

Nhân dòng và phân tích trình tự gen mã hoá dị nguyên Der f 1 từ mạt bụi nhà *Dermatophagoides farinae* trong *Escherichia coli*

Trần Thị Quỳnh Trang¹, Vũ Mai Phương¹, Nguyễn Thị Linh Chi¹, Nguyễn Thị Phương Liên¹, Nguyễn Quang Huy², Nguyễn Thị Anh Đào¹, Nguyễn Duy Hà³, Nguyễn Quỳnh Uyên¹, Hoàng Văn Vinh^{1*}

¹Viện Vi sinh vật và Công nghệ Sinh học, Đại học Quốc gia Hà Nội, 144 Xuân Thủy, phường Cầu Giấy, Hà Nội, Việt Nam

²Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội, 334 Nguyễn Trãi, phường Thanh Xuân, Hà Nội, Việt Nam

³Viện Y học Dự phòng Quân đội, 21 Trung Liệt, phường Đồng Đa, Hà Nội, Việt Nam

Ngày nhận bài 22/11/2024; ngày chuyển phân biện 25/11/2024; ngày nhận phân biện 20/12/2024; ngày chấp nhận đăng 25/12/2024

Tóm tắt:

Kháng nguyên Der f 1 của *Dermatophagoides farinae* (*D. farinae*) là một trong những tác nhân chính gây dị ứng mạt bụi nhà. Mục tiêu của nghiên cứu là tách dòng gen mã hóa kháng nguyên Der f 1 từ *D. farinae* được thu thập tại Việt Nam làm cơ sở cho việc phát triển protein Der f 1 tái tổ hợp. Sản phẩm khuếch đại toàn bộ trình tự mã hóa cho Der f 1 (thực hiện bằng phản ứng PCR 3 bước, sử dụng cDNA tổng hợp từ RNA của *D. farinae*) được nhân dòng trong vector pGEM-T và biến nạp vào vi khuẩn *Escherichia coli* DH5α. Plasmid tái tổ hợp được sàng lọc bằng PCR khuẩn lạc và giải trình tự theo phương pháp Sanger. Trình tự gen Der f 1 và amino acid suy diễn tương ứng có độ tương đồng 99% so với các trình tự nucleotide và amino acid tham chiếu trên GenBank (mã số lần lượt là AB034946.1 và ABL84751.1). Trình tự gen Der f 1 từ *D. farinae* thu thập tại Việt Nam đã được đăng ký trên GenBank với mã số PP417680.1. Kết quả nghiên cứu cung cấp dữ liệu nền tảng quan trọng cho các nghiên cứu về protein Der f 1 tái tổ hợp, phục vụ cho chẩn đoán và điều trị dị ứng mạt bụi nhà bằng liệu pháp miễn dịch đặc hiệu tại Việt Nam.

Từ khóa: *Dermatophagoides farinae*, dị nguyên Der f 1, nhân dòng, PCR.

Chỉ số phân loại: 1.6, 2.6, 3.5

Cloning and analysis of the antigen-coding gene Der f 1 from house dust mite *Dermatophagoides farinae* in *Escherichia coli*

Thi Quynh Trang Tran¹, Mai Phuong Vu¹, Thi Linh Chi Nguyen¹, Thi Phuong Lien Nguyen¹, Quang Huy Nguyen², Thi Anh Dao Nguyen¹, Duy Ha Nguyen³, Quynh Uyen Nguyen¹, Van Vinh Hoang^{1*}

¹Institute of Microbiology and BioTechnology, Vietnam National University - Hanoi, 144 Xuan Thuy Street, Cau Giay Ward, Hanoi, Vietnam

²Hanoi University of Science, Vietnam National University - Hanoi, 334 Nguyen Trai Street, Thanh Xuan Ward, Hanoi, Vietnam

³Military Institute of Preventive Medicine, 21 Trung Liet Street, Dong Da Ward, Hanoi, Vietnam

Received 22 November 2024; revised 20 December 2024; accepted 25 December 2024

Abstract:

Der f 1, an allergen from *Dermatophagoides farinae* (*D. farinae*), is one of the major causes of house dust mite (HDM) allergy. The aim of this study was to clone the gene encoding the Der f 1 allergen from *D. farinae* collected in Vietnam, providing a basis for the development of recombinant Der f 1 protein. The amplified product containing the full-length coding sequence of Der f 1 (obtained through a three-step PCR using cDNA synthesised from the RNA of *D. farinae*) was cloned into the pGEM-T vector and transformed into *Escherichia coli* DH5α. The recombinant plasmids were screened by colony PCR and sequenced using the Sanger method. The sequence of Der f 1 and its deduced amino acid had 99% similarities to the reference nucleotide and amino acid sequences in GenBank, with the accession number of AB034946.1 and ABL84751.1, respectively. The Der f 1 gene sequence from *D. farinae* collected in Vietnam has been deposited in GenBank under the accession number PP417680.1. The results of the study provide essential foundational data for further research on recombinant Der f 1 protein, supporting the diagnosis and treatment of house dust mite allergy through specific immunotherapy in Vietnam.

