

NGHIÊN CỨU ỨNG DỤNG CÔNG NGHỆ GIẢI TRÌNH TỰ THỂ HỆ MỚI TRÊN HỆ GENE VI KHUẨN *BURKHOLDERIA PSEUDOMALLEI*

Lê Minh Quý¹, Lê Thị Bích Trâm^{1*}, Bùi Liêm Chính¹
Trần Trọng Hội¹, Trần Thị Hạnh¹, Đinh Bá Tuấn¹
Vũ Thị Ngọc Anh², Trần Thị Lệ Quyên², Trịnh Thành Trung²

TÓM TẮT

Mục tiêu: Bổ sung một kỹ thuật mới trong định danh vi khuẩn *B. pseudomallei* bằng công nghệ giải trình tự gene thể hệ mới.

Đối tượng và phương pháp: Ứng dụng kỹ thuật giải trình tự gene thể hệ mới trên hệ thống máy MiSeq (Illumina - Hoa Kỳ), giải trình tự toàn bộ hệ gene của hai chủng vi khuẩn *B. pseudomallei*. Đánh giá chất lượng giải trình tự gene bằng phần mềm FastQC. Định danh vi khuẩn bằng trình tự 16S rRNA sử dụng phần mềm Blast.

Kết quả: Phần mềm FastQC cho thấy điểm chất lượng Q đạt trên Q30 (độ chính xác của đọc trình tự là trên 99,9%), chỉ số N50 của hai chủng vi khuẩn là 36.559 bp và 74.123 bp. Phần mềm Blast cho kết quả ở cả 2 chủng vi khuẩn là *B. pseudomallei* với độ bao phủ và độ tương đồng đều đạt 100%. Trình tự kiểu gene (ST) được xác định là ST 541, với kiểu alen của các "house-keeping" genes là ace: 3, gltB: 4, gmhD: 2, lepA: 3, lipA: 5, narK: 4 và ndh: 1. Sử dụng công cụ ABRicate, gene kháng kháng sinh beta-lactam được xác định ở cả hai chủng. Sử dụng cơ sở dữ liệu VFDB, xác định được 165 gene độc lực ở chủng NTW 21.1 và 171 gene độc lực ở chủng SS 1 240'.

Từ khóa: Melioidosis, *B. pseudomallei*, định danh, giải trình tự hệ gene.

ABSTRACT

Objective: Implemented a new technique for bacterial identification of *B.pseudomallei* by next-generation sequencing technology.

Subjects and methods: Applications of next-generation sequencing technique on MiSeq system (Illumina - USA) to sequence the entire gene set of two strains of *B. pseudomallei* bacteria. The quality of the gene sequencing had evaluated by FastQC software. Bacterial identification had performed using 16S rRNA sequencing with Blast software.

Results: FastQC software indicated that the quality scores (Q scores) reached Q30 (sequence read accuracy above 99.9%). The N50 index for the two bacterial strains was 36,559 bp and 74,123 bp, respectively. Blast software confirmed that both strains had been identified as *B. pseudomallei*, with 100% coverage and similarity. The gene sequence type (ST) was determined as ST 541, with allele types for "house-keeping" genes as follows: ace: 3, gltB: 4, gmhD: 2, lepA: 3, lipA: 5, narK: 4, and ndh: 1. Using the ABRicate tool, beta-lactam antibiotic resistance genes were identified in both strains. Utilizing the VFDB database, 165 virulence genes were identified in the NTW 21.1 strain, and 171 virulence genes were identified in the SS 1 240' strain.

Keywords: Melioidosis, *B. pseudomallei*, identification, whole-genome sequencing (WGS).

Chịu trách nhiệm nội dung: Lê Thị Bích Trâm, Email: lebichtram94@gmail.com

Ngày nhận bài: 08/5/2023; mời phản biện khoa học: 5/2023; chấp nhận đăng: 15/6/2023.

