĩa quan trọng trong xác định mục tiêu thuốc tiềm năng vì nó có thể là chu trình cần thiết cho sự tồn tại và khả năng gây bệnh của vi khuẩn [10]. Sau khi so sánh các con đường chuyển hóa giữa MRSA và *H. sapiens* tại máy chủ KEGG, 27 con đường chuyển hóa đặc trưng của MRSA đã được xác định (Phụ lục 2). Trong tổng số 158 protein thiết yếu được kiểm tra tại máy chủ KEGG, kết quả cho thấy có 49 protein tham gia vào 11 con đường chuyển hóa đặc trưng ở MRSA bao gồm sinh tổng hợp peptidoglycan, sinh tổng hợp các chất chuyển hóa thứ cấp, hệ thống phosphotransferase, sinh tổng hợp kháng sinh, quorum sensing, hệ thống hai thành phần, đề kháng Vancomycin, hệ thống tiết của vi khuẩn, chuyển hóa của vi khuẩn trong môi trường đa dạng, chuyển hóa methane và sinh tổng hợp D-alanine (Hình 3; Phụ lục 3). Vai trò của một số con đường đặc trưng và các protein tham gia vào các con đường này được thảo luận chi tiết trong phần tiếp theo.



Hình 3. Số lượng các mục tiêu thuốc tiềm năng và các con đường chuyển hóa đặc trưng tương ứng (một vài protein có thể tham gia vào nhiều hơn một con đường).

Sinh tổng hợp peptidoglycan

Peptidoglycan là thành phần chính cấu tạo nên thành tế bào của hầu hết các vi khuẩn [12]. Nó có vai trò cấu trúc và hoạt động như một màng chọn lọc các phân tử từ môi trường bên ngoài. Do đó, tổng hợp peptidoglycan là một trong những con đường sinh học quan trọng

nhất ở vi khuẩn và đã được nghiên cứu nhiều trong những năm qua [12-14]. Con đường tổng hợp bắt đầu trong tế bào chất, tiếp tục ở màng tế bào chất và kết thúc tại chu chất, nơi các tiền chất được polyme hóa thành lớp peptidoglycan. Một số protein liên quan đến con đường này, ví dụ như các enzym Mur và các protein gắn penicillin (PBPs) đã được nghiên cứu và được xem là các mục tiêu hợp lý cho việc thiết kế các hợp chất kháng khuẩn [13, 14]. Vì vậy, nghiên cứu các đặc điểm sinh học của peptidoglycan có thể là một phần giải pháp đối với vấn đề vi khuẩn kháng kháng sinh, cũng như cải thiện sự hiểu biết về sinh lý vi khuẩn [13]. Từ kết quả KAAS, 13 enzym tham gia sinh tổng hợp peptidoglycan đã được xác định, trong đó murA, murC, murD, murE, murF, murG, femA, femB, femX, ddl, bacA và mraY không có sự tương đồng với hệ protein người và đồng thời rất cần thiết cho sự tồn tại của MRSA. Việc ức chế các protein này do đó có thể được coi là một chiến lược tốt để chống lại MRSA.

Chuyển hóa D-alanine

D-Ala-D-Ala là dipeptide điểm cuối của tiền chất peptidoglycan, UDP-MurNac-pentapeptide và là thành phần cốt yếu trong liên kết ngang của peptidoglycan. Nó được tổng hợp trong tế bào chất của vi khuẩn bởi enzyme D-alanine: D-alanine ligase (Ddl) [15]. Enzyme Ddl xúc tác sự tổng hợp các dipeptide D-Ala-D-Ala và D-Ala-D-Ser hoặc D-Ala-D-Lac depsipeptide ở bước đầu của quá trình tổng hợp peptidoglycan. Chức năng của enzyme này cần thiết cho sự phát triển của vi khuẩn và mang tính đặc hiệu, khiến chúng trở thành mục tiêu hợp lý cho sự phát triển của các kháng sinh mới [16]. Do đó, việc phát triển các hợp chất ức chế D-alanine-D-alanine ligase và các enzym liên quan sẽ đem lại tiềm năng trong việc ức chế sự phát triển không chỉ của MRSA mà còn của các vi khuẩn gây bệnh có thành tế bào.

Hệ thống phosphotransferase

Hệ thống phosphotransferase (PTS) của vi khuẩn thực hiện cả chức năng xúc tác và điều hòa. Nó xúc tác sự vận chuyển và phosphoryl hóa của nhiều loại đường và dẫn xuất, đồng thời cũng thực hiện nhiều chức năng điều hòa liên

quan đến chuyển hóa carbon, nitơ, phosphate và yếu tố độc lực của một số vi khuẩn gây bệnh [17]. Các cơ chế điều hòa qua trung gian PTS dựa trên sự phosphoryl hóa trực tiếp của protein mục tiêu hoặc các tương tác phụ thuộc vào sự phosphoryl hóa dẫn đến sự kích hoạt hoặc ức chế chức năng của các protein mục tiêu [18]. Những đặc tính này giải thích tại sao PTS kiểm soát rất nhiều chức năng tế bào và phát triển rất nhiều cơ chế điều hòa khác nhau, mỗi cơ chế có thể tối ưu để thích nghi với nhu cầu của từng loài vi khuẩn [17]. Do đó, PTS đã trở thành một nguồn đầy tiềm năng cho việc tìm kiếm các mục tiêu thuốc. Nghiên cứu của chúng tôi cho thấy rằng có 9 enzyme tham gia vào hệ thống PTS có thể được sử dụng như những mục tiêu thuốc mới ở MRSA là glucose-specific IIC component, fructose-specific IIB component, enzyme I, alpha-glucoside-specific IIC component, maltose/glucose-specific IIC component, N-acetylglucosamine-specific IIC component, beta-glucoside-specific IIA component, sugar-specific IIA component, mannitol-specific IIA component.

