

KHAI THÁC VÀ CHỌN GEN MÃ HÓA CHITINASE TỪ NGUỒN DỮ LIỆU DNA METAGENOME

Nguyễn Thị Thơm¹, Nguyễn Thị Hồng Hà¹, Phạm Bích Ngọc^{1,✉}, Chu Hoàng Hà¹, Hồ Bích Hải², Lê Mai Hương³

¹Viện Công nghệ sinh học - Viện Hàn lâm Khoa học và Công nghệ Việt Nam

²Viện Công nghệ thông tin - Viện Hàn lâm Khoa học và Công nghệ Việt Nam

³Viện Hóa học các hợp chất thiên nhiên - Viện Hàn lâm Khoa học và Công nghệ Việt Nam

✉Người chịu trách nhiệm liên lạc. E-mail: pbnogoc@ibt.ac.vn

Ngày nhận bài: 18.12.2018

Ngày nhận đăng: 15.03.2019

TÓM TẮT

Khai thác cơ sở dữ liệu DNA metagenome đất vùng rễ của cây cà phê tái canh 2 năm tuổi tại tỉnh Đăk Lăk và so sánh trên dữ liệu CAZy đã tìm thấy 56 trình tự nucleotide mã hóa chitinase. Kết quả phân tích các trình tự trên BLASTP đều dự đoán các protein có độ bao phủ từ 90% trở lên với hệ số tương đồng từ 33% - 99 % so với các gen chitinase đã được công bố. Trong nghiên cứu này, với mục đích tìm kiếm các gen mã hóa chitinase mới có hoạt tính diệt tuyến trùng và bệnh nấm hại rễ ở cà phê, từ 8/56 trình tự với độ tương đồng nhỏ hơn 75%, đã lựa chọn được trình tự ChiA_2. Kết quả phân tích trên BLASTP cho thấy, ChiA_2 có giá trị điểm tối đa cao nhất, đồng thời độ tương đồng của trình tự đạt 72%. Bằng phần mềm InterProScan, trình tự gen *ChiA_2* được dự đoán có trung tâm hoạt động chứa 9 gốc acid amin **FDGIDIDWE** nằm ở vị trí acid amin thứ 134-142. Đồng thời, trình tự này cũng được khảo sát về các tín hiệu tiết ngoại bào bằng sử dụng phần mềm SignalP 4.1 server cho thấy ChiA_2 có trình tự của tín hiệu tiết dài 30 acid amin nằm ở đầu N. Trình tự nucleotide gen *ChiA_2* bao gồm 1167 nucleotide, mã hóa cho 389 acid amin với khối lượng phân tử được dự đoán khoảng 43 kDa. Đoạn gen *ChiA_2* với kích thước khoảng 1,2 kb được tổng hợp, thiết kế thành công vào vector pET-21(a+) và biến nạp vào tế bào *E. coli* BL21 (DE3) bằng phương pháp sốc nhiệt. Protein ChiA_2 đã được chứng minh biểu hiện đặc hiệu trong dòng tế bào *E. coli* BL21 (DE3) tái tổ hợp và có hoạt tính phân giải Chitin đạt 4,9 U/ml. Như vậy, trong nghiên cứu này, phân tích các dữ liệu metagenome từ mẫu đất vùng rễ đã cung cấp các thông tin di truyền tương đối đầy đủ về các gen mã hoá chitinase, bước đầu đã biểu hiện thành công và đánh giá hoạt tính được một trình tự ChiA_2 mã hoá Chitinase 2.

Từ khóa: chitinase, DNA metagenome, *E. coli*, gen mới, tái tổ hợp

MỞ ĐẦU

Năm 2018, Việt Nam đã xuất khẩu được trên 1,1 triệu tấn cà phê, đạt kim ngạch 2,3 tỷ USD, trong đó cà phê Robusta đạt 925.000 tấn, cà phê hòa tan và rang xay là 89.000 tấn. Sản lượng xuất khẩu của cà phê tăng 34,8% về lượng và tăng 17,2% về kim ngạch (VIFOCA, 2018). Với ý nghĩa kinh tế to lớn từ xuất khẩu của các sản phẩm cà phê đã đưa Việt Nam trở thành một trong những quốc gia xuất khẩu cà phê lớn nhất trên thế giới. Tuy nhiên, do khai thác quá mức, quá trình thâm canh trong điều kiện không có cây che bóng, thoái hóa đất, sâu bệnh hại... làm cây cà phê nhanh chóng già cỗi. Việc đẩy mạnh tái canh, thay thế vườn cây già cỗi gây không ít trở ngại

cho ngành cà phê Việt Nam do chi phí trong quá trình tái canh cao, tỷ lệ sâu bệnh nhiều, đặc biệt nghiêm trọng là bệnh tuyến trùng hại rễ, bệnh nấm hại rễ. Đây cũng là những khó khăn chính ảnh hưởng nghiêm trọng đến sản xuất cà phê trên thế giới.