¹Viện Khoa học và Công nghệ, Bộ Công an

²Viện Vi sinh vật và Công nghệ sinh học, Đại học Quốc gia Hà Nội

1. ĐẶT VẤN ĐỀ

Burkholderia pseudomallei là vi khuẩn Gram âm hình que, sống hoại sinh trong đất và nước bề mặt ở vùng nhiệt đới, đặc biệt là ở Đông Nam Á và Bắc Úc. Vi khuẩn *B. pseudomallei* là tác nhân gây bệnh truyền nhiễm melioidosis (còn gọi là bệnh Whitmore) ở người cùng nhiều loài động vật nuôi

hoặc hoang dã. Trên người, bệnh có biểu hiện lâm sàng đa dạng, tiến triển viêm phổi và nhiễm khuẩn huyết cấp tính, gây tử vong nhanh với tỉ lệ tử vong cao nếu bệnh nhân không được chẩn đoán đúng và điều trị kháng sinh theo đúng khuyến cáo. Gần đây, việc nâng cao năng lực chẩn đoán và xét nghiệm vi sinh đã giúp nhiều cơ sở y tế phát hiện

ra ca bệnh, từ đó ghi nhận số ca nhiễm melioidosis ngày một gia tăng ở nhiều nước. Theo dự báo, mỗi năm trên thế giới có khoảng 189.000 người nhiễm melioidosis, trong đó có đến 89.000 người tử vong; riêng ở Việt Nam mỗi năm có khoảng 10.000 người nhiễm melioidosis và gần 5.000 người tử vong [6].

Vi khuẩn này vẫn đang được xem là một tác nhân sinh học đe dọa nghiêm trọng sức khỏe con người [10] và đã được liệt kê vào danh sách tác nhân sinh học nhóm B [11]. Trong lĩnh vực an ninh - quốc phòng, *B. pseudomallei* cùng một số vi khuẩn nguy hiểm khác được quan tâm nghiên cứu và kiểm soát bởi khả năng gây bệnh nhanh, tỉ lệ tử vong cao và có nguy cơ bị sử dụng làm vũ khí sinh học. *B. pseudomallei* được định danh bằng nhiều phương pháp khác nhau, như dựa vào đặc tính sinh hóa và nhạy cảm kháng sinh, kĩ thuật real-time PCR đặc hiệu, giải trình tự gene 16S rRNA và *recA*... Độ chính xác của các phương pháp đã được so sánh và đánh giá [1]; song, định danh *B. pseudomallei* bằng các kĩ thuật sinh học phân tử vẫn được coi là tiêu chuẩn vàng trong các nghiên cứu [2].

Nhằm tiếp cận phương pháp mới trong nghiên cứu vi khuẩn *B. pseudomallei* tại phòng thí nghiệm, chúng tôi thực hiện nghiên cứu ứng dụng kĩ thuật giải trình tự gene thế hệ mới để giải trình tự toàn bộ hệ gene của vi khuẩn *B. pseudomallei*. Đây là kĩ thuật được đánh giá có hiệu suất và tính chính xác cao, ngày càng được phát triển và ứng dụng mạnh mẽ trên nhiều lĩnh vực, bao gồm cả nghiên cứu hệ gene vi sinh vật. Cùng với kết quả định danh vi khuẩn, nghiên cứu này sử dụng các công cụ phần mềm đánh giá một số thông tin về dịch tễ học, tính kháng thuốc và mối quan hệ di truyền của các chủng trong cùng một loài.

2. ĐỐI TƯỢNG, PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng, nguyên liệu, thiết bị nghiên cứu

- Đối tượng nghiên cứu: 2 chủng vi khuẩn *B. pseudomallei*, kí hiệu NTW 21.1 và SS 1 240' (mã trình tự tương ứng là 22-P6H09-PBS-01 và 22-P6H09-PBS-02), do Viện Vi sinh vật và Công nghệ sinh học (Đại học Quốc gia Hà Nội) cung cấp. Các chủng này được phân lập từ mẫu bệnh phẩm lâm sàng và môi trường trong chùm ca bệnh xảy ra tại Sóc Sơn (Hà Nội) năm 2019.

- Nguyên liệu nghiên cứu:

+ Hóa chất sử dụng trong tách chiết hệ gene vi khuẩn: tris (Serva, Đức), EDTA-disodium salt dihydrate (BioBasic, Canada), sodium chloride (Merck, Đức), cetyltrimethyl ammonium bromide

(CTAB) (Serva, Đức), lysozyme (Amersham Pharmacia Biotech Inc, Hoa Kỳ), proteinase K, RNase A (ThermoFisher, Hoa Kỳ), sodium dodecyl sulfate (SDS) (Fisher Chemical, Hoa Kỳ), chloroform (Avantor Performance Materials, Hà Lan), isoamyl alcohol (Merck, Đức), ethanol absolute (Merck, Đức), isopropanol (Merck, Đức).

