

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



Lê Thu Ngọc

**NGHIÊN CỨU PHÂN LẬP VÀ BIỂU HIỆN GEN MÃ HÓA
PROTEIN ĐỘC TỔ DIỆT SÂU ĐỤC QUẢ ĐẬU TƯƠNG
(*Etiella zinckenella* Treitschke) TỪ VI KHUẨN *Bacillus*
thuringiensis BẢN ĐỊA**

LUẬN ÁN TIẾN SĨ SINH HỌC ỨNG DỤNG

Hà Nội – 2024

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ

Lê Thu Ngọc

NGHIÊN CỨU PHÂN LẬP VÀ BIỂU HIỆN GEN MÃ HÓA
PROTEIN ĐỘC TỔ DIỆT SÂU ĐỤC QUẢ ĐẬU TƯƠNG
(*Etiella zinckenella* Treitschke) TỪ VI KHUẨN *Bacillus*
thuringiensis BẢN ĐỊA

LUẬN ÁN TIẾN SĨ SINH HỌC ỨNG DỤNG

Ngành: Công nghệ sinh học

Mã số: 9.42.02.01

Xác nhận của Học viện
Khoa học và Công nghệ

Người hướng dẫn 1
(Ký, ghi rõ họ tên)

Người hướng dẫn 2
(Ký, ghi rõ họ tên)

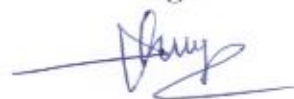
Hà Nội – 2024

LỜI CAM ĐOAN

Tôi xin cam đoan luận án: “Nghiên cứu phân lập và biểu hiện gen mã hóa protein độc tố diệt sâu đục quả đậu tương (*Etiella zinckenella* Treitschke) từ vi khuẩn *Bacillus thuringiensis* bản địa” là công trình nghiên cứu của chính mình dưới sự hướng dẫn khoa học của tập thể hướng dẫn. Luận án sử dụng thông tin trích dẫn từ nhiều nguồn tham khảo khác nhau và các thông tin trích dẫn được ghi rõ nguồn gốc. Các kết quả nghiên cứu của tôi được công bố chung với các tác giả khác đã được sự nhất trí của đồng tác giả khi đưa vào luận án. Các số liệu, kết quả được trình bày trong luận án là hoàn toàn trung thực và chưa từng được công bố trong bất kỳ một công trình nào khác ngoài các công trình công bố của tác giả. Luận án được hoàn thành trong thời gian tôi làm nghiên cứu sinh tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Hà Nội, ngày 15 tháng 10 năm 2024

Tác giả



Lê Thu Ngọc

LỜI CẢM ƠN

Tôi xin bày tỏ lòng biết ơn sâu sắc đến TS. Lê Thị Minh Thành và PGS. TS Phạm Bích Ngọc đã tận tình hướng dẫn, giúp đỡ tôi trong suốt quá trình thực hiện luận án.

Tôi xin chân thành cảm ơn TS. Đỗ Tiến Phát – Viện Công nghệ Sinh học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam đã giúp đỡ và tạo điều kiện thuận lợi cho tôi trong quá trình thực hiện các thí nghiệm nghiên cứu trong đề tài luận án.

Trong quá trình thực hiện đề tài, tôi đã nhận được sự giúp đỡ nhiệt tình của tập thể cán bộ Phòng Công nghệ Tế bào Thực vật, Phòng Công nghệ ADN Ứng dụng và Trung tâm Giống và Bảo tồn nguồn gen Vi sinh vật – Viện Công nghệ Sinh học. Tôi xin chân thành cảm ơn sự giúp đỡ quý báu đó.

Tôi xin cảm ơn Ban lãnh đạo và cán bộ phòng Đào tạo cũng như các phòng chức năng của Học viện Khoa học và Công nghệ Việt Nam đã giúp đỡ và tạo điều kiện để nghiên cứu sinh hoàn thành luận án.

Cuối cùng, tôi xin gửi lời cảm ơn sâu sắc đến gia đình, người thân, bạn bè và các đồng nghiệp đã luôn động viên, cổ vũ tinh thần cho tôi trong suốt thời gian vừa qua.

Hà Nội, ngày 15 tháng 10 năm 2024

Nghiên cứu sinh



Lê Thu Ngọc

MỤC LỤC

Trang

Trang phụ bìa	
LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
MỤC LỤC	iii
DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT	vii
DANH MỤC CÁC BẢNG	viii
DANH MỤC CÁC HÌNH VẼ, ĐỒ THỊ	ix
MỞ ĐẦU	1
1. Tính cấp thiết của đề tài	1
2. Mục tiêu nghiên cứu.....	2
3. Nội dung nghiên cứu	3
4. Ý nghĩa khoa học và thực tiễn của đề tài	3
5. Những đóng góp mới của luận án	4
Chương 1. TỔNG QUAN NGHIÊN CỨU	5
1.1. Đậu tương và côn trùng gây hại	5
1.1.1. Tầm quan trọng của cây đậu tương	5
1.1.2. Tình hình sản xuất đậu tương trên thế giới và Việt Nam.....	6
1.1.3. Sâu đục quả đậu tương <i>E. zinckenella</i> Treitschke và biện pháp phòng trừ	8
1.2. Protein độc tố diệt côn trùng có nguồn gốc từ <i>Bacillus thuringiensis</i> .	12
1.2.1. Phân loại và danh pháp	12
1.2.2. Hoạt tính và phổ diệt côn trùng	17
1.2.3. Cấu trúc protein Cry	19
1.2.4. Cơ chế diệt côn trùng	22
1.3. Một số thành tựu về tạo cây đậu tương chuyển gen kháng sâu	24
1.3.1. Tình hình nghiên cứu và lưu hành các giống đậu tương biến đổi gen kháng sâu hại trên thế giới	24
1.3.2. Nghiên cứu tạo giống đậu tương kháng sâu ở Việt nam	27
Chương 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU	30
2.1. ĐỐI TƯỢNG NGHIÊN CỨU	30
2.1.1. Vật liệu	30