Các enzyme ngoại bào đặc biệt là enzyme chitinase đóng vai trò rất quan trọng trong đấu tranh sinh học. Nhiều nghiên cứu cho thấy chitinase ngoại bào từ các loài nấm kí sinh tuyến trùng có khả năng phân hủy màng trứng tuyến trùng và phân hủy vách tế bào của các loài vi nấm gây hại. Nấm kí sinh côn trùng có khả năng vượt qua hàng rào bảo vệ của côn trùng bằng cách sản xuất ra nhiều enzyme ngoại bào nhằm phân giải protein và chitin tạo điều kiện thuận lợi cho quá trình xâm nhập vào lớp biểu bì và

gây nhiễm trùng (St Leger, 1991). Ngoài ra, một số enzyme khác cũng liên quan đến quá trình kí sinh của nấm vào tuyến trùng như collagenase, lysosymes và protease (Sharon *et al.*, 2001; Khan *et al.*, 2004). Trên thế giới, gen mã hóa chitinase trên đối tượng vi sinh vật đã được nhiều tác giả nghiên cứu, tập trung chủ yếu vào các gen mã hóa chitinase từ một số chủng nấm. Việc sử dụng các hệ biểu hiện để tạo enzyme tái tổ hợp cũng rất quan trọng cho sự phát triển các enzyme mới và sản xuất lượng lớn các enzyme này. Gen mã hóa chitinase đã được nhiều tác giả nghiên cứu biểu hiện ở các hệ thống biểu hiện khác nhau. Barboza-Corona và cộng sự (2003) đã nhân dòng và biểu hiện gen mã hóa endochitinase từ vi khuẩn *B. thuringiensis* trong vi khuẩn *E. coli* DH5 α . Dựa trên các cơ sở khoa học về khả năng diệt tuyến trùng và vi nấm gây bệnh của enzyme chitinase, nhóm nghiên cứu tiến hành khai thác gen chitinase mới có hoạt tính diệt tuyến trùng mạnh từ dữ liệu metagenome nhằm tạo chủng và chế phẩm enzyme chitinase thô để phòng trừ các loại tuyến trùng và vi nấm gây hại.

Hiện nay, metagenomics (phân tích DNA của các vi sinh vật trong môi trường mà không cần nuôi cấy) là kỹ thuật mới, kết hợp các kỹ thuật công nghệ sinh học khác nhau từ công nghệ gen, giải trình tự đến tin sinh học để nghiên cứu thành phần, chức năng và động học của quần thể vi sinh vật. Hiện nay số lượng lớn dữ liệu thu được từ công nghệ metagenomics đã được công bố rộng rãi, đã tạo điều kiện thuận lợi cho các nhà khoa học nghiên cứu phân tích các gen mới góp phần phục vụ mục tiêu canh tác nông nghiệp bền vững của Việt Nam. Việc áp dụng kỹ thuật metagenomics vào nghiên cứu hệ gen của vi sinh vật đất, đã và sẽ đóng vai trò to lớn trong các nghiên cứu tiếp theo. Tuy nhiên, tại Việt Nam chưa có nhiều các nghiên cứu ứng dụng kỹ thuật metagenomics trong nghiên cứu hệ vi sinh vật đất nói chung cũng như hệ vi sinh vật vùng rễ cà phê nói riêng. Vì vậy, nghiên cứu này tập trung đánh giá sự có mặt của gen chitinase trong đất vùng rễ của cây cà phê thông qua kỹ thuật metagenomics.

NGUYÊN LIỆU VÀ PHƯƠNG PHÁP

Nguyên liệu

Nguồn dữ liệu metagenomics của vi sinh vật đất vùng rễ cây cà phê tái canh 2 năm tuổi, được giải trình tự bằng hệ thống giải trình tự HiSeq Illumina và được chú giải chức năng dựa trên CAZy (Do *et al.*, 2014). Dữ liệu gồm có 1.291.826 đoạn ORF

được dự đoán lắp ráp, tương ứng với 1.291.826 gen được dự đoán lắp ráp, và 673.899 gen được dự đoán chú giải chức năng.

Chủng *Bacillus* sp. từ bộ sưu tập chủng vi sinh vật của Phòng Công nghệ tế bào thực vật có hoạt tính chitinase được sử dụng làm đối chứng dương trong nghiên cứu đánh giá hoạt tính của protein ChiA₂.