+ Hóa chất, vật tư, sinh phẩm sử dụng cho giải trình tự hệ gene vi khuẩn: MiSeq Reagent Micro Kit v2 và Illumina DNA Prep (M) Tagmentation 24 Samples (Illumina, Hoa Kỳ); các hóa chất vật tư tiêu khác sử dụng theo đúng hướng dẫn của hãng Illumina.

- Thiết bị nghiên cứu: hệ thống máy MiSeq (Illumina, Hoa Kỳ), máy định lượng DNA Quantus (Promega, Hoa Kỳ), máy ProFlex PCR System (Thermo Scientific, Hoa Kỳ), máy NanoDrop 2000 (Thermo Fisher Scientific, Hoa Kỳ), micropipette và các dụng cụ sinh học phân tử cần thiết khác.

2.2. Phương pháp nghiên cứu

- Tách chiết hệ gene vi khuẩn: hệ gene vi khuẩn tách chiết theo quy trình của Gabor và cộng sự [8] với một số sửa đổi. Vi khuẩn được hoạt hóa từ ống bảo quản lạnh sâu (âm 80°C) bằng cách cấy ria trên môi trường Muller Hinton thạch và nuôi ở 37°C trong 18 giờ. Chọn 2-3 khuẩn lạc có hình thái giống nhau, hòa tan trong 500 µL nước MQ, sau đó li tâm thu cặn ở 8.000 vòng/phút trong 10 phút (lặp lại bước rửa tế bào với nước MQ 2 lần). Hòa cặn tế bào với 400 µL dung dịch li giải (100 mM Tris-HCl, 100 mM Na₂EDTA, 1,5 M NaCl, 1% CTAB, pH 8,0), rồi bổ sung 5 µL proteinase K (600 U/mL), trộn đều và ủ ở 37°C trong 1 giờ. Tiếp theo, thêm 120 µL SDS 20% (w/v) và 20 µL lysozyme (50 mg/mL) vào ống và ủ ở 65°C trong 2,5 giờ. Sau khi bổ sung vào ống 2 µL RNase A (10 mg/mL) và ủ ở 56°C trong 20 phút (lặp lại quá trình 2 lần), thêm tiếp vào ống 550 µL CI (chloroform:isoamyl alcohol = 49:1), trộn đều bằng cách đảo ngược ống 3-5 lần và li tâm 13.000 vòng/phút ở 4°C trong 10 phút. Thu dịch nổi sau li tâm (300 µL) vào một ống eppendorf mới, bổ sung 300 µL dung dịch CI và lặp lại quy trình đảo trộn, li tâm như bước trên. Sau đó, thu 100 µL dịch nổi vào ống eppendorf mới, bổ sung 60 µL isopropanol lạnh và kết tủa DNA qua đêm ở 4°C. Tủa DNA được li tâm thu hồi, rửa với ethanol 70% và làm khô ở 56°C. Hòa tan lại DNA tủa trong 100 µL nước MQ và kiểm tra nồng độ, độ tinh sạch bằng máy NanoDrop 2000.

- Giải trình tự hệ gene vi khuẩn: hệ gene của 2 chủng vi khuẩn *B. pseudomallei* sau khi tách chiết được lập thư viện với bộ kit Illumina DNA Prep (M)

Tagmentation 24 Samples (Illumina, Hoa Kỳ) và giải trình tự sử dụng hệ thống máy Illumina MiSeq (paired-end [PE]).