2.1.2. Hóa chất và thiết bị	32
2.2. PHƯƠNG PHÁP NGHIÊN CỨU.....	33
2.2.1. Sàng lọc các chủng <i>Bacillus thuringiensis</i> có hoạt tính diệt sâu đục quả đậu tương	33
2.2.2. Giải trình tự hệ gen vi khuẩn <i>Bacillus thuringiensis</i> bằng công nghệ SMRT Sequencing của PacBio và khai thác cơ sở dữ liệu hệ gen <i>Bt</i> để tìm các gen mã hóa protein độc tố diệt sâu bằng công cụ tin sinh học	35
2.2.3. Phương pháp phân lập, tách dòng gen	39
2.2.4. Biểu hiện gen mã hóa protein độc tố diệt côn trùng trong hệ biểu hiện <i>E. coli</i> BL21 và đánh giá hoạt lực diệt sâu của protein tái tổ hợp.....	41
2.2.5. Cải biến mã di truyền và biểu hiện gen <i>cry2Ab39</i> trong lá cây thuốc lá <i>N. benthamiana</i> bằng công nghệ biểu hiện gen tạm thời thông qua <i>Agrobacterium tumefaciens</i>	46
Chương 3. KẾT QUẢ NGHIÊN CỨU	53
3.1. Kết quả sàng lọc chủng <i>Bt</i> có hoạt lực diệt sâu đục quả đậu tương cao từ các chủng trong Bộ sưu tập <i>Bt</i> bản địa Việt Nam.....	53
3.1.1. Kết quả hoạt hóa và nhân nuôi thu sinh khối các chủng <i>Bt</i>	53
3.1.2. Kết quả thử khả năng diệt sâu đục quả đậu tương của các chủng <i>Bt</i> nghiên cứu.....	54
3.1.3. Kết quả xác định đặc điểm tinh thể độc tố và động thái sinh trưởng gây chết sâu của các chủng vi khuẩn <i>Bt</i> có hoạt tính diệt sâu đục quả đậu tương $\geq 85\%$	56
3.2. Phân tích, xác định các gen độc tố <i>Bt</i> mới tiềm năng từ cơ sở dữ liệu trình tự hệ gen của 4 chủng <i>Bt</i> có độc lực cao với sâu đục quả đậu tương	59
3.2.1. Kết quả giải trình tự hệ gen 4 chủng <i>Bt</i> bằng công nghệ SMRT Sequencing của PacBio	59
3.2.1.1. Kết quả tách chiết DNA hệ gen có khối lượng phân tử cao từ các chủng <i>Bt</i>	59
3.2.1.2. Kết quả tạo thư viện DNA	60
3.2.1.3. Kết quả giải trình tự.....	62
3.2.1.4. Kết quả lắp ráp <i>denovo</i> hệ gen	63
3.2.1.5. Kết quả chú giải hệ gen lắp ráp <i>denovo</i> các chủng <i>Bt</i>	63
3.2.2. Kết quả phân tích các trình tự gen độc tố <i>Bt</i> từ cơ sở dữ liệu trình tự hệ gen 4 chủng <i>Bt</i> bằng công cụ <i>BtToxin_scanner</i> và xác định các gen mới tiềm năng trong diệt trừ sâu đục quả đậu tương	65