Phương pháp

Xác định các họ Glycosyl Hydrolases (GH) có chứa enzyme chitinase theo CAZy (<http://www.cazy.org>)

Sử dụng dữ liệu phân loại của CAZy để xác định có bao nhiêu họ GH chứa enzyme này, sau đó tìm hiểu về cấu trúc không gian, các loại amino acid cho điện tử và proton trong quá trình hoạt động của enzyme. Kết quả sẽ được tổng hợp thành bảng phân loại danh sách các họ GH chứa enzyme này để tìm hiểu các thông tin sâu hơn.

Khai thác trình tự mã hóa chitinase từ dữ liệu DNA metagenome của vi sinh vật nội sinh trên cây cà phê

Sau khi có các gen mã hóa cho chitinase khác nhau, sử dụng BLASTP để so sánh với trình tự nucleotide của các ORF thuộc metagenome của vi sinh vật trong đất vùng rễ cà phê với mong muốn xác định được gen tiềm năng nhất tiến tới tách dòng và biểu hiện protein chitinase tái tổ hợp.

Thiết kế vector biểu hiện pET-21a(+) mang gen mã hóa chitinase

Vector pUC19/chiA₂ với điểm cắt *Bam*HI và *Sac*I nằm tương ứng ở hai đầu 5' và 3' và pET-21a(+) đồng thời được cắt tạo đầu so le bằng cặp enzyme giới hạn *Bam*HI và *Sac*I. Đoạn gen *chiA*₂ được tách ra từ vector pUC19 và vector pET-21a(+) thu được sau đó được tinh sạch và sử dụng làm nguyên liệu cho phản ứng nối ghép dưới sự xúc tác của T4-ligase để tạo thành vector pET21/chiA₂. Sản phẩm nối ghép được biến nạp vào *E. coli* bằng phương pháp sốc nhiệt theo Cohen và cộng sự (1972). Tiến hành sàng lọc các thể biến nạp mang plasmid tái tổ hợp mong muốn bằng phương pháp PCR sử dụng các cặp mồi đặc hiệu cho gen trên theo chu trình: 95°C - 3 phút, 30 chu kì (95°C - 30 giây, 58°C - 30 giây, 72°C - 1 phút), 72°C - 10 phút, ủ mẫu ở 4°C và phương pháp cắt kiểm tra bằng enzyme hạn chế. Plasmid tái tổ hợp pET21/chiA₂ sau đó được biến nạp vào chủng *E. coli* BL21 (DE3) để biểu hiện gen.

Biểu hiện protein tái tổ hợp trong tế bào vi khuẩn *E. coli*

Một khuẩn lạc chứa plasmid tái tổ hợp được nuôi cấy trong 10 ml môi trường LB lỏng có kháng sinh ampicillin ở 37°C, lắc 200 vòng/phút qua đêm. Cấy chuyển 1 ml dịch khuẩn đã nuôi qua đêm sang 100 ml môi trường LB mới và nuôi ở 37°C, lắc 200 vòng/phút đến khi giá trị OD₆₀₀ = 0,4-0,6 thì bổ sung IPTG tới nồng độ 1 mM và chuyển sang nuôi lắc 200 vòng/phút ở 28°C trong 6 giờ để cảm ứng biểu hiện protein ngoại lai. Tế bào được thu lại bằng cách ly tâm dịch nuôi cấy ở 6000 vòng/phút trong 15 phút và hòa trong dung dịch đệm PBS 1X. Bổ sung lysozyme tới nồng độ 0,1 mg/ml, ủ trên đá 1 giờ và tiến hành siêu âm phá tế bào, sau đó ly tâm 13000 vòng/phút trong 30 phút, thu cả dịch nổi và cặn để kiểm tra bằng phương pháp điện di SDS-PAGE. Kỹ thuật điện di biến tính trên gel polyacrylamid (SDS-PAGE) được thực hiện theo Laemmli (1970) với các thiết bị của hãng Bio-rad.

Kỹ thuật lai miễn dịch (Western Blotting)

Protein tổng số phân tích trên gel acrylamide được thấm chuyển sang màng nitrocellulose PVDF bằng thiết bị lai (Mini-Transblotter) ở điều kiện 22 mA trong 16 h ở 4°C trong dung dịch đệm chuyển (25 mM Tris-HCl, 192 mM Glycine, 20% Methanol, pH 8,3).

Màng được ủ trong đệm TBS (20 mM Tris-HCl, 500 mM NaCl, pH 7,5) với 5% sữa tách béo trong 2 h ở nhiệt độ phòng. Tiếp theo đó ủ màng với dung dịch pha loãng của kháng thể thứ nhất trong 1 h ở nhiệt độ phòng. Rửa màng 3 lần mỗi lần 10 phút trong đệm TBS với 0,5% sữa để loại bỏ kháng thể dư thừa. Ủ tiếp với kháng thể thứ hai đã được liên kết với alkaline phosphatase trong 1 h ở nhiệt độ phòng.

Rửa màng 3 lần mỗi lần 10 phút trong đệm TTBS (TBS, 0,05% Tween 20).