- Phân tích đặc tính sinh học: đánh giá chất lượng các đoạn read thu được sau giải trình tự bằng phần mềm FastQC 0.11.9 [3], làm sạch bằng phần mềm Trimmomatic 0.38.1 (sliding-window: 4:20, min-len: 20) [5]. Hệ gene được lắp ghép bằng phần mềm SPAdes 3.15.4 (-careful). Loại bỏ các contig có kích thước < 500 bp. Đánh giá chất lượng và độ hoàn thiện của hệ gene lắp ghép *de novo* bằng phần mềm BUSCO 5.4.6 [22] (lineage, *Burkholderiales*). Kiểm tra các thông số chất lượng hệ gene lắp ghép bằng phần mềm QUAST 5.2.0. Chú giải hệ gene bằng phần mềm Prokka 1.14.6 [19]. Xác định các gene kháng kháng sinh và gene độc lực căn cứ vào cơ sở dữ liệu ResFinder [7], NCBI Bacterial Antimicrobial Resistance Reference Gene Database [9] và VFDB [13] với công cụ ABRicate 1.0.1 [21]. Xác định kiểu trình tự (sequence type - ST) từ trình tự toàn bộ hệ gene và so sánh với cơ sở dữ liệu PubMLST bằng phần mềm MLST 2.22.0 [20]. Sử dụng phần mềm barnap 0.9 [18] để xác định trình tự gene ribosome 16S của các chủng từ dữ liệu lắp ghép hệ gene.

Bộ cơ sở dữ liệu bao gồm trình tự hệ gene 131 chủng vi khuẩn *B. pseudomallei* tải xuống từ cơ sở dữ liệu NCBI RefSeq (ngày 20/5/2023). Các gen lõi (core gene) có mặt ở > 99% số chủng trong bộ dữ liệu và có độ tương đồng > 92% được lựa chọn bằng phần mềm Roary 3.13.0 [15]. Xây dựng cây phân loại dựa trên thuật toán maximum-likelihood bằng phần mềm IQ-Tree 2.1.2 [14] với các gene lõi này. Từ cây phân loại tổng thể, các chủng có quan hệ gần gũi với chủng trong nghiên cứu này được lựa chọn để xây dựng cây phân loại cùng với các chủng *B. pseudomallei* có nguồn gốc từ Việt Nam [17], với nhóm ngoài (out-group) là chủng *B. pseudomallei* MSHR4503 có nguồn gốc từ Úc. Xây dựng cây phân loại dựa trên thuật toán neighbor-joining với 5.118

gene lõi, 1.000 bootstrap bằng phần mềm MEGA11 [12]. Danh sách 14 chủng *B. pseudomallei* sử dụng để xây dựng cây phân loại này được cung cấp bởi danh sách các chủng sử dụng trong xây dựng cây phân loại neighbor-joining.

3. KẾT QUẢ NGHIÊN CỨU VÀ BÀN LUẬN

3.1. Đánh giá chất lượng giải trình tự và chất lượng lắp ghép

Quy trình giải trình tự gene thể hệ mới thực hiện với vi khuẩn *B. pseudomallei* trên hệ thống giải trình tự MiSeq của Illumina bao gồm nhiều bước. Chất lượng mỗi bước đều ảnh hưởng đến chất lượng chung của việc giải trình tự. Do vậy, sau khi hoàn tất các bước thực nghiệm quy trình, chúng tôi tiến hành đánh giá chất lượng giải trình tự (chất lượng base, chất lượng adapter) bằng phần mềm FastQC. Kết quả đánh giá cho thấy, chất lượng giải trình tự ở cả hai mẫu vi khuẩn đều bảo đảm để có thể thực hiện các phân tích tiếp theo. Chất lượng này thể hiện ở các tiêu chí đạt được nêu ra trong báo cáo tiêu chuẩn của FastQC, gồm: không có trình tự chất lượng kém, điểm chất lượng Q đạt trên Q30 (độ chính xác của đọc trình tự > 99,9%), các adapter đã bị loại bỏ hoàn toàn.

Kiểm tra chất lượng lắp ghép hệ gene bằng phần mềm QUAST cho thấy không có sự bất cập nhầm (mismatches) trong quá trình lắp ghép, chỉ số N50 của 2 chủng thu được là 36.559 bp và 74.123 bp (các contigs có chiều dài này sẽ chiếm 50% bộ gene). Chỉ số N50 lớn hơn ở trình tự 22-P6H09-BPS-02 chứng tỏ quá trình giải trình tự đạt chất lượng tốt hơn do thu được nhiều contigs có chiều dài lớn hơn.

3.2. Phân tích dữ liệu hệ gene và định danh vi khuẩn

Việc phân tích dữ liệu hệ gene và định danh đối với vi khuẩn được thực hiện sau khi đánh giá chất lượng giải trình tự hoàn tất (kết quả thể hiện ở bảng 1).