3.3. Kết quả phân lập, tách dòng và xác định trình tự các gen cry mã hóa protein độc tố mới có tiềm năng diệt sâu đục quả đậu tương từ các chủng Bt lựa chọn.	70
3.3.1. Kết quả phân lập, tách dòng các gen độc tố cry.....	70
3.3.2. Kết quả giải và so sánh trình tự của các gen cry đã phân lập với các gen độc tố Bt đã công bố.	72
3.4. Kết quả thiết kế vector, biểu hiện các gen độc tố cry của các chủng Bt trong tế bào <i>E. coli</i> và đánh giá hoạt tính diệt sâu đục quả đậu tương của protein tái tổ hợp	81
3.4.1. Thiết kế các vector biểu hiện pET32a(+) mang gen cry1Na, cry2Ab, cry2Ah và cry1Be.....	81
3.4.2. Biểu hiện các gen cry1Be, cry1Na, cry2Ab, và cry2Ah trong tế bào <i>E. coli</i> BL21	83
3.4.3. Thử hoạt tính diệt sâu đục quả đậu tương của các protein tái tổ hợp thu được.	85
3.4.4. Tối ưu điều kiện biểu hiện gen cry2Ab trong tế bào <i>E. coli</i> BL21 nhằm tăng lượng protein tái tổ hợp trong pha tan.	87
3.4.4.1. Ảnh hưởng của nhiệt độ nuôi cấy cảm ứng	88
3.4.4.2. Ảnh hưởng của nồng độ IPTG	89
3.4.5. Tinh sạch Trx-His-Stag-Cry2Ab	91
3.4.6. Xác định giá trị liều gây chết trung bình LC ₅₀ của protein tái tổ hợp Cry2Ab tinh sạch với ấu trùng sâu đục quả đậu tương <i>E. zinckenella</i> Treitschke	92
3.5. Tối ưu mã gen cry2Ab39 để phù hợp với hệ thống biểu hiện ở thực vật và biểu hiện tạm thời protein độc tố Cry2Ab39 trong lá thuốc lá <i>N. benthamiana</i> bằng phương pháp agroinfiltration.	94
3.5.1. Tối ưu mã di truyền của gen cry2Ab39.....	94
3.5.2. Thiết kế các vector chuyển gen pFGC/35S_cry2Ab39opt và pFGC/rbcS1A-pro_cry2Ab39	98
3.5.2.1. Thiết kế các vector chuyển gen pFGC/35S_cry2Ab39opt	98
3.5.2.2. Thiết kế các vector chuyển gen pFGC/rbcS1A-pro_cry2Ab39opt	100
3.5.3. Biểu hiện tạm thời protein tái tổ hợp Cry2Ab39 trong lá thuốc lá <i>N. benthamiana</i>	103
3.5.3.1. Ảnh hưởng của promoter đến sự biểu hiện tạm thời của gen cry2Ab39opt.....	104
3.5.3.2. Ảnh hưởng của mật độ vi khuẩn xâm nhiễm đến sự biểu hiện tạm thời của gen cry2Ab39opt	105

3.5.3.3. Tối ưu thời gian thu hoạch lá sau lây nhiễm	106
3.5.4. Tinh sạch protein tái tổ hợp Cry2Ab39 từ lá thuốc lá	108
Chương 4. THẢO LUẬN.....	110
4.1. Ứng dụng phương pháp giải trình tự thế hệ mới trong tìm kiếm và khai thác các gen độc tố diệt sâu <i>Bt</i> mới.	110
4.2. Ảnh hưởng của nhiệt độ nuôi cấy cảm ứng đến sự biểu hiện protein tái tổ hợp Cry2Ab39 dạng tan trong tế bào <i>E. coli</i>	111
4.3. Cải biến mã di truyền gen <i>cry2Ab39</i> có nguồn gốc từ vi khuẩn để phù hợp với hệ biểu hiện của thực vật.....	114
4.4. Lựa chọn promoter phù hợp để biểu hiện gen <i>cry2Ab39opt</i> trong lá thuốc lá.....	115
4.5. Điều kiện tối ưu để biểu hiện tạm thời Cry2Ab39 trong lá thuốc lá	116
4.6. Tiềm năng ứng dụng của protein độc tố Cry2Ab39	117
KẾT LUẬN VÀ KIẾN NGHỊ	119
DANH MỤC CÔNG TRÌNH CÔNG BỐ LIÊN QUAN ĐẾN LUẬN ÁN.	121
DANH MỤC TÀI LIỆU THAM KHẢO.....	122
PHỤ LỤC.....	139

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

APS	Ammonium persulfate
bp	Base pair
<i>Bt</i>	<i>Bacillus thuringiensis</i>
DAB	3,3'-diaminobenzidine
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
EDTA	Axit etylen diamin tetra acetic
EtBr	Ethydium Bromide
gDNA	Genomic deoxyribonucleic acid
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ISAAA	International service for the acquisition of agri-biotech applications
kb	Kilo base
kDa	Kilo dalton
LB	Luria-Bertani
OD	Optical Density
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
QC	Quality control
SDS	Sodium dodexyl sulfat
SEM	Scanning electron microscope
RNA	Ribonucleic acid
TAE	Tris acetid EDTA
TEMED	Tetramethylethylenediamine

DANH MỤC CÁC BẢNG

	Trang
Bảng 1.1. Các nhóm phân loại trong hệ thống danh pháp các protein độc tố diệt côn trùng đã sửa đổi	15
Bảng 1.2. Các giống đậu tương chuyển gen kháng sâu đã thương mại hóa	26
Bảng 2.1. Các cặp mồi được sử dụng trong nghiên cứu	31
Bảng 3.1. Mật độ tế bào của một số chủng <i>Bt</i> nghiên cứu sau 72 giờ nuôi cấy	54
Bảng 3.2. Hoạt tính diệt sâu đục quả đậu tương của các chủng <i>Bt</i> nghiên cứu	55
Bảng 3.3. Kết quả đo nanodrop và Qubit các mẫu DNA	60
Bảng 3.4. Kết quả giải trình tự các chủng vi khuẩn SP14.2, Đ6.1, Đ6.2 và BD8.2 bằng hệ máy PacBio SEQUEL	62
Bảng 3.5. Kết quả lắp ráp <i>denovo</i> các chủng vi khuẩn <i>Bt</i>	63
Bảng 3.6. Danh sách các trình tự được dự đoán mã hóa protein độc tố <i>Bt</i> khai thác được từ cơ sở dữ liệu trình tự hệ gen 4 chủng <i>Bt</i> bằng công cụ BtToxin_scanner	66
Bảng 3.7. Danh sách các protein độc tố tiềm năng được dự đoán có tính mới và độ tương đồng với các protein diệt côn trùng <i>Bt</i> đã công bố	69
Bảng 3.8. Sự sai khác về trình tự nucleotide giữa các gen độc tố được phân lập với trình tự gen tiên đoán khai thác từ dữ liệu giải trình tự hệ gen	72
Bảng 3.9. Mức độ tương đồng của 6 protein độc tố đã phân lập được gen mã hóa với các protein đã công bố trên Genbank	74
Bảng 3.10. Vị trí các nucleotide và axit amin sai khác của các gen độc tố phân lập được so với gen tham chiếu	79
Bảng 3.11. Kết quả thử hoạt tính diệt sâu đục quả <i>E. zinckenella</i> Treitschke của 4 protein độc tố tái tổ hợp	87
Bảng 3.12. Giá trị liều gây chết trung bình của Trx-His-Stag-Cry2Ab đối với sâu non tuổi 2 <i>Etiella zinckenella</i> và các tham số của mô hình hồi quy Probit	93
Bảng 3.13. So sánh các thông số về trình tự giữa gen <i>cry2Ab39</i> kiểu dại và gen được cải biến <i>cry2Ab39opt</i>	97