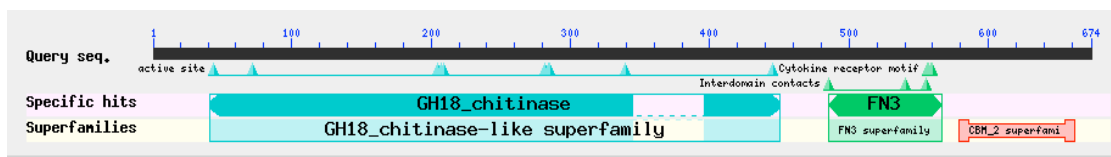
Vạch protein được phát hiện bằng cách nhuộm màng trong dung dịch nhuộm (100 mM Tris; 100 mM NaCl; 5 mM MgCl₂; 75 µg/ml BCIP (5-bromo, 4-chloro, 3-indolylphosphate); 150 µg/ml NBT (nitroblue tetrazolium); pH 8,9), quan sát màu, ngừng phản ứng nhuộm bằng cách rửa màng với nước.

KẾT QUẢ VÀ THẢO LUẬN

Khai thác trình tự mã hóa cho chitinase từ cơ sở dữ liệu DNA metagenome của vi sinh vật đất vùng rễ cây cà phê

Từ dữ liệu DNA metagenomes của vi sinh vật đất vùng rễ, chúng tôi khai thác được 10 nhóm gen phân giải chitin bao gồm alpha-N-acetylglucosaminidase, beta-N-acetylglucosaminidase, chitinase, N-acetylglucosamine-6-sulfatase, N-acetylglucosamine-6-phosphate deacetylase, UDP-2,3-diacetylglucosamine hydrolase, glucosamine-6-phosphate deaminase, exo-1,4-beta-D-glucosaminidase, N-acetylglucosamine-1-phosphodiester alpha-N- và acetylglucosaminidase.

Dựa trên dữ liệu CAZy tìm được 56 trình tự gen hoàn chỉnh mã hóa chitinase trong dữ liệu metagenome của vi sinh vật trên cây cà phê. Trong số đó, có 8 trình tự được xem là gen mới với độ tương đồng <75% so với các gen chitinase đã được công bố. Khảo sát lại vùng bảo thủ của 8 trình tự được lựa chọn bằng BLASTP, chúng tôi thấy các trình tự đều chứa các vùng đặc thù cho chitinase (specific hit) và có vị trí hoạt động (active site) như ở hình 1.



Hình 1. Độ đặc hiệu trình tự của các trình tự chứa gen chitinase thuộc họ glycosyl 18

Kết quả đánh giá *in silico* các gen chitinase khai thác

Các gen mã hoá chitinase sau khi được khai thác từ cơ sở dữ liệu metagenome được so sánh độ tương đồng dựa trên các trình tự có trên ngân hàng gen của NCBI (<https://www.ncbi.nlm.nih.gov/>), và

tập trung vào các trình tự đã ước đoán chức năng ban đầu. Các trình tự có độ tương đồng nhỏ hơn 50% sẽ tiếp tục được sàng lọc bằng nhiều phần mềm dự đoán cấu trúc và chức năng của protein như BLASTP, EXPASY. Kết quả BLASTP giúp cho việc ước đoán rất nhiều các thông số của enzyme như vị trí các vùng bảo tồn, họ enzyme, nguồn vi

sinh vật chứa enzyme... Trước khi đưa vào thực nghiệm, các gen này tiếp tục được phân tích sử dụng phần mềm dự đoán về tính chất như khả năng chịu kiềm/axit (Lin *et al.*, 2013). Cấu trúc bậc 3 của protein được so sánh bằng phần mềm PHYRE2.

Từ các trình tự khai thác được, các gen được lựa chọn theo các tiêu chí sau: (1) là các trình tự hoàn thiện, trình tự mới; (2) có chỉ số điểm tối đa (max score) khi blast với mẫu dò; (3) được dự đoán bằng KEGG tương đồng với mẫu dùng khai thác. Kết quả

so sánh đặc điểm của 8 trình tự mới chúng tôi nhận thấy rằng giá trị E thể hiện độ tin cậy, chỉ số E càng thấp kết quả càng đáng tin cậy, đối với tham số nucleotide giá trị E phải nhỏ hơn 1E-6 vì vậy trình tự 1 không thỏa mãn yêu cầu. Với 7 trình tự còn lại, giá trị Max score của trình tự 7 (ChiA_2) đạt giá trị cao nhất đồng thời độ tương đồng của trình tự đạt 72%. Vì vậy, chúng tôi sử dụng trình tự này (bảng 2) để biểu hiện và đánh giá tiếp về vùng bảo thủ, dự đoán cấu trúc enzyme.