Keywords: cloning, Der f 1 allergen, *Dermatophagoides farinae*, PCR.

Classification numbers: 1.6, 2.6, 3.5

*Tác giả liên hệ: Email: vinh.imbt@gmail.com

1. Đặt vấn đề

Mạt bụi nhà (MBN) được xem là một trong những nguồn dị nguyên chính trong nhà, gây ra hơn 50% tổng số bệnh nhân bị dị ứng. *D. farinae*, một trong ba loài MBN chính gây dị ứng, thường gặp ở các nước có khí hậu khô lạnh [1]. Tuy nhiên, theo một nghiên cứu mới của D. Kim và cs (2024) [2], tỷ lệ phát hiện *D. farinae* trong bụi đệm giường tại Việt Nam lên đến 97%. Một số nghiên cứu khác cho thấy, tại miền Nam Việt Nam, tỷ lệ bệnh nhân bị dị ứng đường hô hấp có test lấy da (Skin Prick Test - SPT) dương tính và bị dị ứng với *D. farinae* là cao nhất (lần lượt là 59,8 và 54,3%) [3, 4]. Trong số 37 dị nguyên đã được xác định của *D. farinae*, các nhóm dị nguyên 1, 2 và 23 được xác định là dị nguyên chính do có khả năng liên kết cao với IgE và hoạt tính kháng nguyên mạnh [5]. Dị nguyên Der f 1 được xác định là dị nguyên chính (thuộc nhóm 1) và có hoạt tính cysteine protease. Trình tự nucleotide tại vùng biến đổi của gen mã hóa cho Der f 1 có sự đa dạng giữa các chủng *D. farinae* thu thập từ các khu vực khác nhau trên thế giới [6].

Các bệnh dị ứng do MBN nói chung và *D. farinae* nói riêng có thể gây ra các vấn đề hô hấp nghiêm trọng như viêm mũi dị ứng và hen suyễn, làm giảm chất lượng cuộc sống của bệnh nhân. Gần đây, việc ứng dụng các protein tái tổ hợp trong liệu pháp miễn dịch đặc hiệu chẩn đoán và điều trị dị ứng đã cho thấy những ưu điểm vượt trội (số lượng, độ tinh khiết, hoạt tính kháng nguyên, độ ổn định giữa các mẻ sản xuất và giảm nguy cơ gây phản ứng phụ) so với dịch chiết thô [7]. Các hệ tế bào chủ biểu hiện như *E. coli*, nấm men, hệ tế bào baculovirus-insect (baculovirus-insect cell system), hệ tế bào động vật có vú và tế bào thực vật đã được phát triển và sử dụng với mục đích biểu hiện các protein tái tổ hợp này [8, 9]. Hệ biểu hiện *E. coli* là hệ biểu hiện phổ biến nhất với nhiều ưu điểm như sự phát triển nhanh của tế bào chủ, thời gian sinh trưởng ngắn, dễ thao tác, hiệu suất sản xuất protein tái tổ hợp cao và giá thành thấp [10].

Tại Việt Nam, mặc dù tỷ lệ bệnh nhân dị ứng MBN nói chung và dị ứng với *D. farinae* nói riêng là khá cao, nhưng các nghiên cứu về vấn đề này vẫn chủ yếu tập trung vào điều tra, khảo sát tỷ lệ mắc cảm với từng loài dựa trên kết quả test lấy da và xét nghiệm IgE đặc hiệu [3, 4]. Ngoài ra, nguồn nguyên liệu dị nguyên sử dụng trong các xét nghiệm dị ứng vẫn phụ thuộc hoàn toàn vào nguồn dị nguyên nhập khẩu từ nước ngoài, dẫn đến một số nhược điểm về chi phí, thời gian, thủ tục, chất lượng và tính đặc trưng của dị nguyên. Do vậy, nghiên cứu này được thực hiện với mục đích tách dòng gen mã hóa cho kháng nguyên Der f 1 của *D. farinae* thu thập tại Việt Nam để làm nền tảng cho việc biểu hiện protein tái tổ hợp Der f 1, hướng tới việc phát triển các bộ kit chẩn đoán và liệu pháp miễn dịch đặc hiệu, phù hợp với nhu cầu trong nước cũng như giảm sự phụ thuộc vào nguồn nguyên liệu nhập khẩu.

2. Nguyên liệu và phương pháp nghiên cứu

2.1. Nguyên liệu

Mẫu MBN *D. farinae* được thu nhận, định danh bằng phương pháp hình thái học và sinh học phân tử, lưu giữ tại -80°C [11]. Chủng vi khuẩn *E. coli* DH5α được cung cấp bởi hãng Fermentas (Thermo Fisher Scientific, Mỹ).