Chuyển hóa của vi khuẩn trong môi trường đa dạng

Vi khuẩn thường gặp phải vô số các điều kiện bất lợi (stress) ảnh hưởng đến khả năng phát triển cả trong điều kiện tự nhiên và trong cơ thể vật chủ. Những stress này bao gồm sự hạn chế dinh dưỡng, stress oxy hóa, tổn thương thành tế bào, phơi nhiễm kháng sinh và các stress gây ảnh hưởng tăng trưởng khác [19-21]. Tế bào phản ứng với các stress này bằng cách kích thích các đáp ứng bảo vệ và/hoặc thích nghi nhằm mục đích gia tăng khả năng tồn tại của vi khuẩn. Các đáp ứng này thường dẫn đến những thay đổi đối với sinh lý tế bào, thúc đẩy sản xuất màng sinh học, kích hoạt cơ chế kháng thuốc và cảm ứng đột biến kháng thuốc [20, 22]. Do đó, việc xác định các protein kiểm soát và mạng lưới điều tiết là điều cần thiết để kiểm soát hoặc thay đổi hành vi vi khuẩn [22]. Đặc biệt, việc xác định các protein liên quan đến stress quan trọng sẽ cho thấy các mục tiêu cho một thao tác cụ thể để thúc đẩy hoặc hạn chế sự phát triển đề kháng kháng sinh. Trong nghiên cứu này, tổng cộng chín protein (narG, ureA,

ureB, ureC, gpi, hemL, ackA, pta và fbaA) trong số 158 protein thiết yếu ở MRSA đã được xác định tham gia vào các quá trình sinh tổng hợp khác nhau trong việc chống lại những điều kiện bất lợi của môi trường. Do đó, các protein này rất cần thiết cho sự tồn tại của MRSA đồng thời cũng thể hiện sự không tương đồng với hệ protein người. Việc thiết kế thuốc dựa trên các protein mục tiêu này có thể được coi là một chiến lược tiềm năng để điều trị tác động gây ra bởi MRSA.

Hệ thống hai thành phần

Hệ thống hai thành phần (TCS) bao gồm các cảm biến histidine kinase và các protein điều hòa đáp ứng được xem là những yếu tố quan trọng nhất trong việc truyền tín hiệu ở vi khuẩn [23]. Do tầm quan trọng của chúng đối với sự tồn tại, độc lực và sự phát triển của tế bào vi khuẩn, các hệ thống này nằm trong số những protein được nghiên cứu nhiều nhất. Trong những năm gần đây, một loạt các nghiên cứu về tin sinh học, di truyền và sinh hóa đã cung cấp những hiểu biết chi tiết về nhiều cơ chế phân tử của việc phát hiện tín hiệu và tạo ra đáp ứng thích hợp bằng hệ thống hai thành phần [23-25]. Quan trọng hơn, đã có nhiều nghiên cứu chỉ ra rằng có sự đa dạng đáng kể trong các cơ chế được sử dụng bởi từng loài vi khuẩn. Vì nhiều lý do, TCS đã trở thành mục tiêu thuyết phục cho việc thiết kế các hợp chất kháng khuẩn [25]. Trong nghiên cứu này, chúng tôi đã tìm thấy có 4 protein thuộc hệ thống hai thành phần ở MRSA bao gồm K⁺-transporting ATPase A chain, chromosomal replication initiator protein, vancomycin resistance sensor histidine kinase VraS và sensor histidine kinase. Các protein này do đó là những mục tiêu đáng chú ý trong việc ngăn chặn sự phát triển và gây bệnh của MRSA.

Quorum sensing

Quorum sensing (QS) là quá trình giao tiếp tế bào – tế bào, cho phép vi khuẩn chia sẻ thông tin về mật độ tế bào và điều chỉnh biểu hiện gen cho phù hợp [26]. Quá trình này cho phép vi khuẩn sử dụng năng lượng hợp lý khi tác động của các quá trình đó lên môi trường hoặc trên vật chủ sẽ được tối đa hóa. Một trong số nhiều

đặc điểm được kiểm soát bởi QS là sự biểu hiện của các yếu tố độc lực ở vi khuẩn gây bệnh [27]. Ở các loài staphylococci, sự điều hòa này rất quan trọng đối với thời gian biểu hiện yếu tố độc lực trong quá trình nhiễm khuẩn và sự phát triển của tình trạng bệnh cấp tính, trong khi hoạt động thấp của QS có liên quan đến nhiễm khuẩn mạn tính, chẳng hạn như sự hình thành màng sinh học. Theo đó, các thuốc ức chế QS đang được đánh giá khả năng kiểm soát các trường hợp nhiễm khuẩn *S. aureus* cấp tính [28]. Dựa trên kết quả phân tích KAAS, ba protein có thể là những mục tiêu tiềm năng cho các nghiên cứu ức chế hệ thống QS ở *S. aureus* cũng như MRSA đã được xác định. Các protein này bao gồm preprotein translocase subunit SecA, preprotein translocase subunit SecY, anthranilate synthase component II.

Sinh tổng hợp các chất chuyển hóa thứ cấp

Các tương tác vi khuẩn - vi khuẩn hoặc tương tác vi khuẩn - vật chủ là yếu tố quan trọng để vi khuẩn có thể phát triển trong những điều kiện môi trường khác nhau. Những tương tác này liên quan đến nhiều khía cạnh sinh học, bao gồm các thay đổi về mặt hóa lý, trao đổi chất, sinh tổng hợp chuyển hóa chất thứ cấp, truyền tín hiệu và trao đổi thông tin di truyền [29]. Do đó, các tương tác này là kết quả của quá trình đồng tiến hóa dẫn đến sự thích ứng và chuyển hóa, cho phép sự phát triển ở những điều kiện khác nhau bằng cách giảm stress sinh học và phi sinh học hoặc trao đổi các yếu tố tăng trưởng và tín hiệu. Nhiều cơ chế có thể tham gia vào các tương tác này, chẳng hạn như tăng sản xuất các chất chuyển hóa thứ cấp, hệ thống quorum sensing và sự hình thành màng sinh học. Đơn vị tương tác cuối cùng là biểu hiện gen chịu trách nhiệm cho việc sản xuất các phân tử liên quan của mỗi sinh vật để đáp ứng với một kích thích môi trường [30]. Do đó, kiến thức về các cơ chế liên quan đến quá trình sinh tổng hợp các chất chuyển hóa thứ cấp có thể là chìa khóa để phát triển các tác nhân cụ thể có thể tránh hoặc làm gián đoạn sự truyền nhiễm trong quá trình nhiễm khuẩn và do đó làm giảm hiệu quả phòng thủ hoặc tấn công của mầm bệnh. Ngoài ra, việc nghiên cứu các cơ chế này có thể góp phần vào sự hiểu biết về sinh bệnh