Bảng 1. So sánh tương đồng của 8 trình tự lựa chọn.

Tên trình tự	Ký hiệu	Điểm tối đa	Tổng điểm	Độ bao phủ (%)	Giá trị E	Độ tương đồng (%)
Trình tự 1	Node 10334	45,4	45,4	52	0.07	33
Trình tự 2	Node 448702	53,9	53,9	51	0.000001	65
Trình tự 3	Node 32291	139	139	89	2E-36	44
Trình tự 4	Node 384	301	301	96	5E-97	41
Trình tự 5	Node 54	340	340	98	7E-113	47
Trình tự 6	Node 8845	205	205	86	2E-61	53
Trình tự 7	Node 31	479	479	89	3E-166	72
Trình tự 8	Node 4449	455	455	96	2E-156	66

Bảng 2. Đặc tính của trình tự gen ChiA_2 lựa chọn biểu hiện.

Tên enzyme	Ký hiệu	Mã trình tự	GH	Nhóm	Số aa	Vị trí trung tâm hoạt động	pH hoạt động
Chitinase	ChiA_2	Node_31	GH18	II	389	134-142	Kiềm



ChiA_2

Hình 2. Cấu trúc bậc 3 của gen chiA_2.

Dựa trên phần mềm Phyre2, cấu trúc các trình tự amino acid từ các gen đã được ước đoán. Kết quả thể hiện trong hình 2 xác định gen *ChiA_2* có cấu trúc xác định gồm 8 vòng xoắn α/β . Đồng thời, trình tự ChiA_2 có độ tương đồng cao nhất 39% với cấu trúc

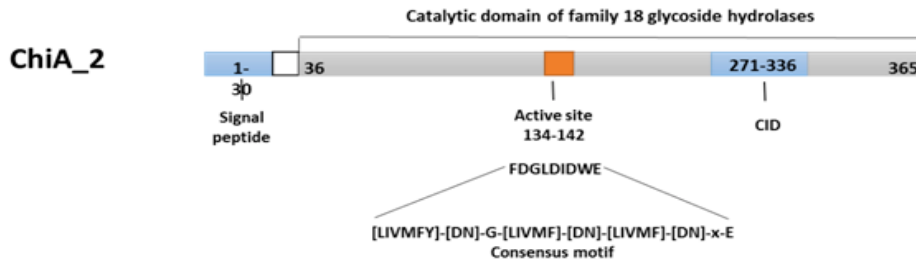
chitinase a1 từ *Bacillus circulans* với độ tin cậy 100%. Trình tự này cũng được khảo sát về các tín hiệu tiết có trên trình tự sử dụng phần mềm SignalP 4.1 server, chúng tôi nhận thấy ChiA_2 có trình tự của tín hiệu tiết dài 30 acid amin nằm ở đầu N. Đồng thời chúng tôi cũng khai thác mô hình cấu trúc của protein dựa theo phần mềm InterProScan và Prosite (Karin *et al.*, 2014).

Kết quả hình 3 cho thấy, trình tự ChiA_2 bao gồm 1167 nucleotide, mã hóa cho 389 acid amin với khối lượng phân tử được dự đoán khoảng 43 kDa và trung tâm hoạt động chứa 9 gốc acid amin FDGIDIDWE nằm ở vị trí acid amin thứ 134-142.

Trình tự *ChiA_2* mã hóa cho chitinase đã được tối ưu mã, thêm điểm cắt của enzyme hạn chế. Đồng thời chúng tôi cũng thiết kế thêm trình tự 6xHis nằm ở đầu 5' đoạn gen sau trình tự acid amin mở đầu để tiện cho việc tinh sạch protein tái tổ hợp sau này cũng như bổ sung thêm trình tự c-myc nằm đầu 3'

của đoạn gen để thuận lợi cho việc đánh giá biểu hiện gen. Sau đó đoạn gen được đặt tổng hợp tại

công ty Sinh hóa Phù Sa và đưa vào vector pUC19 để chuẩn bị cho biểu hiện gen.



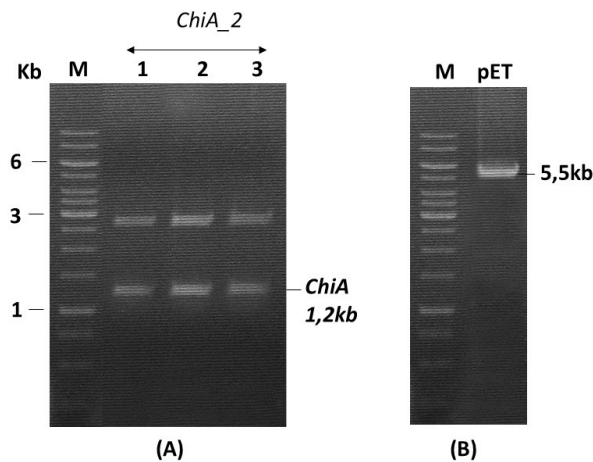
Hình 3. Phân tích mô hình cấu trúc protein.