DANH MỤC CÁC HÌNH VẼ, ĐỒ THỊ

	Trang
Hình 1.1. Vòng đời của sâu đục quả đậu tương <i>Etiella zinckenella</i> Treischke	9
Hình 1.2. Ảnh chụp SEM các tinh thể protein (hình lưỡng tháp tam giác) cùng các bào tử từ chủng <i>Bt</i> LIPMKA	13
Hình 1.3. Sơ đồ khái quát về hệ thống danh pháp 4 bậc để định danh các nội độc tố (Cry và Cyt) và các độc tố dạng tiết (Vip và Sip)	14
Hình 1.4. Phổ vật chủ của các nội độc tố dạng tinh thể δ -endotoxin có nguồn gốc từ <i>Bt</i>	18
Hình 1.5. Mô hình cấu trúc không gian ba chiều của protein độc tố Cry1A.	20
Hình 1.6. Sự thay đổi cấu trúc không gian 3 chiều trong quá trình phân cắt hoạt hóa tiền độc tố Cry1Ac	21
Hình 1.7. Cơ chế hoạt động của độc tố Cry theo mô hình tạo lỗ màng	23
Hình 2.1. Quy trình xác định các gen độc tố diệt côn trùng từ dữ liệu giải trình tự hệ gen các chủng <i>Bt</i>	38
Hình 2.2. Sơ đồ quá trình nhân dòng và thiết kế các vector biểu hiện pET32a(+) mang các gen <i>cry</i>	42
Hình 2.3. Sơ đồ quá trình thiết kế các vector chuyển gen pFGC/35S_cry2Ab39opt và pFGC/rbcS1A-pro_cry2Ab39op	49
Hình 3.1. Hình thái khuẩn lạc đặc trưng của một số chủng <i>Bt</i> được hoạt hóa trên môi trường thạch LB	53
Hình 3.2. Hình ảnh thử hoạt tính diệt sâu non đục quả đậu tương tuổi 2-3 của một số chủng <i>Bt</i> nghiên cứu..	56
Hình 3.3. Hình thái bào tử và tinh thể của các chủng <i>Bt</i> trên kính hiển vi điện tử quét với độ phóng đại 5000x	57
Hình 3.4. Đồ thị biểu diễn sự ảnh hưởng của động thái sinh trưởng đến hoạt tính diệt sâu đục quả đậu tương của 4 chủng <i>Bt</i> nghiên cứu	58
Hình 3.5. DNA tổng số tách chiết từ 4 chủng <i>Bt</i> nghiên cứu	60
Hình 3.6. Kết quả điện di và phân tích sản phẩm cắt DNA genome 4 chủng <i>Bt</i> bằng g-TUBE™ trên máy bioanalyzer	61
Hình 3.7. Kết quả phân tích thư viện DNA bằng máy bioanalyzer	61
Hình 3.8. Kết quả phân chia các trình tự mã hóa vào các hệ thống con	64
Hình 3.9. Thành phần và tỉ lệ các dạng protein độc tố <i>Bt</i> khai thác được từ cơ sở dữ liệu trình tự hệ gen 4 chủng <i>Bt</i> bằng công cụ BtToxin_scanner	66