học của vi khuẩn và sự phát triển của các loại thuốc kháng khuẩn mới. Rõ ràng là các hợp chất thứ cấp đóng vai trò đáng kể trong sự đề kháng của nhiều trường hợp nhiễm khuẩn [31]. Do đó, tác động đến các mục tiêu này có thể bổ sung cho sự phát triển của các liệu pháp điều trị mới. Dựa trên kết quả phân tích KEGG, chúng tôi đã tìm thấy ở MRSA có 9 protein tham gia vào các quá trình sinh tổng hợp hợp chất chuyển hóa thứ cấp đặc trưng cho MRSA, trong đó bao gồm glucose-6-phosphate isomerase, glutamate-1-semialdehyde 2,1-aminomutase, tryptophan synthase beta chain, glycerol-3-phosphate acyltransferase, mannose-6-phosphate isomerase, fructose-bisphosphate aldolase, class II, 5-amino-6-(5-phospho-D-ribitylamino) uracil phosphatase, glycerol-3-phosphate acyltransferase, anthranilate synthase component II. Việc sử dụng các protein này làm mục tiêu thiết kế thuốc có thể sẽ dẫn đến sự phát triển các hợp chất ức chế sự phát triển vi khuẩn, không chỉ ở MRSA mà còn có thể ở nhiều vi khuẩn khác.

Hệ thống tiết của vi khuẩn

Quá trình tiết là một nhiệm vụ quan trọng của tế bào sinh vật nhân sơ để tương tác tốt hơn với môi trường xung quanh hoặc vật chủ của chúng. Đặc biệt, việc sản xuất các protein ngoại bào đặc biệt quan trọng đối với sự tồn tại và thích ứng của vi khuẩn [32]. Hệ thống tiết ở vi khuẩn có cấu trúc đa thành phần bao gồm nhiều nhóm protein và những chaperon điều hướng đi kèm. Vì vai trò của hệ thống này đối với nhóm vi khuẩn gây bệnh trong việc vận chuyển một loạt các yếu tố gây bệnh ra ngoài màng tế bào tế bào vi khuẩn và thiết lập khả năng trao đổi thông tin với môi trường, việc hiểu được các thành phần và chức năng của hệ thống vận chuyển protein này sẽ dẫn đến giải pháp phù hợp để ngăn chặn khả năng gây bệnh của chúng. Hơn nữa, các hệ thống này có thể được sử dụng cho mục đích cạnh tranh sinh học [33]. Kết quả từ phân tích KEGG cho thấy hai protein tham gia vào các hệ thống tiết vi khuẩn - secY và secA (cũng tham gia vào quorum sensing) - có thể được sử dụng như các mục tiêu thuốc tiềm năng chống lại MRSA.

Vancomycin resistance

Vancomycin là kháng sinh glycopeptide được sử dụng để điều trị nhiễm khuẩn Gram dương, đặc biệt trong các trường hợp nhiễm khuẩn do MRSA [34]. Thông thường, nó được sử dụng như kháng sinh điều trị hàng cuối; tuy nhiên, đã có những báo cáo lâm sàng về việc phân lập các chủng MRSA giảm tính nhạy cảm với Vancomycin (*S. aureus* kháng Vancomycin trung gian [VISA]) và gần đây hơn là một số chủng *S. aureus* kháng Vancomycin hoàn toàn (VRSA) [35]. Một số chủng VRSA mang gen kháng Vancomycin từ *Enterococcus faecalis* kháng Vancomycin đã được biết làm thay đổi cấu trúc thành tế bào và một số quá trình trao đổi chất, nhưng cơ chế đề kháng trong các trường hợp VISA phân lập được còn chưa được nghiên cứu nhiều [36]. Trong nghiên cứu này, chúng tôi đã tìm thấy bốn protein thiết yếu của *S. aureus* (ddl, murF, murG và mraY) có thể tham gia vào sự đề kháng Vancomycin. Mặc dù đây chỉ là nghiên cứu *in silico*, tuy nhiên những protein này vẫn có thể sử dụng làm mục tiêu tham khảo trong những nghiên cứu về cơ chế đề kháng Vancomycin ở *S. aureus* trong tương lai.

4. Kết luận

Tóm lại, bằng phương pháp phân tích loại trừ trong hệ protein sử dụng các công cụ tin học tính toán kết hợp với các cơ sở dữ liệu sinh học đã dẫn đến việc xác định và mô tả một số protein thiết yếu không tương đồng với các protein vật chủ. Vì vậy, những protein này cần cho sự tồn tại của mầm bệnh và do đó có thể được ứng dụng làm mục tiêu cho quá trình thiết kế thuốc dựa trên mục tiêu. Bên cạnh đó, các con đường trao đổi chất đặc trưng thu được trong nghiên cứu này cũng có thể được sử dụng làm thư viện cho việc khám phá các mục tiêu thuốc mới trong tương lai.

Lời cảm ơn

Nghiên cứu được tài trợ bởi Trường Đại học Bách khoa – Đại học Quốc gia TP. HCM trong khuôn khổ đề tài mã số TNCS-KTHH-2017-12.