Thiết kế vector biểu hiện pET-21a(+)/chiA

Sau khi khai thác cơ sở dữ liệu DNA metagenome và xác định được trình tự đoạn gen tiềm năng mã hóa enzyme chitinase. Đoạn gen *ChiA_2* được đặt tổng hợp và đặt trong vector tách dòng pUC19 với điểm cắt *Bam*HI và *Sac*I ở hai đầu 5' và 3'. Sau đó, chúng tôi tiến hành cắt vector pUC19/*chiA_2* bằng cặp enzyme *Bam*HI và *Sac*I để thu đoạn gen *ChiA_2*. Gen *ChiA_2* được gắn vào vector biểu hiện pET-21a(+) thông qua các điểm cắt trong đồng của *Bam*HI và *Sac*I. Sản phẩm cắt được điện di kiểm tra trên gel agarose 0,8%.

Kết quả điện di hình 4 cho thấy đoạn gen *ChiA_2* (kích thước tính toán khoảng 1,2kb) đã được tách ra từ vector pUC19. Trong khi đó, vector pET21-a(+) cũng đã được mở vòng thành công, bằng DNA thu được là một băng sáng, rõ nét và có kích

thước khoảng 5,5Kb phù hợp với kích thước lý thuyết. Đoạn gen *ChiA_2* và pET21-a(+) được tinh sạch và sử dụng cho phản ứng ghép nối dưới sự xúc tác của enzyme T4-ADN ligase để tạo plasmid tái tổ hợp pET21/*ChiA_2*. Sau khi lai ghép chúng tôi tiến hành biến nạp sản phẩm lai vào tế bào *E. coli* bằng phương pháp sốc nhiệt và nuôi trên môi trường có bổ sung kháng sinh chọn lọc ampicillin. Sản phẩm biến nạp được nuôi lắc trong môi trường LB lỏng ở 37°C trong 45 phút trước khi cấy trải trên môi trường thạch đĩa LB qua đêm ở 37°C. Theo cơ chế chọn lọc bằng kháng sinh, các dòng tế bào *E. coli* nguyên thủy không chứa plasmid ngoại lai sẽ bị tiêu diệt trên môi trường chứa kháng sinh. Trong khi đó vector biểu hiện pET-21a(+) có chứa trình tự gen kháng kháng sinh ampicillin, vì vậy chỉ những dòng tế bào nào nhận được plasmid mới có khả năng mọc trên môi trường chứa kháng sinh này.



Hình 4. Kết quả điện di sản phẩm cắt vector pUC19/*chiA* và pET-21a(+) bằng *Bam*HI và *Sac*I. (A): Sản phẩm cắt plasmid pUC19/*ChiA*; (B): Sản phẩm cắt mở vòng vector pET-21a(+).

Những khuẩn lạc phát triển được trên môi trường sàng lọc được lựa chọn để kiểm tra tiến hành thông qua Colony PCR bằng cặp mồi đặc hiệu cho gen *ChiA_2*. Các khuẩn lạc dương tính với PCR được tiếp tục nuôi để tách plasmids và kiểm tra lại bằng enzyme cắt hạn chế.

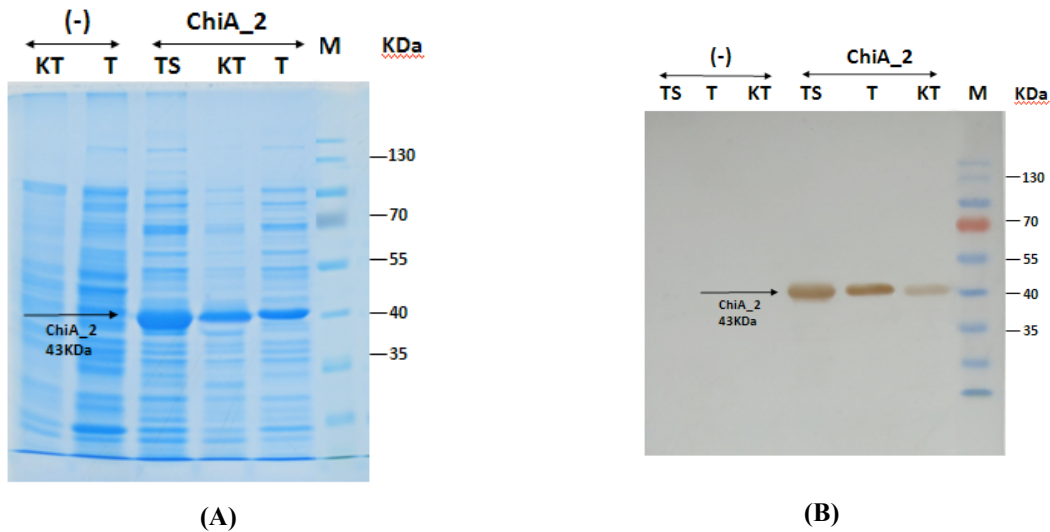
Kiểm tra sự biểu hiện của chitinase tái tổ hợp