Vector tách dòng pGEM-T do Promega cung cấp và các cặp mồi được tổng hợp bởi hãng IDT (Mỹ). Các hóa chất và bộ kit đạt tiêu chuẩn trong lĩnh vực sinh học phân tử đã được sử dụng trong nghiên cứu.

Các hoá chất sử dụng trong nghiên cứu như Master Mix (Promega), Agarose (Promega), RedSafe (Intron)... đều có độ tinh khiết cao, đạt tiêu chuẩn sử dụng trong nghiên cứu sinh học phân tử.

2.2. Phương pháp nghiên cứu

Tách chiết RNA và tổng hợp cDNA [12]:

Quá trình tách chiết RNA tổng số từ mẫu *D. farinae* sử dụng bộ kit SV Total RNA Isolation System (Promega) được bổ sung thêm bước xử lý RNA với RQ1 RNase free DNase nhằm đảm bảo xử lý hoàn toàn DNA và thực hiện theo hướng dẫn của nhà sản xuất. Nồng độ và độ tinh khiết (tỷ lệ $OD_{260\text{ nm}}/OD_{280\text{ nm}}$) của RNA tổng số được xác định bằng máy Nanodrop (ThermoFisher). Nồng độ RNA trong mẫu được xác định theo công thức sau:

$$\text{Nồng độ RNA } (\mu\text{g/ml}) = 40 \times OD_{260\text{ nm}} \times n$$

trong đó: $OD_{260\text{ nm}}$ là độ hấp thụ của dung dịch RNA ở 260 nm; 40 là nồng độ RNA ứng với $OD_{260\text{ nm}}=1$; n là hệ số pha loãng.

Sản phẩm cDNA thu nhận sau khi sử dụng RNA tổng số làm khuôn cho quá trình phiên mã ngược (bằng bộ kit SuperScript III Reverse Transcriptase, Invitrogen) được bảo quản ở -20°C để sử dụng cho các nghiên cứu sau.

Thiết kế các cặp mồi:

Hai cặp mồi được thiết kế sử dụng phần mềm Primer-BLAST dựa trên vùng bảo thủ của 32 trình tự gen mã hóa cho Der f 1 của các chủng *D. farinae* được công bố trên cơ sở dữ liệu GenBank là AB034946.1, DQ185509.1, EF139428.1, EF139429.1, EF139430.1, X65196.1, HQ009399.1, EU095368.1, AY283290.1, JN222804.1, KJ542085.1, KJ542084.1, KJ542083.1, KJ542082.1, KJ542081.1, KJ542080.1, KJ542079.1, KJ542078.1, KJ542077.1, KJ542076.1, KJ542075.1, KJ542074.1, KJ542073.1, KJ542072.1, KJ542071.1, KJ542070.1, KJ542069.1, KJ542068.1, KJ542067.1, KJ542066.1, KJ542065.1 và KJ542064.1.

Phản ứng PCR và điện di kiểm tra sản phẩm PCR [13]:

Phản ứng PCR được thực hiện với các cặp mồi F1/R2, F3/R3 và M13 F/M13 R, có thành phần là H₂O: 6,2 µl, GoTaq G2 Hot Start Green Mastermix, 2X: 10,0 µl, mồi F 10 µM: 0,4 µl, mồi R 10 µM: 0,4 µl, DNA khuôn: 3,0 µl và chương trình là: 95°C - 2 phút, (95°C - 30 giây, 59°C - 30 giây, 72°C - 60 giây) x 45 chu kỳ, 72°C - 5 phút. PCR khuôn lạc cũng được tiến hành tương tự như trên với DNA khuôn là dịch ly giải các khuẩn lạc đơn mọc trên môi trường LB có ampicillin với cặp mồi F1/R2 và cặp mồi M13 F/M13 R ở nhiệt độ gần mồi 60°C. Phản ứng PCR gộp mở rộng (Overlap Extension PCR) để tạo khuôn cho phản ứng khuếch đại toàn bộ gen mã hóa Der f 1, được thực hiện bằng cách trộn các sản phẩm PCR của cặp mồi F1/R1; F2/R2 và sử dụng chu trình nhiệt như sau: 95°C - 2 phút, (95°C - 30 giây, 59°C - 60 giây, 72°C - 60 giây) x 40 chu kỳ, 72°C - 5 phút. Các sản phẩm PCR được điện di trên gel agarose 1%, trong đệm TBE 1X và chụp ảnh bằng hệ thống soi gel ImageQuant™ LAS 500 (GE Healthcare, Mỹ).

Tách dòng sản phẩm PCR trong vector pGEM-T [14]:

Sản phẩm PCR được tách dòng trong vector pGEM-T theo hướng dẫn của bộ kit pGEM-T Vector Systems (Promega). Các khuẩn lạc đơn trên đĩa LB có ampicillin (50 µg/ml) được sàng lọc bằng PCR khuẩn lạc (PCR colony) với cặp mồi F1/R2 và M13 F/M13 R. Plasmid từ dịch nuôi cấy các khuẩn lạc (trong môi trường LB lỏng có bổ sung 50 µg/ml ampicillin) cho kết quả PCR dương tính với cả 2 cặp mồi trên, được tách bằng kit ZR Plasmid Miniprep (Zymo Research), rồi được giải trình tự theo phương pháp Sanger (Công ty DNA sequencing, Việt Nam).