Tài liệu tham khảo

- [1] S.J. Peacock, G.K. Paterson, Mechanisms of methicillin resistance in *Staphylococcus aureus*, *Annu. Rev. Biochem.* 84(1) (2015) 577-601.
- [2] F.D. Lowy, Antimicrobial resistance: the example of *Staphylococcus aureus*, *J. Clin. Invest.* 111(9) (2003) 1265-1273.
- [3] P.D. Stapleton, P.W. Taylor, Methicillin resistance in *Staphylococcus aureus*: mechanisms and modulation, *Sci. Prog.* 85(1) (2002) 57-72.
- [4] R. Uddin, K. Saeed, Identification and characterization of potential drug targets by subtractive genome analyses of methicillin resistant *Staphylococcus aureus*, *Comput. Biol. Chem.* 48 (2014) 55-63.
- [5] T. Hossain, M. Kamruzzaman, T.Z. Choudhury, H.N. Mahmood, A.H.M.N. Nabi, Md.I. Hosen, Application of the subtractive genomics and molecular docking analysis for the identification of novel putative drug targets against *Salmonella enterica* subsp. *Enterica serovar poona*, *BioMed Res. Int.* Vol. 2017 (2017) Article ID 3783714, 9 pages.
- [6] N. Haag, K. Velk, C. Wu, In silico identification of drug targets in methicillin/multidrug-resistant *Staphylococcus aureus*, In: BIOTECHNO 2011, The Third International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies, IARIA, Nice, 2011, pp. 91-99.
- [7] M. Hossain, D.U.S Chowdhury, J. Farhana, M.T. Akbar, A. Chakraborty, S. Islam, A. Mannan, Identification of potential targets in *Staphylococcus aureus* N315 using computer aided protein data analysis, *Bioinformation.* 9(4) (2013) 187-192.
- [8] S. Pundir, M.J. Martin, C. O'Donovan, Chapter 2. Protein Knowledgebase, *Methods Mol. Biol.* 1558 (2017) 41-55.
- [9] R.C. Edgar, Search and clustering orders of magnitude faster than BLAST, *Bioinformatics.* 26(19) (2010) 2460-2461.
- [10] M.A. Hasan, S.M. Alauddin, M. Al-Amin, S.M. Nur, A. Mannan, Identification of putative drug targets in vancomycin-resistant *Staphylococcus aureus* (VRSA) using computer aided protein data analysis, *Gene.* 575(1) (2016) 132-143.
- [11] T. UniProt Consortium, UniProt: the universal protein knowledgebase, *Nucleic Acids Res.* 46(5) (2018) 2699.
- [12] A. Gautam, R. Vyas, R. Tewari, Peptidoglycan biosynthesis machinery: A rich source of drug targets, *Crit. Rev. Biotechnol.* 31(4) (2011) 295-336.

- [13] C.K. Teo, D.I. Roper, Core steps of membrane-bound peptidoglycan biosynthesis: recent advances, insight and opportunities, *Antibiotics*. 4(4) (2015) 495-520.
- [14] Y. Liu, E. Breukink, The membrane steps of bacterial cell wall synthesis as antibiotic targets, *Antibiotics*. 5(3) (2016) 28.
- [15] A. Kovač, J. Konc, B. Vehar, J.M. Bostock, I. Chopra, D. Janežič, S. Gobec, Discovery of new inhibitors of d-alanine:d-alanine ligase by structure-based virtual screening, *J. Med. Chem.* 51(23) (2008) 7442-7448.
- [16] I. Tytgat, E. Colacino, P.M. Tulkens, J.H. Poupaert, M. Prevost, F. Van Bambeke, DD-ligases as a potential target for antibiotics: past, present and future, *Curr. Med. Chem.* 16(20) (2009) 2566-2580.
- [17] J. Deutscher, C. Francke, P.W. Postma, How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria, *Microbiol. Mol. Biol. Rev.* 70(4) (2006) 939-1031.
- [18] J. Deutscher, F.M.D. Aké, M. Derkaoui, A. C. Zébré, T.N. Cao, H. Bouraoui, T. Kentache, A. Mokhtari, E. Milohanic, P. Joyet, The bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions, *Microbiol. Mol. Biol. Rev.* 78(2) (2014) 231-256.
- [19] F.C. Fang, E.R. Frawley, T. Tapscott, A. Vazquez-Torres, Bacterial stress responses during host infection, *Cell Host Microbe*. 20(2) (2016) 133-143.
- [20] G. Renzone, C. D'Ambrosio, S. Arena, R. Rullo, L. Ledda, L. Ferrara, A. Scaloni, Differential proteomic analysis in the study of prokaryotes stress resistance, *Ann Ist Super Sanita* 41(4) (2005) 459-468.
- [21] G. Storz, R. Hengge-Aronis, Bacterial stress responses, ASM Press, Washington DC, 2000.
- [22] R.A. VanBogelen, K.D. Greis, M. Blumenthal, T.H. Tani, R. Matthews, Mapping regulatory networks in microbial cells, *Trends Microbiol.* 7(8) (1999) 320-328.
- [23] E. Bem, N. Velikova, M.T. Pellicer, P. van Baarlen, A. Marina, J.M. Wells, Bacterial histidine kinases as novel antibacterial drug targets, *ACS Chem. Biol.* 10(1) (2015) 213-224.
- [24] S. Wang, Bacterial two-component systems: structures and signaling mechanisms, In C. Huang (Eds.), *Protein Phosphorylation in Human Health*, IntechOpen, London, 2012, pp. 439-466.
- [25] S. Tiwari, S.B. Jamal, S.S. Hassan, P.V.S.D. Carvalho, S. Almeida, D. Barh, P. Ghosh, A. Silva, T.L.P. Castro, V. Azevedo, Two-component signal transduction systems of pathogenic bacteria as targets for antimicrobial therapy: an overview, *Front. Microbiol.* 8 (2017) 1878.
- [26] S.T. Rutherford, B.L. Bassler, Bacterial quorum sensing: its role in virulence and possibilities for its control, *Cold Spring Harb. Perspect. Med.* 2(11) (2012) a012427.
- [27] J. Park, R. Jagasia, G.F. Kaufmann, J.C. Mathison, D.I. Ruiz, J.A. Moss, M.M. Meijler, R.J. Ulevitch, K.D. Janda, Infection control by antibody disruption of bacterial quorum sensing signaling, *Chem. Biol.* 14 (2007) 1119-1127.
- [28] M. Otto, Virulence factors of the coagulase-negative staphylococci, *Front. Biosci.* 9 (2004) 841-863.
- [29] R.M. Braga, M.N. Dourado, W.L. Araújo, Microbial interactions: ecology in a molecular perspective, *Braz. J. Microbiol.* 47(Suppl 1) (2016) 86-98.
- [30] L. Demain, A. Fang, The natural functions of secondary metabolites, *Adv. Biochem. Eng. Biotechnol.* 69 (2000) 1-39.
- [31] H. Nikaido, Multidrug resistance in bacteria, *Annu. Rev. Biochem.* 78 (2009) 119-146.
- [32] E.R. Green, J. Meccas, Bacterial secretion systems – an overview, *Microbiol. Spectr.* 4(1) (2016) 10.1128/microbiolspec.VMBF-0012-2015.
- [33] E.G. Parizad, E.G. Parizad, I. Pakzad, A. Valizadeh, A review of secretion systems in pathogenic and non-pathogenic bacteria, *Biosci. Biotech. Res. Asia.* 13(1) (2016) 135-145.
- [34] A. Srinivasan, J.D. Dick, T.M. Perl, Vancomycin resistance in Staphylococci, *Clin. Microbiol. Rev.* 15(3) (2002) 430-438.
- [35] S. Gardete, A. Tomasz, Mechanisms of vancomycin resistance in *Staphylococcus aureus*, *J. Clin. Invest.* 124(7) (2014) 2836-2840.
- [36] W.A. McGuinness, N. Malachowa, F.R. DeLeo, Vancomycin resistance in *Staphylococcus aureus*, *Yale J. Biol. Med.* 90(2) (2017) 269-281.