Để kiểm tra khả năng biểu hiện tạo protein tái tổ hợp của chủng vi khuẩn mang gen *ChiA_2*, chúng tôi đã chọn ngẫu nhiên một khuẩn lạc *E. coli* BL21 (DE3) mang gen *ChiA_2* đem nuôi lắc biểu hiện trong môi trường LB lỏng có chứa kháng sinh. Dòng đối chứng âm được chọn là vector pET-21a(+) không mang gen cũng được nuôi cấy song song cùng điều kiện. Sau 5 h nuôi cấy, ly tâm thu sinh khối tế bào, sau đó đưa các mẫu về cùng nồng độ và thực hiện phá vỡ tế bào bằng siêu âm. Mẫu protein được biến tính bằng đệm biến tính và điện di kiểm tra trên gel SDS-PAGE 12,6%. Kết quả điện di được quan

sát sau khi nhuộm với thuốc nhuộm (Hình 5A).

Kết quả phân tích ở Hình 5A cho thấy protein tổng số từ dòng mang gen *ChiA_2* đã xuất hiện bằng protein có kích thước khoảng 43 kDa phù hợp với kích thước lý thuyết của proetin ChiA₂. Trong khi đó ở đường chạy đối chứng (chúng không mang gen) không xuất hiện bằng protein này, điều này chứng tỏ protein ChiA₂ đã được biểu hiện trong chủng vi khuẩn *E. coli* BL21 (DE3) tái tổ hợp.

Trong nghiên cứu của chúng tôi, đoạn gen mã hóa ChiA₂ được thiết kế gắn đuôi -cmyc để thuận tiện cho phản ứng lai Western blot kiểm tra sự có mặt của protein ChiA₂. Do vậy, chúng tôi tiến hành lai Western blot với kháng thể anti-cmyc. Kết quả được thể hiện trong hình 5.B. Kết quả Western blot cho thấy, ChiA₂ đã được biểu hiện đặc hiệu trong dòng tế bào *E. coli* BL21 (DE3) tái tổ hợp. Cùng với đó, protein cũng được tập trung cả ở pha tan và không tan



Hình 5. Kết quả đánh giá sự biểu hiện của chitinase tái tổ hợp. **(A):** Kết quả phân tích SDS-PAGE protein ChiA-2 tái tổ hợp biểu hiện trong *E. coli* BL21 (DE3); **(B):** Kết quả lai Western blot protein chitinase biểu hiện *E. coli* BL21 (DE3)/ChiA-2 với kháng thể kháng c-myc. **TS:** Protein tổng số; **T:** Protein pha tan; **KT:** Protein pha không tan

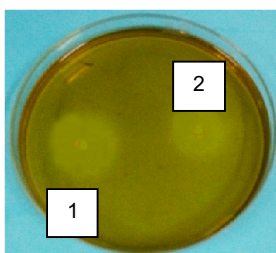
Kết quả đánh giá hoạt tính protein ChiA tái tổ hợp

Sau khi kiểm tra mức độ biểu hiện tái tổ hợp của protein ChiA₂, chúng tôi tiến hành đánh giá sơ bộ hoạt tính phân giải cơ chất chitin. Kết quả thử hoạt tính được thể hiện qua Bảng 3.

Dịch protein ChiA₂ được xác định hoạt tính

chitinase trên đĩa thạch có chứa chitin huyền phù 0,5%, các giếng đều xuất hiện vòng phân giải chitin màu trắng sau khi nhuộm bằng lugol 1,0%. Kết quả cho thấy protein ChiA₂ tái tổ hợp có hoạt tính tốt hơn chủng *Bacillus* sp. với hoạt độ lần lượt là 4,9 và 4,4 U/ml. Điều này cho thấy, gen *ChiA_2* đã được biểu hiện và có hoạt tính chitinase trong dòng tế bào *E. coli* BL21 (DE3).

Bảng 3. Kết quả hoạt tính chitinase từ Protein ChiA tái tổ hợp.