Hình 3.10. Các vùng chức năng đặc trưng cho các nhóm protein độc tố <i>Bt</i> được phát hiện trên các trình tự ứng viên khai thác được từ cơ sở dữ liệu hệ gen của 4 chủng <i>Bt</i>	68
Hình 3.11. Kết quả điện di trên gel agarose kiểm tra các bước phân lập và nhân dòng các gen mã hóa protein độc tố <i>Bt</i>	71
Hình 3.12. Kết quả giống hàng thể hiện sự sai khác về trình tự nucleotide và axit amin giữa gen độc tố <i>cry2Ab</i> , <i>cry1Be</i> , <i>cry2Ah</i> và các gen tham chiếu	78
Hình 3.13. Kết quả điện di trên gel agarose kiểm tra các bước thiết kế các vector biểu hiện pET32a(+) mang gen <i>cry1Na</i> , <i>cry2Ab</i> , <i>cry2Ah</i> và <i>cry1Be</i>	82
Hình 3.14. Kết quả biểu hiện các gen <i>cry1Be</i> , <i>cry1Na</i> , <i>cry2Ab39</i> và <i>cry2Ah</i> trong tế bào <i>E. coli</i> BL21	84
Hình 3.15. Kết quả thử hoạt tính diệt sâu đục quả đậu tương <i>E. zinckenella</i> Treitschke của các protein tái tổ hợp	86
Hình 3.16. Phân tích độ hòa tan của protein tái tổ hợp Trx-His-Stag-Cry2Ab bằng SDS-PAGE (a) và Western blot (b)	88
Hình 3.17. Điện di SDS-PAGE phân tích sự biểu hiện của Trx-His-Stag-Cry2Ab ở các nhiệt độ khác nhau	89
Hình 3.18. Điện di SDS-PAGE phân tích sự biểu hiện của Trx-His-Stag-Cry2Ab ở các nồng độ chất cảm ứng IPTG khác nhau	90
Hình 3.19. Điện di SDS-PAGE phân tích sản phẩm tinh sạch protein tái tổ hợp Trx-His-Stag-Cry2Ab	91
Hình 3.20. Kết quả thử hoạt tính diệt sâu đục quả đậu tương <i>E. zinckenella</i> Treitschke của protein tái tổ hợp Trx-His-Stag-Cry2Ab.	92
Hình 3.21. Đường hồi quy Probit được xây dựng dựa trên số liệu về nồng độ độc tố Trx-His-Stag-Cry2Ab sử dụng và tỉ lệ sâu chết để ước tính giá trị LC50.	93
Hình 3.22. Trình tự gen đã cải biến <i>cry2Ab39opt</i> .	96
Hình 3.23. Sự phân bố tỉ lệ phần trăm các mã bộ ba phù hợp của gen <i>cry2Ab39</i> với hệ biểu hiện <i>N. benthamiana</i> trước và sau khi được cải biến mã di truyền.	97
Hình 3. 24. Kết quả thiết kế vector chuyển gen pFGC/35S_ <i>cry2Ab39opt</i>	99
Hình 3.25. Kết quả phân lập, nhân dòng và phân tích trình tự promoter <i>rbcS1A</i>	101
Hình 3.26. Kết quả thiết kế vector chuyển gen pFGC/ <i>rbcS1A</i> -pro_ <i>cry2Ab39opt</i>	102
Hình 3.27. Phân tích sự biểu hiện của gen <i>cry2Ab39opt</i> dưới sự điều khiển của promoter 35S và <i>rbcS1A</i> -pro bằng Western blot sử dụng kháng thể anti-6X His.	104

- Hình 3.28.** Phân tích ảnh hưởng của mật độ khuẩn *Agrobacterium* lây nhiễm đến biểu hiện tạm thời của gen cry2Ab39opt trong *N. benthamiana* bằng Western blot sử dụng kháng thể anti-6X His 106
- Hình 3.29.** Phân tích sự biểu hiện của protein Cry2Ab39 trong lá *N. benthamiana* sau các ngày biến nạp bằng Western blot sử dụng kháng thể anti-6X His 107
- Hình 3.30.** Phân tích kết quả tinh sạch protein Cry2Ab39 bằng SDS-PAGE (a) và lai miễn dịch Western blot (b) 108

MỞ ĐẦU

1. Tính cấp thiết của đề tài

Cây đậu tương (*Glycine max*) là một trong những cây trồng quan trọng, có giá trị kinh tế cao không chỉ ở Việt Nam mà còn đối với nhiều quốc gia trên thế giới. Hạt đậu tương có thể được dùng làm nguồn thức ăn giàu đạm cho người và gia súc, đồng thời là nguồn nguyên liệu quan trọng cho nhiều ngành công nghiệp chế biến. Ngoài ra, trồng cây đậu tương còn có tác dụng cải tạo đất, tăng năng suất các cây trồng khác nhờ hoạt động cố định N_2 của vi khuẩn *Rhizobium* cộng sinh trên rễ cây họ đậu. Ở Việt Nam, đậu tương được gieo trồng từ rất sớm trước cả cây đậu xanh và đậu đen. Tuy nhiên, với phương pháp canh tác truyền thống nhỏ lẻ, bộ giống năng suất thấp, giá thành đậu tương trong nước không có khả năng cạnh tranh với đậu tương nhập khẩu. Hiện sản xuất đậu tương nội địa mới chỉ đủ cung cấp cho khoảng 8–10% nhu cầu tiêu thụ trong nước, còn lại phụ thuộc đến 90% nguồn nguyên liệu đậu tương nhập khẩu, đa phần để chế biến thức ăn chăn nuôi. Một trong những nguyên nhân ảnh hưởng lớn đến năng suất đậu tương là sâu bệnh, trong đó sâu đục quả *Etiella zinckenella* Treitschke là đối tượng gây hại nghiêm trọng và khó phòng trừ. Do giai đoạn sâu non sống trong quả đậu nên giải pháp dùng thuốc trừ sâu hóa học là không hiệu quả và có nguy cơ tích tụ hàm lượng lớn thuốc bảo vệ thực vật trong hạt đậu tương. Vì vậy, việc nghiên cứu tìm ra các thuốc trừ sâu vi sinh có hoạt lực cao trong diệt trừ sâu đục quả đậu tương cũng như tạo các giống đậu tương chuyển gen kháng sâu đục quả là một giải pháp tiềm năng.