Phụ lục 1

**Các protein thiết yếu ở MRSA
(Xác định từ phân tích với CSDL DEG)**

Số thứ tự	Mã số UniprotKB/ Uniparc	Số protein phù hợp tại DEG	Mã số DEG	Số thứ tự	Mã số UniprotKB/ Uniparc	Số protein phù hợp tại DEG	Mã số DEG
1	UPI0001AB17B9	1	DEG10020135;	80	UPI00000D9E46	1	DEG10020207;
2	UPI0001C1200D	1	DEG10020284;	81	UPI0001AB1796	1	DEG10020118;
3	UPI00003B153D	1	DEG10020219;	82	UPI0000009B95	1	DEG10020207;
4	UPI0001C11CF9	1	DEG10020066;	83	UPI00003B1774	1	DEG10020207;
5	UPI0001AE9B8D	1	DEG10020201;	84	UPI000005480C	1	DEG10020053;
6	UPI0001C11C84	1	DEG10020016;	85	UPI00000543CB	1	DEG10020107;
7	UPI00003B1501	1	DEG10020067;	86	UPI00001356AB	1	DEG10020207;
8	UPI00003B169C	1	DEG10020135;	87	UPI0004483B3E	1	DEG10020252;
9	UPI00027E5DE3	2	DEG10020150; DEG10020004;	88	UPI000005476E	1	DEG10020069;
10	UPI00003B14FC	2	DEG10020297; DEG10020064;	89	UPI0001AE9B80	1	DEG10020296;
11	UPI00003B15DE	1	DEG10020115;	90	UPI00000D7844	1	DEG10020296;
12	UPI0001C120BC	2	DEG10020297; DEG10020064;	91	UPI0000054778	1	DEG10020068;
13	UPI00000CAA08	1	DEG10020219;	92	UPI0002550C16	1	DEG10020059;
14	UPI00003B18AA	1	DEG10020129;	93	UPI000195C503	1	DEG10020245;
15	UPI00003B1989	1	DEG10020290;	94	UPI00003B142E	1	DEG10020026;
16	UPI00019CD777	1	DEG10020290;	95	UPI000006032D	1	DEG10020059;
17	UPI00000CADEA	1	DEG10020218;	96	UPI00000374E9	1	DEG10020195;
18	UPI00003B14DA	2	DEG10020060; DEG10020295;	97	UPI00003B14CB	1	DEG10020058;
19	UPI0001C120B6	2	DEG10020060; DEG10020295;	98	UPI0001C120F6	1	DEG10020241;
20	UPI00003B170A	1	DEG10020170;	99	UPI00000D7773	1	DEG10020239;
21	UPI00004B5BAA	1	DEG10020104;	100	UPI0001C12081	1	DEG10020241;
22	UPI0000054B5C	1	DEG10020092;	101	UPI0004467E3F	1	DEG10020151;
23	UPI000005480D	1	DEG10020282;	102	UPI00003B13CC	1	DEG10020017;

Số thứ tự	Mã số UniprotKB/Uniparc	Số protein phù hợp tại DEG	Mã số DEG	Số thứ tự	Mã số UniprotKB/Uniparc	Số protein phù hợp tại DEG	Mã số DEG
24	UPI00003B1429	1	DEG10020026;	103	UPI00003B141B	2	DEG10020025; DEG10020076;
25	UPI0000054B55	2	DEG10020093; DEG10020140;	104	UPI00003B1A11	1	DEG10020017;
26	UPI00003B184D	1	DEG10020227;	105	UPI000013918D	1	DEG10020030;
27	UPI00003B18DE	1	DEG10020290;	106	UPI00000DA04D	1	DEG10020249;
28	UPI00000522AA	1	DEG10020283;	107	UPI00000601EC	1	DEG10020241;
29	UPI00000DCBDB	1	DEG10020290;	108	UPI000012FCE0	1	DEG10020221;
30	UPI00000D4753	1	DEG10020116;	109	UPI00003B187B	2	DEG10020060; DEG10020243;
31	UPI000012F9AC	1	DEG10020229;	110	UPI00000D76A9	1	DEG10020271;
32	UPI0000696C65	1	DEG10020290;	111	UPI0000054BF8	1	DEG10020246;
33	UPI00025509B6	1	DEG10020109;	112	UPI00003B18C5	1	DEG10020281;
34	UPI0001F632B7	1	DEG10020007;	113	UPI0001F63712	1	DEG10020139;
35	UPI00003B12E0	1	DEG10020001;	114	UPI000004801B	1	DEG10020264;
36	UPI00003B1851	1	DEG10020229;	115	UPI00003B185D	2	DEG10020020; DEG10020223;
37	UPI00003B1436	1	DEG10020033;	116	UPI000013C17D	1	DEG10020015;
38	UPI000012F989	1	DEG10020108;	117	UPI0000052289	1	DEG10020046;
39	UPI000012E109	1	DEG10020028;	118	UPI0000054C69	1	DEG10020268;
40	UPI0005ADADD6	1	DEG10020258;	119	UPI0000054BB3	1	DEG10020251;
41	UPI00000549F0	1	DEG10020148;	120	UPI0000110EF1	1	DEG10020301;
42	UPI00003B1766	1	DEG10020203;	121	UPI0000054ADA	1	DEG10020273;
43	UPI00000543F1	1	DEG10020106;	122	UPI00001331BF	1	DEG10020138;
44	UPI00003B1660	1	DEG10020024;	123	UPI0000133CC4	1	DEG10020127;
45	UPI00024334A0	1	DEG10020012;	124	UPI0000054501	1	DEG10020027;
46	UPI000013B7E6	1	DEG10020135;	125	UPI0000054C5E	1	DEG10020267;
47	UPI0002B149BD	1	DEG10020266;	126	UPI00000D766C	1	DEG10020185;
48	UPI0000054C2C	1	DEG10020263;	127	UPI00000D77A0	1	DEG10020280;
49	UPI00000D7824	1	DEG10020143;	128	UPI0001B3ABFF	1	DEG10020091;