Chủng	Vòng hoạt tính	Hoạt tính (U/ml)
ChiA2-BL21 (1)	1,2	4,9
Chủng <i>Bacillus</i> sp. (2)	0,8	4,4

KẾT LUẬN

Từ cơ sở dữ liệu DNA metagenome của mẫu đất vùng rễ cây cà phê tái canh, đã xác định được 56 trình tự mã hóa chitinase, trong đó có 8 gen mới với độ tương đồng <75% so với các gen chitinase đã được công bố. Trình tự gen *ChiA_2* đã được tổng hợp, tách dòng và biểu hiện thành công protein ChiA₂ trong dòng tế bào *E. coli* BL21 (DE3). Protein chitinase tái tổ hợp có hoạt tính phân giải chitin từ với hoạt độ 4,9 U/ml.

Lời cảm ơn: Công trình được thực hiện bằng kinh phí Đề tài “Nghiên cứu metagenome của vi sinh vật đất vùng rễ một số cây trồng ở Việt Nam: cây thuốc có củ (cây nghệ), cây công nghiệp (cà phê) nhằm tăng năng suất và chất lượng cây trồng”, mã số DTĐLCN.14/14, Bộ Khoa học và Công nghệ.

TÀI LIỆU THAM KHẢO

Barboza Corona JE, Nieto Mazzocco E, Velazquez Robledo Robledo R, Salcedo Haman dez R, Bautista M, Jiménez B, Ibarra JE (2003). Cloning, sequencing and expression of vegetative insecticidal protein into mother cell of *Bacillus thuringiensis*. *Biochem Biophys Res Commun* 310:158-162.

Do TH, Nguyen TT, Nguyen TN, Le QG, Nguyen C, Kimura K, Truong NH (2014) Mining biomass-degrading

genes through Illumina-based de novo sequencing and metagenomic analysis of free-living bacteria in the gut of the lower termite *Coptotermes gestroi* harvested in Vietnam. *J Biosci Bioeng* 118: 665–671.

Khan Faisal I, Tahir Husain, Ramzi Hejazi ((2004) An overview and analysis of site remediation technologies. *J Environ Manag* 71: 95–122.

Karin Hjort, Ilaria Presti , Annelie Elväng , Flavia Marinelli , Sara Sjöling (2014) Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics. *Appl Microbiol Biotechnol* 98:2819–2828.

Lin H, Chen W, Ding H (2013) AcalPred: a sequence-based tool for discriminating between acidic and alkaline enzymes. *PloS One* 8: e75726.

Sharon E, Bar-Eyal M, Chet I, Herrera-Estrella A, Kleinfeld O, Spiegel Y(2001) Biological Control of the Root-Knot Nematode *Meloidogyne javanica* by *Trichoderma harzianum*. *Phytopathology* 91(7):687-93.

St. Leger RJ (1991) Integument as a barrier to microbial infections. In: Binnington, K. and Retnakaran, A., Eds., *Physiology of the Insect Epidermis* Commonwealth Scientific and Industrial Research Organization, Melbourne, 286-308.

<http://www.cazy.org>

<https://www.ncbi.nlm.nih.gov>

MINING AND SELECTING GENES ENCODING CHITINASE FROM DNA METAGENOME DATA

Nguyen Thi Thom¹, Nguyen Thi Hong Ha¹, Pham Bich Ngoc¹, Chu Hoang Ha¹, Ho Bich Hai², Le Mai Huong³

¹Institute of Biotechnology, Vietnam Academy of Science and Technology (VAST)

²Institute of Information Technology, Vietnam Academy of Science and Technology (VAST)

³Institute of Natural products chemistry, Vietnam Academy of Science and Technology (VAST)

SUMMARY

In this study, genes encoding chitinases were obtained by mining the rhizosphere metagenome database of

2-year-old replanted coffee trees in Dak Lak province and annotating using CAZy data. These sequences predictably are covered 90% of chitinase domains by BLASTp and range from 33% to 99% identity in comparison with published sequence data of chitinase genes. In order to find new functional chitinase gene(s), among 8/56 chitinase genes with sequence identity less than 75%, ChiA_2 sequence showed the highest max score and significant homology (72%) was initially selected. Analysis of *in silico* sequence identified a protein motif for a signal peptide of 30 amino acid residues at the N- terminal using SignalP 4.1 server. By InterProscan software, the *ChiA_2* gene sequence is predicted to have conserved active site **FDGIDIDWE** located at 134-142. The ChiA_2 sequence consisting of 1167 nucleotides, encoding for 389 amino acids with a predicted molecular mass of 43 kDa was synthesized, cloned into the pET21 a(+) vector and transformed into *E. coli* BL21 (DE3) by heat shock. ChiA_2 protein was specifically expressed in the recombinant *E. coli* BL21 (DE3) cell with an activity of 4.9 U/ml. In this study, analyzing metagenomic data from the soil sample of rhizosphere provided relatively sufficient genetic information of chitinase gene and initially successful expression and functional evaluation of ChiA_2 sequence, encoding Chitinase 2.

Keywords: *chitinase, DNA metagenome, E. coli, new gene, recombinant*