So sánh trình tự nucleotide và trình tự amino acid:

Trình tự amino acid suy diễn từ trình tự nucleotide được xác định bằng công cụ Translate Tool trên ExPASy. Các trình tự amino acid và trình tự nucleotide được tải lên GenBank, thực hiện các bước so sánh với các trình tự đã được công bố tương ứng.

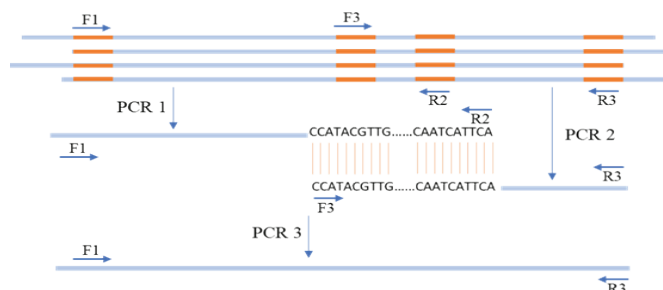
3. Kết quả

3.1. Thiết kế mồi cho phản ứng PCR khuếch đại gen mã hóa Der f 1

Dựa trên các vùng bảo thủ của 32 trình tự cDNA mã hóa cho kháng nguyên Der f 1 của *D. farinae*, trình tự 2 cặp mồi sử dụng trong phản ứng khuếch đại gen Der f 1 được chúng tôi thiết kế là F1: 5' TTCATCAAAAATGAAATTCGTTTGGCC 3' (có chiều dài là 28 nucleotide), R2: 5' TGAATGATTGTTTCGTCCATCATAATG 3' (có chiều dài là 26 nucleotide), F3: 5' CCATACGTTGCACGAGAACAAC 3' (có chiều dài là 22 nucleotide), và R3: 5' CAAATGTTACATGATTACAACATATGG 3' (có chiều dài là 28 nucleotide). Kích thước sản phẩm PCR của 2 cặp mồi F1/R2 và F3/R3 theo tính toán lý thuyết lần lượt là 786 và 394 bp, với 197 bp gộp nhau.

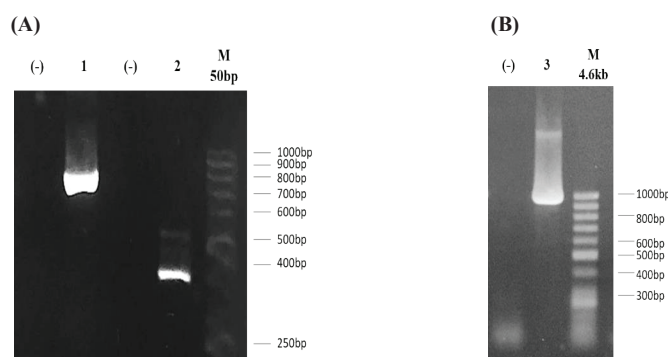
3.2. Lắp ráp cDNA hoàn chỉnh mã hóa Der f 1

RNA tổng số sau tách chiết đáp ứng yêu cầu về nồng độ (2806,90 ng/µl) và độ sạch (tỷ lệ OD_{260 nm}/OD_{280 nm}=1,83), được sử dụng để tạo cDNA, và cDNA này được sử dụng làm khuôn để khuếch đại toàn bộ trình tự mã hóa cho Der f 1 qua phản ứng PCR 3 bước liên tiếp (hình 1).



Hình 1. Sơ đồ thiết kế các mồi và phản ứng PCR 3 bước liên tiếp nhằm khuếch đại toàn bộ gen mã hóa cho Der f 1. PCR 1, PCR 2: sử dụng lần lượt cặp mồi F1/R2, F3/R3 và cDNA của *D. farinae* làm khuôn; PCR 3: sử dụng cặp mồi F1/R3 và khuôn là sản phẩm thu được khi thực hiện phản ứng nối kéo dài của PCR 1 và PCR 2; vùng màu vàng: trình tự vùng nucleotide bảo thủ của 32 trình tự mã hóa cho Der f 1.