Số thứ tự	Mã số UniprotKB/Uniparc	Số protein phù hợp tại DEG	Mã số DEG	Số thứ tự	Mã số UniprotKB/Uniparc	Số protein phù hợp tại DEG	Mã số DEG
50	UPI00003B1749	1	DEG10020191;	129	UPI0000167807	1	DEG10020285;
51	UPI00000D78AB	1	DEG10020232;	130	UPI0002550F93	1	DEG10020208;
52	UPI00003B1685	1	DEG10020155;	131	UPI000012AF6D	1	DEG10020081;
53	UPI0001AE8711	1	DEG10020290;	132	UPI00005FE6DA	1	DEG10020238;
54	UPI000243207D	1	DEG10020045;	133	UPI00003B15F4	2	DEG10020300; DEG10020130;
55	UPI0005ADF421	1	DEG10020167;	134	UPI0000054638	1	DEG10020037;
56	UPI0001C11E13	1	DEG10020156;	135	UPI0000054BFA	1	DEG10020258;
57	UPI000005428C	1	DEG10020176;	136	UPI00003B1734	2	DEG10020018; DEG10020183;
58	UPI0000126061	1	DEG10020074;	137	UPI00003B173F	1	DEG10020011;
59	UPI0000054C2E	1	DEG10020193;	138	UPI00003B18AF	1	DEG10020285;
60	UPI000013602F	1	DEG10020075;	139	UPI00003B185E	1	DEG10020231;
61	UPI0002550CE7	1	DEG10020006;	140	UPI0001AE9A2F	1	DEG10020182;
62	UPI0000060A16	2	DEG10020025; DEG10020076;	141	UPI00000549D5	1	DEG10020285;
63	UPI0000054B94	1	DEG10020249;	142	UPI0001AE8448	1	DEG10020231;
64	UPI0000E56B10	1	DEG10020249;	143	UPI0000060333	1	DEG10020212;
65	UPI0001C11DC6	1	DEG10020241;	144	UPI0001AE996D	1	DEG10020285;
66	UPI00000543F2	1	DEG10020105;	145	UPI0000054ADD	1	DEG10020285;
67	UPI0002432AA7	1	DEG10020008;	146	UPI0001C33AD6	1	DEG10020007;
68	UPI0000060A50	10	DEG10020250; DEG10020013; DEG10020014; DEG10020286; DEG10020287; DEG10020222; DEG10020083; DEG10020097; DEG10020057; DEG10020298;	147	UPI00003B1663	1	DEG10020152;
69	UPI00000B9AC7	1	DEG10020290;	148	UPI00003B175E	1	DEG10020202;
70	UPI0001AE84D3	1	DEG10020169;	149	UPI00000603AC	1	DEG10020114;
71	UPI00003B15EB	1	DEG10020126;	150	UPI00000D773B	1	DEG10020010;

Số thứ tự	Mã số UniprotKB/Uniparc	Số protein phù hợp tại DEG	Mã số DEG	Số thứ tự	Mã số UniprotKB/Uniparc	Số protein phù hợp tại DEG	Mã số DEG
72	UPI00003B1A12	1	DEG10020299;	151	UPI0001C11DAA	1	DEG10020136;
73	UPI00003B144A	1	DEG10020044;	152	UPI00000543C7	1	DEG10020110;
74	UPI0000054C73	1	DEG10020162;	153	UPI000006061E	1	DEG10020158;
75	UPI0000134FAF	1	DEG10020272;	154	UPI00000CAB19	1	DEG10020213;
76	UPI00003B1779	2	DEG10020209; DEG10020102;	155	UPI000005498D	1	DEG10020002;
77	UPI000005449B	1	DEG10020213;	156	UPI00000B98FA	1	DEG10020244;
78	UPI000013352D	1	DEG10020142;	157	UPI0001C11E0E	1	DEG10020154;
79	UPI0000054B87	1	DEG10020213;	158	UPI0001C1203D	1	DEG10020230;