Việc sử dụng các tác nhân sinh học (vi khuẩn, vi nấm) hay các hợp chất tự nhiên có hoạt tính sinh học để phòng trừ hiệu quả sâu bệnh cho cây trồng đã được tiến hành rộng rãi trong nhiều thập kỉ qua. Trong đó, phổ biến nhất là thuốc trừ sâu vi sinh *Bt* (*Bacillus thuringiensis*), chiếm tới 90% thị trường thuốc trừ sâu sinh học trên thế giới nhờ khả năng diệt côn trùng phổ rộng và hiệu quả cao, nhưng vẫn đảm bảo thân thiện với môi trường và an toàn với con người cũng như các sinh vật không chủ đích khác. Ngoài ra, các gen mã hóa protein độc tố diệt côn trùng trong hệ gen vi khuẩn *Bt* đã được phân lập và ghép vào hệ gen thực vật để tạo ra các giống cây trồng biến đổi gen có khả năng kháng sâu bệnh. Việc sử dụng cây trồng

Bt trong nông nghiệp mang lại nhiều lợi ích, bao gồm việc kiểm soát côn trùng gây hại hiệu quả hơn, giảm thiểu sử dụng thuốc trừ sâu hóa học, tạo điều kiện và duy trì quần thể thiên địch trong các khu vực canh tác và cho phép thực hành nông nghiệp bền vững hơn. Với cây đậu tương, trên thế giới hiện nay có 6 giống đậu tương chuyển gen *Bt* được phép canh tác và thương mại hóa, có thể kháng một số loại sâu hại thuộc bộ Cánh vảy như sâu ăn lá *Anticarsia gemmatalis* Hübner, *Chrysodeixis includens* Walker..., sâu đa thực *Helicoverpa armigera*. Tuy nhiên, hầu như chưa có báo cáo nào đánh giá khả năng kháng sâu đục quả *E. zinckenella* Treitschke của 6 giống đậu tương đã thương mại hóa này cũng như các giống đậu tương chuyển gen *Bt* khác đang được nghiên cứu.

Chính vì vậy, việc tìm ra các chủng *Bt* cùng nguồn gen độc tố có khả năng diệt sâu *E. zinckenella* Treitschke mang tính quan trọng và cấp thiết, tạo tiền đề cho các nghiên cứu ứng dụng công nghệ gen để phát triển giống đậu tương kháng sâu đục quả. Hiện nay, Viện Công nghệ sinh học đang lưu giữ bộ sưu tập *Bt* (VBtC) gồm hơn 3000 chủng phân lập từ 52 tỉnh thành của Việt Nam. Với ưu thế là một trong những quốc gia có tính đa dạng sinh học cao nhất thế giới, việc nghiên cứu sàng lọc và khai thác các chủng *Bt* bản địa của Việt Nam để tìm được các gen đặc hiệu diệt côn trùng đích mong muốn có triển vọng rất cao. Kết quả của những nghiên cứu này sẽ cho phép các nhà khoa học nước ta chủ động trong việc tạo được các giống đậu tương chuyển gen kháng sâu đục quả thích ứng tốt với điều kiện khí hậu đặc thù của Việt Nam, góp phần tăng năng suất và sức cạnh tranh của đậu tương nội địa.

Xuất phát từ tình hình nghiên cứu và thực tiễn nói trên, chúng tôi đã thực hiện đề tài luận án: “**Nghiên cứu phân lập và biểu hiện gen mã hóa protein độc tố diệt sâu đục quả đậu tương (*Etiella zinckenella* Treitschke) từ vi khuẩn *Bacillus thuringiensis* bản địa**”.

2. Mục tiêu nghiên cứu

Phát hiện, phân lập được gene mã hoá protein độc tố mới, có độc tính cao từ nguồn vi khuẩn *B. thuringiensis* bản địa của Việt Nam để làm vật liệu chuyển gen nhằm diệt sâu đục quả đậu tương *Etiella zinckenella* Treitschke.

3. Nội dung nghiên cứu

- i) Tuyển chọn các chủng *Bt* có khả năng diệt sâu đục quả đậu tương từ các chủng *Bt* trong bộ sưu tập *Bt* bản địa Việt Nam.
- ii) Giải trình tự và khai thác dữ liệu trình tự hệ gen của một vài chủng *Bt* bản địa có hoạt tính diệt sâu đục quả đậu tương với hiệu quả $\geq 85\%$ để tìm ra các gen mã hóa protein độc tố mới tiềm năng.
- iii) Phân lập gen mã hóa protein độc tố *Bt* mới có tiềm năng diệt sâu đục quả *E. zinckenella* Treitschke từ các chủng *Bt* bản địa có hoạt tính $\geq 85\%$.
- iv) Biểu hiện gen mã hóa protein độc tố *Bt* mới trong tế bào *Escheriachi coli* và đánh giá khả năng diệt sâu đục quả đậu tương *E. zinckenella* Treitschke của protein tái tổ hợp.
- v) Cải biến mã di truyền và tối ưu hóa điều kiện biểu hiện gen mã hóa độc tố *Bt* mới trong hệ biểu hiện thực vật.

4. Ý nghĩa khoa học và thực tiễn của đề tài

➤ Ý nghĩa khoa học

Kết quả nghiên cứu đã tìm ra các gen mã hóa protein diệt sâu mới từ các chủng *Bt* bản địa của Việt Nam, góp phần làm giàu cơ sở dữ liệu về các protein độc tố có nguồn gốc từ vi khuẩn *Bt* đã được phát hiện và công bố.