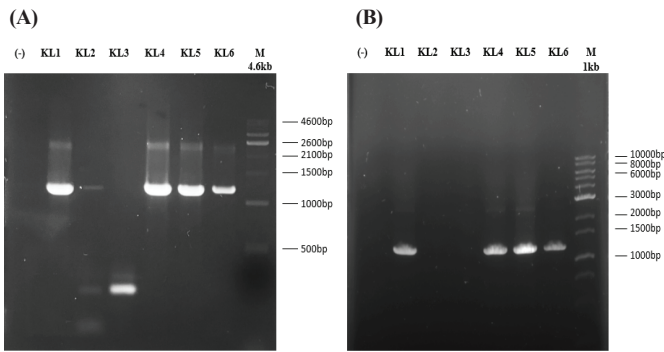
Kết quả các phản ứng PCR 1 và PCR 2 (lần lượt sử dụng cặp mồi F1/R2, F3/R3) cho thấy tính đặc hiệu (1 băng duy nhất) với kích thước phù hợp theo tính toán lý thuyết của 2 đoạn cDNA vùng 5' và 3' của gen Der f 1 (hình 2A). Sản phẩm PCR 3 cũng thu được 1 băng đậm với kích thước khoảng 1000 bp, phù hợp với tính toán theo lý thuyết của cDNA hoàn chỉnh mã hóa cho Der f 1 (hình 2B).



Hình 2. Sản phẩm PCR sử dụng cặp mồi F1/R2, F3/R3 (hình 2A) và F1/R3 (hình 2B). 1, 2: sản phẩm PCR 1 và PCR 2 lần lượt sử dụng cặp mồi F1/R2, F3/R3 và cDNA của *D. farinae* làm khuôn; 3: sản phẩm PCR 3 sử dụng cặp mồi F1/R3 và khuôn là sản phẩm thu được khi thực hiện phản ứng nối kéo dài của PCR 1 và PCR 2; (-): đối chứng âm sử dụng nước làm khuôn cho phản ứng PCR; M: thang chuẩn DNA.

3.3. Tách dòng cDNA mã hóa cho Der f 1 trong vector pGEM - T

Sau khi được thôi gel bằng kit Wizard SV Gel and Clean up system (Promega) sản phẩm PCR khuếch đại cDNA hoàn chỉnh mã hóa cho Der f 1 (kích thước khoảng 1000 bp), được đưa vào vector pGEM - T theo hướng dẫn của nhà sản xuất và biến nạp vào tế bào khả biến của *E. coli* DH5α. Các khuẩn lạc đơn trên đĩa LB có ampicillin được tiến hành sàng lọc bằng PCR khuẩn lạc với cặp mồi F1/R3 và M13 F/M13 R (hình 3).



Hình 3. Sản phẩm PCR khuẩn lạc với cặp mồi F1/R3 (hình 3A) và cặp mồi M13 F/M13 R (hình 3B) của các khuẩn lạc sàng lọc. KL1, KL2, KL3, KL4, KL5, KL6: sản phẩm PCR sử dụng khuôn là DNA tổng số của các khuẩn lạc đơn tương ứng trên môi trường LB có ampicillin; (-): đối chứng âm sử dụng nước làm khuôn cho phản ứng PCR; M: thang chuẩn DNA.

Dermatophagoides farinae mRNA for Der f 1 allergen preproenzyme, complete cds
Sequence ID: [AB034946.1](#) Length: 1081 Number of Matches: 1