Bảng 1. Kết quả phân tích dữ liệu hệ gene 2 chủng vi khuẩn *B. pseudomallei*.

Chủng	Mã trình tự	Kích thước hệ gene (bp) (%GC)	Kiểu trình tự (ST)	Số vùng mã hóa protein (CDS)	Số lượng ARN	Kích thước contigs dài nhất	N50	Kết quả định danh (gene 16S)
NTW 21.1	22-P6H09-BPS-01	6.993.710 (68,24%)	541	5.777	82	113.236	36.559	<i>B. pseudomallei</i> K96243
SS 1 240'	22-P6H09-BPS-02	6.992.098 (68,30%)	541	5.710	75	221.830	74.123	- Độ bao phủ: 100% - Độ tương đồng: 100%

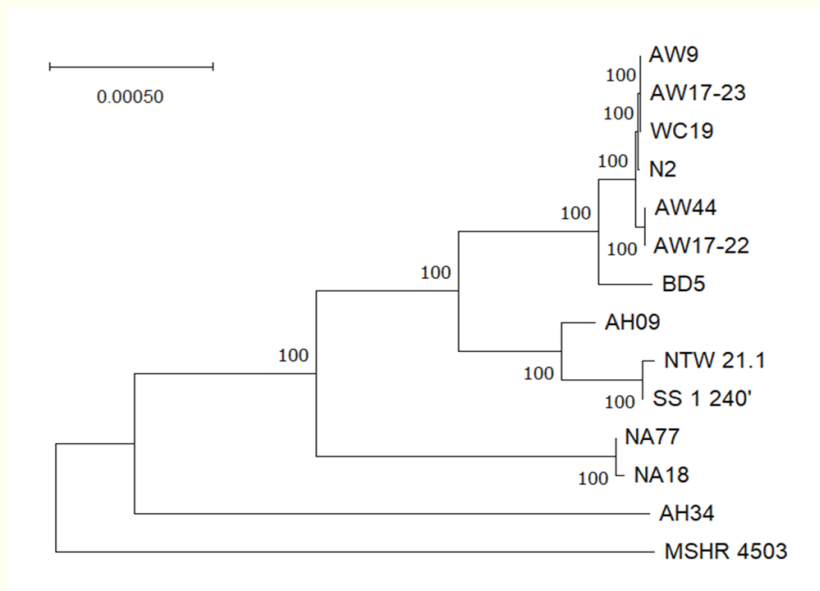
Kích thước hệ gene xác định được ở hai chủng lần lượt là 6.993.701 bp và 6.992.098 bp, với tỉ lệ GC tương ứng là 68,24% và 68,30%. Sử dụng phần mềm Blast [4] với hệ gene tham chiếu trên Genbank (K96243), trích đoạn gene mã hóa cho ribosome 16S (kích thước 1.527 bp) cho kết quả định danh là *B. pseudomallei* ở cả 2 chủng vi khuẩn với các chỉ số mức độ bao phủ (Query cover) và độ tương đồng (Per. Ident) đều đạt 100%. Xác định loại trình tự ST qua công cụ MLST là 541, với kiểu alen của các gene “keeping house” là *ace*: 3, *gltB*: 4, *gmhD*: 2, *lepA*: 3, *lipA*: 5, *narK*: 4 và *ndh*: 1 ở cả hai chủng vi khuẩn. Một số thông số như số vùng mã hóa protein, số lượng ARN, kích thước contigs dài nhất, chỉ số N50... cũng được xác định. Kết quả định danh và các thông số đã nêu khẳng định sự phù hợp với kết quả của nghiên cứu đã thực hiện trước đó [16]. Kết quả giải trình tự đã được đăng kí trên Genbank với mã số JASKZC000000000 và JASKZD000000000.

3.3. Xác định mối liên hệ với các chủng khác

Trên cơ sở kết quả trình tự thu được, chúng tôi xác định mối quan hệ di truyền giữa các chủng trong nghiên cứu này với các chủng từ các quốc gia khác, thông qua việc xây dựng cây phân loại.