Luận án đã khẳng định khả năng diệt hiệu quả ấu trùng sâu đục quả đậu tương *Etiella zinckenella* Treitschke của protein độc tố Cry2Ab39 mới phát hiện, đồng thời chứng minh chiến lược phù hợp để biểu hiện gen mã hóa độc tố này trong thực vật. Đây là cơ sở khoa học cho hướng tiếp cận về thiết kế các cấu trúc phục vụ chuyển gen vào cây đậu tương tăng khả năng kháng sâu đục quả.

Các trình tự gen công bố trên ngân hàng GenBank cùng hai bài báo đăng tải trên các tạp chí khoa học quốc tế là những tư liệu có giá trị tham khảo trong nghiên cứu và giảng dạy.

➤ Ý nghĩa thực tiễn

Chủng *Bt* SP14.2 cùng gen độc tố *cry2Ab39* là nguyên liệu tiền đề cho các nghiên cứu tạo và phát triển các giống đậu tương biến đổi gen có khả năng kháng sâu đục quả *E. zinckenella* Treitschke.

5. Những đóng góp mới của luận án

- Hệ gen 4 chủng *Bt* bản địa của Việt Nam đã được giải trình tự, lắp ráp và chú giải. Từ kết quả khai thác dữ liệu hệ gen 4 chủng này, đã tìm được 4 gen mã hóa độc tố diệt côn trùng được dự đoán có tính mới thuộc hạng 4, bao gồm 2 gen thuộc nhóm *cry1* (*cry1Na*, *cry1Be*) và 2 gen thuộc nhóm *cry2* (*cry2Ab*, *cry2Ah*).
- Đã phân lập, nhân dòng và biểu hiện gen mã hóa độc tố *Bt* mới từ chủng *Bt* SP14.2 thuộc bộ sưu tập *Bt* bản địa của Việt Nam. Trình tự gen này được công bố trên ngân hàng GenBank với mã số MN319700.1, đồng thời protein độc tố mới đã được Hội đồng danh pháp các độc tố có nguồn gốc từ *Bt* định danh là Cry2Ab39. Gen *cry2Ab39* có kích thước 1899 bp, giống 99,05% về trình tự nucleotide với gen tham chiếu *cry2Ab3* (Genbank AF164666.1). Trong khi đó, protein Cry2Ab39 gồm 633 amino axit do gen này mã hóa và protein tham chiếu Cry2Ab3 có độ tương đồng 99,21%. Chủng *Bt* SP14.2 mang gen *cry2Ab39* đã được Cục sở hữu trí tuệ Việt Nam cấp bằng độc quyền sáng chế số 37733 ngày 30/10/2023.
- Protein Cry2Ab39 tái tổ hợp được biểu hiện và tinh sạch từ tế bào *E. coli* BL21 đã được chứng minh có khả năng diệt ấu trùng sâu đục quả đậu tương *Etiella zinckenella* Treitschke với liều gây chết 50% sau 7 ngày là $LC_{50} = 1,74 \mu\text{g/g}$ thức ăn.

Chương 1. TỔNG QUAN NGHIÊN CỨU

1.1. Đậu tương và côn trùng gây hại

1.1.1. Tầm quan trọng của cây đậu tương

Đậu tương (*Glycine max* (L.) Merrill), thuộc họ đậu (Fabaceae), là cây trồng quan trọng đứng vị trí thứ tư trong số các cây lương thực thực phẩm sau lúa mì, lúa nước và ngô [1]. Đậu tương ban đầu được thuần hóa cách đây hơn ba nghìn năm ở vùng Đông Bắc Trung Quốc. Trong nhiều thế kỷ, việc sản xuất, chế biến và tiêu thụ loại cây này vẫn tập trung ở Trung Quốc và Đông Á, nhưng ngày nay đậu tương đã trở thành mặt hàng toàn cầu - với 170 quốc gia giao dịch loại cây này trên thế giới với tổng giá trị ước tính là 58 tỷ USD vào năm 2017 [2].

Khó có thể tìm được loại cây trồng nào lại có tác dụng nhiều mặt và hiệu quả kinh tế cao như cây đậu tương. Về thực phẩm, hạt đậu tương có thành phần dinh dưỡng cao, chứa khoảng 40% protein và 21% dầu trên tổng khối lượng khô. Protein của đậu tương có phẩm chất tốt nhất trong các protein của thực vật, có đầy đủ và cân đối các loại axit amin cần thiết, hoàn toàn có thể thay thế đạm động vật trong bữa ăn hàng ngày của con người. Protein của đậu tương lại dễ tiêu hoá và không có các thành phần tạo thành cholesterol. Lipit của đậu tương chứa tỷ lệ lớn các axit béo chưa no, có hệ số đồng hoá lớn (98%), chỉ số iốt cao (120-137) có tác dụng phòng chống bệnh bướu cổ cho người, đặc biệt đối với vùng trung du và miền núi. Hạt đậu tương còn chứa nhiều loại muối khoáng và có khả năng cung cấp năng lượng khá lớn (4.710 kcal/kg), cho nên người ta đã chế biến hạt đậu tương thành hơn 600 sản phẩm khác nhau. Mặt khác sử dụng protein và lipit đậu tương còn có tác dụng phòng chữa bệnh đái tháo đường, béo phì, huyết áp cao, chảy máu não... Ngày nay người ta mới biết thêm trong hạt đậu tương có chất lecithin có tác dụng làm cơ thể trẻ lâu, tăng trí nhớ và tái sinh các mô, làm cứng xương, tăng sức đề kháng. Ngoài ra nó còn chứa nhiều loại vitamin, là vị thuốc để chữa bệnh, có tác dụng tốt cho tim, gan, thận, dạ dày... [3]. Sản phẩm từ cây đậu tương được sử dụng rất đa dạng như dùng trực tiếp hạt thô hoặc chế biến thành đậu phụ, ép thành dầu đậu nành, nước tương, làm bánh kẹo, sữa đậu nành... Trong chăn nuôi, đậu tương là nguồn cung cấp protein lớn nhất cho gia súc trên toàn cầu [4]. Hiện nay, đậu tương đóng góp 50%