Score	Expect	Method	Identities	Gaps	Strand
1786 bits(967)	0.0	Compositional matrix adjust.	973/976(99%)	0/976(0%)	Plus/Plus
Query 1	TT CATCAAAATGAAATCGTTTGGCCATTGCTCTTTGGTATTGAGCACTGTTTA		60		
Sbjct 34	TT CATCAAAATGAAATCGTTTGGCCATTGCTCTTTGGTATTGAGCACTGTTTA		93		
Query 61	TGCTCGCCAGCTTCAATCAAACTTTTGAAGAAATCAAAAAGCCTTCAACAAAACTA		120		
Sbjct 94	TGCTCGCCAGCTTCAATCAAACTTTTGAAGAAATCAAAAAGCCTTCAACAAAACTA		153		
Query 121	TGCCACGTTGAAGAGGAAGGTTGCCCGTAAAAAATTTTGGAAATCATGAAATATGT		180		
Sbjct 154	TGCCACGTTGAAGAGGAAGGTTGCCCGTAAAAAATTTTGGAAATCATGAAATATGT		213		
Query 181	TGAAGCTAACAAAGGTGCTCAACCAATTTGTCGATTGTGATGGATGAATCAAAAA		240		
Sbjct 214	TGAAGCTAACAAAGGTGCTCAACCAATTTGTCGATTGTGATGGATGAATCAAAAA		273		
Query 241	CCGTTATTTGATGAGTGTGAGGCTTTTGAACAACTCAAACTCAATTCGATTGAATGC		300		
Sbjct 274	CCGTTATTTGATGAGTGTGAGGCTTTTGAACAACTCAAACTCAATTCGATTGAATGC		333		
Query 301	CGAAACAAAGCGTTCGGTATCAATTCGGTTAACTTCATCGGAATGGATTACGATC		360		
Sbjct 334	CGAAACAAAGCGTTCGGTATCAATTCGGTTAACTTCATCGGAATGGATTACGATC		393		
Query 361	ACTGCGAATGTCACCTCAATCGTATGCGAGGAGGCTGTGGTTCATGTTGGGCTTTCTC		420		
Sbjct 394	ACTGCGAATGTCACCTCAATCGTATGCGAGGAGGCTGTGGTTCATGTTGGGCTTTCTC		453		
Query 421	TGGTGTGCGCCCACTGAATCAGCTTATTTGGCTACCGTAACACGCTTTGGATCTTTT		480		
Sbjct 454	TGGTGTGCGCCCACTGAATCAGCTTATTTGGCTACCGTAACACGCTTTGGATCTTTT		513		
Query 481	TGAACAGGAATCGTGCATTCGCGATCTCAACACGATGTCACGGGATACAAATCAAG		540		
Sbjct 514	TGAACAGGAATCGTGCATTCGCGATCTCAACACGATGTCACGGGATACAAATCAAG		573		
Query 541	AGGCATCGAATACATCAACAAATGGTGTGTTGAAGAAAGAGCTATCCATACGTTGC		600		
Sbjct 574	AGGCATCGAATACATCAACAAATGGTGTGTTGAAGAAAGAGCTATCCATACGTTGC		633		
Query 601	ACGAGAACAAATGCGGACGACCAATTCGCAACATTACGGTATCTCAAACTACTGCCA		660		
Sbjct 634	ACGAGAACAAATGCGGACGACCAATTCGCAACATTACGGTATCTCAAACTACTGCCA		693		
Query 661	AGTTATCCACAGATGTGAACAAATCGTGAAGCTTTGACTCAACACACACAGCTAT		720		
Sbjct 694	AGTTATCCACAGATGTGAACAAATCGTGAAGCTTTGACTCAACACACACAGCTAT		753		
Query 721	TGCGGTCTATTATGGCATTAAAGATTTGAGAGCTTTTCAACATTATGATGGACGAACAT		780		
Sbjct 754	TGCGGTCTATTATGGCATTAAAGATTTGAGAGCTTTTCAACATTATGATGGACGAACAT		813		
Query 781	CATTGACATGACATGGTTATCAACAACTATCATGCGGTCAACATTGTCGGTTACGG		840		
Sbjct 814	CATTGACATGACATGGTTATCAACAACTATCATGCGGTCAACATTGTCGGTTACGG		873		
Query 841	AAGTACAAAGCGGCTGATTATTTGGATGCTACGAACAGTTGGGATACCTACCTGGGGTGA		900		
Sbjct 874	AAGTACAAAGCGGCTGATTATTTGGATGCTACGAACAGTTGGGATACCTACCTGGGGTGA		933		
Query 901	TAGCGGATACGGATATTTTCAAGCGGAAACAACTCATGATGATGAAATATCCATA		960		
Sbjct 934	TAGCGGATACGGATATTTTCAAGCGGAAACAACTCATGATGATGAAATATCCATA		993		
Query 961	TGTTGTAATCATGTGA 976				
Sbjct 994	TGTTGTAATCATGTGA 1009				

Hình 4. So sánh trình tự nucleotide của gen Der f 1 được tách dòng và trình tự Der f 1 (mã số AB034946.1) đã được công bố trên GenBank.

Kết quả PCR cho thấy, khuẩn lạc số 1 (KL1), số 4 (KL4), số 5 (KL5) và số 6 (KL6) cho kết quả phản ứng PCR dương tính với đồng thời 2 cặp mồi F1/R3 (hình 3A) và M13 F/M13 R (hình 3B). Khuẩn lạc số 5 được chúng tôi lựa chọn để nuôi cấy, tách plasmid bằng ZR Plasmid Miniprep kit (Zymo Research) theo hướng dẫn của nhà sản xuất để kiểm tra trình tự gen mã hóa cho Der f 1 theo phương pháp Sanger.

3.4. So sánh trình tự cDNA mã hóa cho Der f 1 được tách dòng trong vector pGEM - T và trình tự amino acid suy diễn của chúng

Kết quả so sánh trình tự cDNA mã hóa cho Der f 1 trong nghiên cứu của chúng tôi với trình tự tham chiếu mã hóa cho Der f 1 (mã số AB034946.1) trên GenBank cho thấy, có khác biệt ở 3 vị trí (đánh dấu đỏ) và có 99% tương đồng (hình 4). Trình tự amino acid suy diễn của Der f 1 tách dòng của chúng tôi chỉ có 1 vị trí sai khác (đánh dấu đỏ) và cũng có 99% tương đồng với trình tự amino acid tham chiếu của Der f 1 (mã số ABL84751.1) đã đăng ký trên GenBank (hình 5).