Từ 131 hệ gene có nguồn gốc từ Lào, Thái Lan, Malaysia, Singapore, Hong Kong, Trung Quốc, Úc, thấy 2 chủng trong nghiên cứu này có quan hệ gần gũi nhất với 7 chủng có nguồn gốc từ Singapore (là WC19, BD5, N2, AW9, AW17-23, AW44, AW17-22). Cùng với các hệ gene của 4 chủng từ Việt Nam mới được công bố gần đây, chúng tôi xây dựng cây phân loại (hình 1).

Hai chủng trong nghiên cứu này có quan hệ gần gũi nhất với chủng AH09. Đây là chủng có cùng kiểu trình tự ST 541 và được phân lập từ ruộng mía ở Nghệ An, Việt Nam. Điều đó khẳng định độ tin cậy của kĩ thuật giải trình tự hệ gene vi khuẩn trong nghiên cứu này.



Hình 1. Cây phân loại xây dựng dựa trên 5.118 gene lõi của hai chủng *B. pseudomallei* NTW 21.1, SS 1 240' và các chủng có quan hệ gần gũi.

Việc xây dựng cây phân loại dựa trên trình tự toàn bộ hệ gene rất hữu ích trong nghiên cứu dịch tễ học phân tử, từ đó cung cấp thông tin nhằm truy vết nguồn gốc tác nhân gây bệnh, giúp cho các nhà điều tra, cán bộ y tế nắm bắt được tình hình khi có khủng bố sinh học xảy ra.

3.4. Phân tích xác định gene kháng kháng sinh và gene độc lực

Bên cạnh việc định danh vi sinh vật, việc thu thập dữ liệu trình tự hệ gene cũng cho phép chúng ta sử dụng để phân tích với nhiều mục đích khác nhau, trong đó có việc xác định các gene kháng

kháng sinh và các gene gây độc. Trong nghiên cứu này, chúng tôi tiến hành xác định các gene kháng kháng sinh của hai chủng nghiên cứu (kết quả phân tích thể hiện ở bảng 2).

Kết quả cho thấy cả hai chủng vi khuẩn đều có kiểu hình kháng beta-lactam. Chúng chứa các gene mã hóa cho enzyme phân giải kháng sinh nhóm beta-lactam là oxacillin-hydrolyzing class D beta-lactamase OXA-59 và PenI family class A extended-spectrum beta-lactamase. Các gene có độ tương đồng > 99% so với gene NG_049799.1 của chủng *B. pseudomallei* K96243 và gene NG_048031.1 của chủng *B. pseudomallei* 392a lưu giữ trên Genbank.

Bảng 2. Kết quả phân tích xác định gene kháng kháng sinh hai chủng nghiên cứu.

Chủng	Mã Trình tự	Tên gene	Độ bao phủ (%)	Độ tương đồng (%)	Mã số trên genbank	Protein sản phẩm	Dự đoán kiểu hình kháng
NTW 21.1	22-P6H09-BPS-01	blaOXA-59_1	100	99.88	NG_049799.1	oxacillin-hydrolyzing class D beta-lactamase OXA-59	BETA-LACTAM
		penI_Bp	100	99.32	NG_048031.1	PenI family class A extended-spectrum beta-lactamase	BETA-LACTAM
SS 1 240'	22-P6H09-BPS-02	blaOXA-59_1	100	99.88	NG_049799.1	oxacillin-hydrolyzing class D beta-lactamase OXA-59	BETA-LACTAM
		penI_Bp	100	99.32	NG_048031.1	PenI family class A extended-spectrum beta-lactamase	BETA-LACTAM

Bảng 3. Kết quả phân tích xác định các nhóm gene độc tố của hai chủng nghiên cứu.

Nhóm gene độc	Chủng <i>B. pseudomallei</i>	
	NTW 21.1	SS 1 240'
Di động nhờ actin (Actin-based motility)	1	1
Bám dính (Adherence)	13	13
Chống thực bào (Antiphagocytosis)	27	27
Xâm nhiễm (Invasion)	54	55
Hệ thống tiết (Secretion system)	53	55
Hô hấp kỵ khí (Anaerobic respiration)	1	1
Bơm đẩy (Efflux pump)	1	1
Nội độc tố (Endotoxin)	2	2
Yếu tố bám dính lông (Fimbrial adherence determinants)	1	1
Lẩn trốn miễn dịch (Immune evasion)	1	1
Hấp thu sắt (Iron uptake)	8	10
Độc tố (Toxin)	1	2
Khác	2	2

Sử dụng cơ sở dữ liệu VFDB, đã xác định được các gene độc lực và sản phẩm protein tương ứng của hai chủng nghiên cứu, gồm 165 gene độc lực ở chủng NTW 21.1 và 171 gene độc lực ở chủng SS 1 240'. Các gene độc của hai chủng này chỉ khác nhau ở nhóm xâm nhiễm, hệ thống tiết, hấp thu sắt và độc tố (bảng 3).