Phụ lục 2

Các con đường chuyển hóa đặc trưng của MRSA

Đường chuyển hóa chính	Các đường chuyển hóa tương ứng	Các đường chuyển hóa đặc trưng	Mã số KEGG
Sinh tổng hợp	Chuyển hóa lipid	Sinh tổng hợp bile acid thứ cấp	sau00121
Sinh tổng hợp	Sinh tổng hợp các chất chuyển hóa thứ cấp khác	Sinh tổng hợp monobactam	sau00261
Sinh tổng hợp	Chuyển hóa các terpenoid và polyketide	Thoái hóa geraniol	sau00281
Sinh tổng hợp	Chuyển hóa amino acid	Sinh tổng hợp lysine	sau00300
Sinh tổng hợp	Thoái hóa các xenobiotic và chuyển hóa	Thoái hóa benzoate	sau00362
Sinh tổng hợp	Sinh tổng hợp các chất chuyển hóa thứ cấp khác	Sinh tổng hợp novobiocin	sau00401
Sinh tổng hợp	Chuyển hóa của các amino acid	Chuyển hóa cyanoamino acid	sau00460
Sinh tổng hợp	Chuyển hóa của các amino acid	Chuyển hóa D-Alanine	sau00473
Sinh tổng hợp	Sinh tổng hợp các chất chuyển hóa thứ cấp khác	Sinh tổng hợp streptomycin	sau00521
Sinh tổng hợp	Sinh tổng hợp glycan và chuyển hóa	Sinh tổng hợp peptidoglycan	sau00550
Sinh tổng hợp	Thoái hóa các xenobiotic và chuyển hóa	Thoái hóa xylene	sau00622
Sinh tổng hợp	Thoái hóa các xenobiotic và chuyển hóa	Thoái hóa chloroalkane and chloroalkene	sau00625
Sinh tổng hợp	Thoái hóa các xenobiotic và chuyển hóa	Thoái hóa naphthalene	sau00626
Sinh tổng hợp	Chuyển hóa carbohydrate	Chuyển hóa C5-Branched dibasic acid	sau00660
Sinh tổng hợp	Chuyển hóa năng lượng	Chuyển hóa Methane	sau00680
Sinh tổng hợp	Chuyển hóa các terpenoid và polyketide	Sinh tổng hợp carotenoid	sau00906
		Sinh tổng hợp chất chuyển hóa thứ cấp	sau01110
		Chuyển hóa vi khuẩn trong các môi trường đa dạng	sau01120
		Sinh tổng hợp các kháng sinh	sau01130
		Thoái hóa các hợp chất vòng thơm	sau01220
Yếu tố gây bệnh	Đề kháng thuốc	Kháng beta-Lactam	sau01501
Yếu tố gây bệnh	Đề kháng thuốc	Kháng vancomycin	sau01502
Yếu tố gây bệnh	Đề kháng thuốc	Kháng cationic antimicrobial peptide (CAMP)	sau01503
Yếu tố trao đổi thông tin với môi trường	Truyền tín hiệu	Hệ thống hai thành phần	sau02020
Yếu tố trao đổi thông tin với môi trường	Vận chuyển qua màng	Hệ thống phosphotransferase (PTS)	sau02060
Yếu tố trao đổi thông tin với môi trường	Vận chuyển qua màng	Hệ thống tiết của vi khuẩn	sau03070
Chu trình tế bào	Cộng đồng tế bào prokaryote	Quorum sensing	sau02024

Phụ lục 3

Các protein liên quan đến các con đường chuyển hóa đặc trưng

UniprotKB/ Uniparc ID	Mã số KO	Tên protein; Chức năng; Mã số EC (nếu có)	Chu trình
UPI000005480D	K01428	ureC; urease subunit alpha [EC:3.5.1.5]	Chuyển hóa của vi khuẩn trong môi trường đa dạng
UPI00003B18C5	K01429	ureB; urease subunit beta [EC:3.5.1.5]	Chuyển hóa của vi khuẩn trong môi trường đa dạng
UPI00000D77A0	K01430	ureA; urease subunit gamma [EC:3.5.1.5]	Chuyển hóa của vi khuẩn trong môi trường đa dạng
UPI0001C1200D	K00370	narG, narZ, nxrA; nitrate reductase/nitrite oxidoreductase, alpha subunit [EC:1.7.5.1 1.7.99.-]	Chuyển hóa của vi khuẩn trong môi trường đa dạng, Hệ thống hai thành phần
UPI00003B14DA	K02769	PTS-Fru-EIIB, fruA; PTS system, fructose-specific IIB component [EC:2.7.1.202]	Chuyển hóa của vi khuẩn trong môi trường đa dạng, Phosphotransferase system (PTS)
UPI0001C120B6	K02769	PTS-Fru-EIIB, fruA; PTS system, fructose-specific IIB component [EC:2.7.1.202]	Chuyển hóa của vi khuẩn trong môi trường đa dạng, Phosphotransferase system (PTS)
UPI0001C1203D	K01921	ddl; D-alanine-D-alanine ligase [EC:6.3.2.4]	Chuyển hóa D-Alanine, Sinh tổng hợp Peptidoglycan, Vancomycin resistance
UPI000005480C	K00625	E2.3.1.8, pta; phosphate acetyltransferase [EC:2.3.1.8]	Chuyển hóa methane, Chuyển hóa của vi khuẩn trong môi trường đa dạng
UPI00003B175E	K00925	ackA; acetate kinase [EC:2.7.2.1]	Chuyển hóa methane, Chuyển hóa của vi khuẩn trong môi trường đa dạng
UPI00000D7773	K01624	FBA, fbaA; fructose-bisphosphate aldolase, class II [EC:4.1.2.13]	Chuyển hóa methane, Sinh tổng hợp các chất chuyển hóa thứ cấp, Chuyển hóa của vi khuẩn trong môi trường đa dạng, Sinh tổng hợp kháng sinh
UPI00003B184D	K01546	kdpA; K ⁺ -transporting ATPase ATPase A chain [EC:3.6.3.12]	Hệ thống hai thành phần
UPI00003B12E0	K02313	dnaA; chromosomal replication initiator protein	Hệ thống hai thành phần
UPI000005449B	K07681	vraS; two-component system, NarL family, vancomycin resistance sensor histidine kinase VraS [EC:2.7.13.3]	Hệ thống hai thành phần
UPI0000054B87	K07683	nreB; two-component system, NarL family, sensor histidine kinase NreB [EC:2.7.13.3]	Hệ thống hai thành phần
UPI00003B1989	K20118	PTS-Glc1-EIIC, ptsG, glcA, glcB; PTS system, glucose-specific IIC component	Phosphotransferase system (PTS)
UPI00019CD777	K20118	PTS-Glc1-EIIC, ptsG, glcA, glcB; PTS system, glucose-specific IIC component	Phosphotransferase system (PTS)
UPI0000054B5C	K08483	PTS-El.PTSI, ptsI; phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9]	Phosphotransferase system (PTS)
UPI00003B18DE	K02750	PTS-Glv-EIIC, glvC, malP, aglA; PTS	Phosphotransferase system (PTS)