Der f 1 allergen [Dermatophagoides farinae]

Sequence ID: [ABL84751.1](#) Length: 321 Number of Matches: 1

Score	Expect	Method	Identities	Positives	Gaps
674 bits(1740)	0.0	Compositional matrix adjust.	320/321(99%)	321/321(100%)	0/321(0%)
Query 1	MKFLAIASLLVSTVYARPAKSTFEFFKKAFKNKYATVEEEVARKNFLESKYVEAN		60		
Sbjct 1	MKFLAIASLLVSTVYARPAKSTFEFFKKAFKNKYATVEEEVARKNFLESKYVEAN		60		
Query 61	KGAINHLSDSLDEFKNRYLMSAEAFQKTFDLNAETSAKRINSVWVPSLDLRLSRT		120		
Sbjct 61	KGAINHLSDSLDEFKNRYLMSAEAFQKTFDLNAETSAKRINSVWVPSLDLRLSRT		120		
Query 121	VTPIRMGGGCGSFAFSGVAATESAYLAYRNTSLDSEQLVDCASQHGHDITPRGIE		180		
Sbjct 121	VTPIRMGGGCGSFAFSGVAATESAYLAYRNTSLDSEQLVDCASQHGHDITPRGIE		180		
Query 181	YIQNGVVEERSYPVAREQCCRRPNSQHYGISNYCVPDPVKQIREALTQTHTAIAVI		240		
Sbjct 181	YIQNGVVEERSYPVAREQCCRRPNSQHYGISNYCVPDPVKQIREALTQTHTAIAVI		240		
Query 241	IGIKDLRAFQHYDGRITIRHNGYQPNYHVNIVGSGTGQVDYIVRNSMDTTWGSQGY		300		
Sbjct 241	IGIKDLRAFQHYDGRITIRHNGYQPNYHVNIVGSGTGQVDYIVRNSMDTTWGSQGY		300		
Query 301	GVFQAGNNLMIEQYPPVVM 321				
Sbjct 301	GVFQAGNNLMIEQYPPVVM 321				

Hình 5. So sánh trình tự amino acid suy diễn của gen Der f 1 được tách dòng và trình tự kháng nguyên Der f 1 (mã số ABL84751.1) được công bố trên GenBank.

4. Bàn luận

Các cặp mồi F1/R2 và F3/R3 được thiết kế dựa trên các vùng bảo thủ của 32 trình tự cDNA mã hóa cho kháng nguyên Der f 1 đã cho thấy, khả năng liên kết đặc hiệu với cDNA khuôn của *D. farinae* để tạo ra 2 đoạn cDNA vùng 5' và 3' của gen Der f 1 với 197 nucleotide gối nhau. Ba phản ứng PCR lần lượt dựa vào sự gối lên nhau của 2 đoạn cDNA thành phần đã được sử dụng trong nghiên cứu của Best EA (2000), để khuếch đại toàn bộ gen mã hóa cho Der f 1 [9]. Với kích thước 983 nucleotide, trình tự mã hóa cho Der f 1 trong nghiên cứu của chúng tôi bao gồm ~19 amino acid trình tự peptide tín hiệu, 80 amino acid thuộc trình tự tiền protein chưa hoạt hóa (pro-form) và 222-223 amino acid thuộc trình tự protein trưởng thành [8].

Trình tự cDNA của gen Der f 1 trong nghiên cứu của chúng tôi có 3 vị trí khác biệt so với trình tự nucleotide tham chiếu (mã số AB034946.1). Tuy nhiên, do tính chất thoái hóa của bộ ba mã di truyền mà trình tự amino acid suy diễn tương ứng của gen Der f 1 chỉ có 1 amino acid khác biệt so với trình tự amino acid tham chiếu (mã số ABL84751.1) và amino acid sai khác này không nằm trong trung tâm hoạt động của cysteine protease Der f 1 [15]. Cả trình tự cDNA của gen này và amino acid suy diễn tương ứng đều có 99% tương đồng với 2 trình tự tham chiếu trên. Sự sai khác giữa các trình tự nucleotide và amino acid của Der f 1 trong nghiên cứu này có thể phản ánh sự đa dạng tự nhiên của các chủng *D. farinae* tại các khu vực sinh thái khác nhau. Điều này có thể do sự thích nghi của *D. farinae* với điều kiện môi trường địa phương, dẫn đến sự biến đổi di truyền trong quần thể loài. Trình tự nucleotide của gen mã hóa cho Der f 1 trong nghiên cứu của chúng tôi đã được đăng ký trên GenBank với mã số PP417680.1.

5. Kết luận

Gen mã hóa cho Der f 1 đã được tách dòng thành công trong vector pGEM-T ở vi khuẩn *E. coli*. Trình tự gen mã hóa thu được có độ dài 966 bp và mã hóa cho 321 amino acid. Trình tự nucleotide và amino acid suy diễn của gen Der f 1 từ *D. farinae* thu thập ở Việt Nam không có sự khác biệt đáng kể so với các trình tự tham chiếu. Cụ thể là, cả hai trình tự này đều tương đồng 99% với trình tự nucleotide và trình tự amino acid tham chiếu của Der f 1 với mã số lần lượt là AB034946.1 và ABL84751.1 trên GenBank. Trình tự gen mã hóa cho Der f 1 trong nghiên cứu của chúng tôi đã được đăng ký trên GenBank với mã số PP417680.1.