Phân tích, xác định các gene nêu trên ngoài việc góp phần xác định đặc điểm kháng kháng sinh, đặc điểm sinh độc tố của vi khuẩn để hỗ trợ cho truy vết

nguồn gốc, còn giúp cán bộ y tế định hướng phác đồ điều trị thích hợp khi có tình huống nhiễm độc xảy ra và kiểm soát sự lây lan của vi khuẩn kháng thuốc trong cộng đồng.

4. KẾT LUẬN

Chúng tôi đã nghiên cứu ứng dụng thành công công nghệ giải trình tự thế hệ mới trong giải trình tự toàn bộ hệ gene vi khuẩn *B. pseudomallei* tại phòng thí nghiệm (Viện Khoa học và Công nghệ, Bộ Công an). Bằng việc sử dụng cụ phần mềm phân tích, chúng tôi xác nhận tính chính xác trong định danh vi khuẩn bằng kĩ thuật sinh học phân tử và tính chính xác về kiểu trình tự ST. Bên cạnh đó, chúng tôi cũng đã phân tích được các gene kháng kháng sinh và gene độc lực của vi khuẩn.

TÀI LIỆU THAM KHẢO

- Hoàng Việt Hà, Bùi Nguyễn Hải Linh, Trần Thị Lệ Quyên, Phạm Thị Huyền, Hoàng Quang Trung, Trần Anh Đào, Nguyễn Vũ Trung, Trịnh Thành Trung (2020), "So sánh các kĩ thuật định danh *B. pseudomallei* trong xét nghiệm chẩn đoán bệnh Whitmore", *Tạp chí Khoa học và Công nghệ Việt Nam*, số 62(5), 5/2020.
- Alex R Hoffmaster, David AuCoin, Prasith Baccam, Henry C Baggett, Rob Baird, Saithip Bhengsri, David D Blaney, Paul J Brett, Timothy J.G Brooks, Katherine A Brown, Narisara Chantratita, Allen C Cheng, David A.B Dance, Saskia Decuypere, Dawn Defenbaugh, Jay E Gee, Raymond Houghton, Possawat Jorakate, Ganjana Lertmemongkolchai, Direk Limmathurotsakul, et al. (2015) "Melioidosis Diagnostic Workshop 2013", *Emerg Infect Dis.*; 21(2): e141045. doi: 10.3201/eid2102.141045.
- Andrews S (n.d), *FastQC A Quality Control tool for High Throughput Sequence Data*. www.bioinformatics.babraham.ac.uk/projects/fastqc/.