UniprotKB/ Uniparc ID	Mã số KO	Tên protein; Chức năng; Mã số EC (nếu có)	Chu trình
UPI0000DCBDB	K02791	system, alpha-glucoside-specific IIC component PTS-MalGlc-EIIC, malX; PTS system, maltose/glucose-specific IIC component	Phosphotransferase system (PTS)
UPI0000696C65	K02804	PTS-Nag-EIIC, nagE; PTS system, N-acetylglucosamine-specific IIC component	Phosphotransferase system (PTS)
UPI00000B9AC7	K02755	PTS-Bgl-EIIA, bglF, bglP; PTS system, beta-glucoside-specific IIA component [EC:2.7.1.-]	Phosphotransferase system (PTS)
UPI0001AE8711	K02777	PTS-Glc-EIIA, crr; PTS system, sugar-specific IIA component [EC:2.7.1.-]	Phosphotransferase system (PTS)
UPI00003B187B	K02798	PTS-Mtl-EIIA, mtlA, cmtB; PTS system, mannitol-specific IIA component [EC:2.7.1.197]	Phosphotransferase system (PTS)
UPI0000054BFA	K03076	secY; preprotein translocase subunit SecY	Quorum sensing , Hệ thống tiết của vi khuẩn
UPI00003B14FC	K00370	secA; preprotein translocase subunit SecA	Quorum sensing, Hệ thống tiết của vi khuẩn
UPI0001C120BC	K00370	PTS-Glc1-EIIC, ptsG, glcA, glcB; PTS system, glucose-specific IIC component	Quorum sensing, Hệ thống tiết của vi khuẩn
UPI0005ADADD6	K03076	secY; preprotein translocase subunit SecY	Quorum sensing, Hệ thống tiết của vi khuẩn
UPI0001AB1796	K03621	plsX; glycerol-3-phosphate acyltransferase PlsX [EC:2.3.1.15]	Sinh tổng hợp các chất chuyển hóa thứ cấp
UPI00000601EC	K21064	ycsE, yitU, ywtE; 5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase [EC:3.1.3.104]	Sinh tổng hợp các chất chuyển hóa thứ cấp
UPI00000549F0	K08591	plsY; glycerol-3-phosphate acyltransferase PlsY [EC:2.3.1.15]	Sinh tổng hợp các chất chuyển hóa thứ cấp
UPI000012AF6D	K01810	GPI, pgi; glucose-6-phosphate isomerase [EC:5.3.1.9]	Sinh tổng hợp các chất chuyển hóa thứ cấp, Chuyển hóa của vi khuẩn trong môi trường đa dạng, Sinh tổng hợp kháng sinh
UPI00003B173F	K01845	hemL; glutamate-1-semialdehyde 2,1-aminomutase [EC:5.4.3.8]	Sinh tổng hợp các chất chuyển hóa thứ cấp, Chuyển hóa của vi khuẩn trong môi trường đa dạng
UPI00003B1663	K01696	trpB; tryptophan synthase beta chain [EC:4.2.1.2]	Sinh tổng hợp các chất chuyển hóa thứ cấp, Sinh tổng hợp kháng sinh
UPI0001AE9B80	K01809	manA, MPI; mannose-6-phosphate isomerase [EC:5.3.1.8]	Sinh tổng hợp các chất chuyển hóa thứ cấp, Sinh tổng hợp kháng sinh
UPI00000D7844	K01809	manA, MPI; mannose-6-phosphate isomerase [EC:5.3.1.8]	Sinh tổng hợp các chất chuyển hóa thứ cấp, Sinh tổng hợp kháng sinh
UPI00003B1660	K01658	trpG; anthranilate synthase component II [EC:4.1.3.27]	Sinh tổng hợp các chất chuyển hóa thứ cấp, Sinh tổng hợp kháng sinh, Quorum sensing
UPI00003B1436	K04042	glmU; bifunctional UDP-N-acetylglucosamine pyrophosphorylase / Glucosamine-1-phosphate N-acetyltransferase [EC:2.7.7.23]	Sinh tổng hợp kháng sinh

UniprotKB/ Uniparc ID	Mã số KO	Tên protein; Chức năng; Mã số EC (nếu có)	Chu trình
UPI00000DCBDB	K05362	2.3.1.157] murE; UDP-N-acetylmuramoyl-L- alanyl-D-glutamate-L-lysine ligase [EC:6.3.2.7]	Sinh tổng hợp Peptidoglycan
UPI000012F989	K01925	murD; UDP-N-acetylmuramoylalanine- -D-glutamate ligase [EC:6.3.2.9]	Sinh tổng hợp Peptidoglycan
UPI0000167807	K11694	femA; peptidoglycan pentaglycine glycine transferase (the second and third glycine) [EC:2.3.2.17]	Sinh tổng hợp Peptidoglycan
UPI0002550F93	K01924	murC; UDP-N-acetylmuramate--alanine ligase [EC:6.3.2.8]	Sinh tổng hợp Peptidoglycan
UPI00003B18AF	K11693	femX, fmhB; peptidoglycan pentaglycine glycine transferase (the first glycine) [EC:2.3.2.16]	Sinh tổng hợp Peptidoglycan
UPI00003B185E	K00790	murA; UDP-N-acetylglucosamine 1- carboxyvinyltransferase [EC:2.5.1.7]	Sinh tổng hợp Peptidoglycan
UPI00000549D5	K11695	femB; peptidoglycan pentaglycine glycine transferase (the fourth and fifth glycine) [EC:2.3.2.18]	Sinh tổng hợp Peptidoglycan
UPI0001AE8448	K00790	murA; UDP-N-acetylglucosamine 1- carboxyvinyltransferase [EC:2.5.1.7]	Sinh tổng hợp Peptidoglycan
UPI00003B14CB	K06153	bacA; undecaprenyl-diphosphatase [EC:3.6.1.27]	Sinh tổng hợp Peptidoglycan
UPI00003B1851	K01929	murF; UDP-N-acetylmuramoyl- tripeptide--D-alanyl-D-alanine ligase [EC:6.3.2.10]	Sinh tổng hợp Peptidoglycan, Vancomycin resistance
UPI0001C11E0E	K02563	murG; UDP-N-acetylglucosamine--N- acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N- acetylglucosamine transferase [EC:2.4.1.227]	Sinh tổng hợp Peptidoglycan, Vancomycin resistance
UPI00000543CB	K01000	mraY; phospho-N-acetylmuramoyl- pentapeptide-transferase [EC:2.7.8.13]	Sinh tổng hợp Peptidoglycan, Vancomycin resistance