LỜI CẢM ƠN

Nghiên cứu được thực hiện với sự hỗ trợ kinh phí từ nhiệm vụ khoa học và công nghệ cấp quốc gia “Nghiên cứu chế tạo dị nguyên mật bụi nhà (HDM - House dust mites) bằng công nghệ protein tái tổ hợp định hướng ứng dụng trong chẩn đoán bệnh dị ứng tại Việt Nam”, mã số ĐTĐL.CN-109/21, thuộc Chương trình 562 - Lĩnh vực khoa học sự sống. Chúng tôi xin cảm ơn và cam kết không có bất kỳ xung đột lợi ích nào trong nghiên cứu này.

TÀI LIỆU THAM KHẢO

- [1] L.G. Arlian, M.S. Morgan, J.S. Neal (2002), “Dust mite allergens: Ecology and distribution”, *Current Allergy and Asthma Reports*, **2**, pp.401-411, DOI: 10.1007/s11882-002-0074-2.
- [2] D. Kim, H. Jeon, J.Y. Park, et al. (2024), “Characteristics of house dust mite allergens in Southeast and East Asia with the effect of hygienic practices”, *Indoor Environments*, **1**(2), DOI: 10.1016/j.indenv.2024.100010.
- [3] T.H. Trinh, P.T. Nguyen, T.T. Tran, et al. (2023), “Profile of aeroallergen sensitizations in allergic patients living in southern Vietnam”, *Front Allergy*, **3**, DOI: 10.3389/falgy.2022.1058865.
- [4] N.H.H. Pham, N.D. Tran, H.L. Do (2021), “The prevalence of allergens detected by immunoblotting assay and some related factors among allergy patients at Can Tho Hospital of Dermato-Venereology in 2020”, *Can Tho Journal of Medicine and Pharmacy*, **41**, pp.167-175 (in Vietnamese).
- [5] W.R. Thomas (2015), “Hierarchy and molecular properties of house dust mite allergens”, *Allergology International*, **64**(4), pp.304-311, DOI: 10.1016/j.alit.2015.05.004.
- [6] R.H. Shafique, P.B. Klimov, M. Inam, et al. (2014), “Group 1 allergen genes in two species of house dust mites, dermatophagoides farinae and *D. pteronyssinus* (Acari: Pyroglyphidae): Direct sequencing, characterization and polymorphism”, *PLOS ONE*, **9**(12), DOI: 10.1371/journal.pone.0114636.
- [7] S.V. Susanne, H.H. Hans, V.R. Thomas (2014), “Recombinant house dust mite allergens”, *Methods*, **66**(1), pp.67-74, DOI: 10.1016/j.ymeth.2013.07.034.
- [8] E.A. Best, K.E. Stedman, C.M. Bozic, et al. (2000), “A recombinant group 1 house dust mite allergen, rDer f 1, with biological activities similar to those of the native allergen”, *Protein Expr. Purif.*, **20**, pp.462-471, DOI: 10.1006/prep.2000.1327.
- [9] G. Schmidt, G. Gadermaier, H. Pertl, et al. (2008), “Production of recombinant allergens in plants”, *Phytochem. Rev.*, **7**, pp.539-552, DOI: 10.1007/s11101-008-9099-z.
- [10] D.M. Francis, R. Page (2010), “Strategies to optimize protein expression in *E. coli*”, *Curr. Protoc. Protein Sci.*, **5**(1), DOI: 10.1002/0471140864.ps0524s61.
- [11] T.L.C. Nguyen, Q.U. Nguyen, Q.Q.H. Nguyen, et al. (2024), “Identification of some species of house dust mites isolated in Hanoi, Vietnam using molecular biology”, *Vietnam Journal of Preventive Medicine*, **34**(2), pp.85-93, DOI: 10.51403/0868-2836/2024/1626 (in Vietnamese).
- [12] J.M. Chirgwin, A.E. Przybyla, R.J. MacDonald, et al. (1979), “Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease”, *Biochemistry*, **18**(24), pp.5294-5299, DOI: 10.1021/bi00591a005.
- [13] M. Miura, C. Tanigawa, Y. Fujii, et al. (2013), “Comparison of six commercially-available DNA polymerases for direct PCR”, *Revista do Instituto de Medicina Tropical de Sao Paulo*, **55**(6), pp.401-406, DOI: 10.1590/S0036-46652013000600005.
- [14] C. Ayling (2023), “TA cloning approaches to cloning DNA with damaged ends DNA”, *Methods in Molecular Biology*, **2633**, pp.55-64, DOI: 10.1007/978-1-0716-3004-4_5.
- [15] K. Takahashi, T. Takai, T. Yasuhara, et al. (2001), “Effects of site-directed mutagenesis in the cysteine residues and the N-glycosylation motif in recombinant Der f 1 on secretion and protease activity”, *Int. Arch. Allergy Immunol.*, **124**(4), pp.454-460, DOI: 10.1159/000053780.