4. Bankevich A, Nurk S, Antipov D, Gurevich A.A, Dvorkin M, Kulikov A.S, Lesin V.M, Nikolenko S.I, Pham S, Pribelski A.D, Pyshkin A.V, Sirotkin A.V, Vyahhi N, Tesler G, Alekseyev M.A, Pevzner P.A (2012), "SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing", *Journal of Computational Biology*, 19(5), 455-477. <https://doi.org/10.1089/cmb.2012.0021>.
5. Bolger A.M, Lohse M, Usade B (2014), "Trimmomatic: a flexible trimmer for Illumina sequence data", *Bioinformatics*, 30(15), 2114-20. <https://doi.org/10.1093/bioinformatics/btu170>.
6. Direk Limmathurotsakul, Nick Golding, David Ab Dance, Jane P Messina, David M Pigott, Catherine L Moyes, Dionne B Rolim, Eric Bertherat, Nicholas Pj Day, Sharon J Peacock, Simon I Hay (2016), "Predicted global distribution of *B. pseudomallei* and burden of melioidosis". *Nat Microbiol.*; 1(1):15008. doi: 10.1038/nmicrobiol.2015.8.
7. Ea Zankari, Henrik Hasman, Salvatore Cosentino, Martin Vestergaard, Simon Rasmussen, Ole Lund, Frank M Aarestrup, Mette Voldby Larsen (2012), "Identification of acquired antimicrobial resistance genes", *J Antimicrob Chemother.* 67(11): 2640-4. doi: 10.1093/jac/dks261. Epub 2012 Jul 10.
8. Esther M Gabor, Erik J de Vries, Dick B Janssen (2003), "Efficient recovery of environmental DNA for expression cloning by indirect extraction methods", *FEMS Microbiol Ecol.*; 44(2):153-63. doi: 10.1016/S0168-6496(02)00462-2.
9. Feldgarden M, Brover V, Gonzalez-Escalona N, Frye J.G, Haendiges J, Haft D.H, Hoffmann M, Pettengill J.B, Prasad A.B, Tillman G.E, Tyson G.H, Klimke W, "AMR FinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence", *Sci Rep.* 2021 Jun 16; 11(1):12728. doi: 10.1038/s41598-021-91456-0. PMID: 34135355; PMCID: PMC8208984.
10. Hara Y, Mohamed R, Nathan S, et al. (2013), "Epidemiology of melioidosis in Asia and the Pacific: a systematic review", *PloS Neglected Tropical Diseases* 7(6): e2455. <https://doi.org/10.1371/journal.pntd.0002455>.
11. <http://www.bt.cdc.gov/agent/agentlist-category.asp>.
12. Koichiro Tamura, Glen Stecher, Sudhir Kumar (2021), "MEGA11: Molecular Evolutionary Genetics Analysis version 11", *Molecular Biology and Evolution*, 38:3022-3027.
13. Liu B, Zheng D, Jin Q, Chen L, Yang J (2019), "VFDB 2019: a comparative pathogen-omic platform with an interactive web interface", *Nucleic Acids Research.* 47, D1, D687-D692. <https://doi.org/10.1093/nar/gky1080>.
14. Nguyen L.T, Schmidt H.A, von Haeseler A, Minh B.Q (2014), "IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies", *Molecular Biology and Evolution*, 32(1), 268-274. <https://doi.org/10.1093/molbev/msu300>.
15. Page A.J, Cummins C.A, Hunt M, Wong V.K, Reuter S, Holden M.T.G, Fookes M, Falush D, Keane J.A, Parkhill J (2015), "Roary: rapid large-scale prokaryote pan genome analysis". *Bioinformatics*, 31(22), 3691-3693. <https://doi.org/10.1093/bioinformatics/btv421>.
16. Quyen T.L Tran, Phuc H Phan, Linh N.H Bui, Ha T.V Bui, Ngoc T.B Hoang, Dien M Tran, Trung T Trinh (2022), "Child Melioidosis Deaths Caused by *B. pseudomallei* Contaminated Borehole Water, Vietnam, 2019", *Emerg Infect Dis.*; 28(8): 1689-1693. doi: 10.3201/eid2808.220113.
17. Sabine Lichtenegger, Trung T Trinh, Karoline Assig, Karola Prior, Dag Harmsen, Julian Pesl, Andrea Zauner, Michaela Lipp, Tram A Que, Beatrice Mutsam, Barbara Kleinhappl, Ivo Steinmetz, Gabriel E Wagner (2021), "Development and Validation of a *B. pseudomallei* core genome multilocus sequence typing scheme to facilitate molecular surveillance". *J Clin Microbiol.*; 59(8): e00093-21. doi: 10.1128/JCM.00093-21.
18. Seemann T (2013), *Barnap 0.8: rapid ribosomal RNA prediction*.
19. Seemann T (2014) "Prokka: rapid proka-ryotic genome annotation", *Bioinformatics*, 30(14), 2068-2069, doi.org/10.1093/bioinformatics/btu153.
20. Seemann T (2016), "MLST: Scan contig files against PubMLST typing schemes. <https://github.com/tseemann/mlst>.
21. Seemann T (2016), *ABRicate: mass screening of contigs for antibiotic resistance*.
22. Simão F.A, Waterhouse R.M, Ioannidis P, Kriventseva E.V, Zdobnov E.M. (2015), "BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs", *Bioinformatics*, 31(19), 3210-3212. doi.org/10.1093/bioinformatics/btv351. □