lượng dầu ăn toàn cầu và khoảng 2/3 lượng protein thực vật trên thế giới để làm thức ăn cho người và vật nuôi [5]. Đậu tương đã được khuyến nghị đưa vào chương trình lương thực quốc gia của 13 nước như Mỹ, Nhật Bản, Hàn Quốc, Canada, Anh, Úc...

Về mặt công nghiệp, những sản phẩm của đậu tương là nguyên liệu của nhiều ngành công nghiệp chế biến thực phẩm, công nghiệp ép dầu, công nghiệp sử dụng trực tiếp hoặc gián tiếp nguyên liệu từ đậu tương như như: chế biến cao su nhân tạo, mực in, xà phòng, chất dẻo, tơ nhân tạo, chất đốt lỏng, dầu bôi trơn trong ngành hàng không... Trong lĩnh vực trồng trọt, đậu tương từ lâu đã được coi là thành phần quan trọng trong hệ thống canh tác cây trồng ở vùng nhiệt đới và cận nhiệt đới, thích hợp cho nhiều hệ thống canh tác khác nhau. Đậu tương có khả năng tích lũy đạm từ nitơ trong tự nhiên để nuôi cây và làm giàu cho đất nhờ cộng sinh của vi khuẩn *Rhizobium* ở bộ rễ, cùng với thân và lá, các chất hữu cơ này góp phần làm thay đổi tính chất lý hoá và tăng độ phì nhiêu cho đất. Trong hệ thống thâm canh, trồng đậu tương có tác dụng làm cho cây trồng vụ sau phát triển tốt hơn, góp phần phá vỡ chu kỳ sâu bệnh, chóng nạn ô nhiễm do lạm dụng bón phân hoá học và thuốc trừ sâu. Ngoài ra lá, rễ và thân của cây đậu tương để lại trong đất được phân giải sẽ tăng mùn giúp cho cải tạo đất. Cây đậu tương có thể trồng trên nhiều loại đất, nhiều vụ trong năm, có thể xen canh, gối vụ rất thuận lợi [3].

1.1.2. Tình hình sản xuất đậu tương trên thế giới và Việt Nam

Do có khả năng thích nghi rộng với các điều kiện khí hậu và sinh thái khác nhau nên đậu tương được trồng rộng rãi trên cả năm châu lục, tập trung nhiều nhất ở châu Mỹ, kế tiếp là châu Á. Mức tăng trưởng đột biến trong sản xuất đậu tương toàn cầu được cho là kết quả của hai nguyên nhân chính là việc mở rộng diện tích canh tác và tăng năng suất trung bình nhờ những tiến bộ trong cải tạo di truyền giống và kỹ thuật canh tác. Dữ liệu thống kê cho giai đoạn 1968-2023 cho thấy sự tăng trưởng không ngừng về diện tích canh tác, tổng sản lượng và năng suất trung bình của cây đậu tương trên thế giới. Trong 55 năm, sản lượng đậu tương của thế giới tăng 9,8 lần (từ 40 triệu tấn lên 395 triệu tấn), năng suất trung bình tăng gần gấp đôi (từ 1,5 tấn/ha lên 2,8 tấn/ha) và diện tích toàn cầu dành cho cây trồng này tăng 4,8 lần (từ 28,8 triệu ha năm 1968, lên hơn 139 triệu ha vào năm 2023, tương ứng với

tổng diện tích lớn hơn toàn bộ Nam Phi) [2], [5], [6], [7]. Trong những năm 1980, Mỹ là nước sản xuất hơn 50% sản lượng đậu tương thế giới. Tuy nhiên, hiện nay Brazil đã trở thành các quốc gia sản xuất đậu tương hàng đầu thế giới với tổng sản lượng năm 2023 đạt 153 triệu tấn (chiếm 39% tổng sản lượng đậu tương toàn cầu). Trong khi đó, Mỹ (hơn 113 triệu tấn) và Argentina (49 triệu tấn) đứng ở vị trí thứ hai và thứ ba, chiếm lần lượt 29% và 13% tổng sản lượng đậu tương toàn cầu năm 2023. Ba nước trên cùng với Trung Quốc và Ấn Độ sản xuất gần 90% sản lượng đậu tương của thế giới [7].

Ở Việt Nam, việc canh tác đậu tương có truyền thống lâu đời, ban đầu từ giống hoang dại, sau đó được thuần hóa và được trồng như một cây có giá trị dinh dưỡng cao. Từ năm 1993 cả nước đã hình thành sáu vùng sản xuất đậu tương chủ lực bao gồm vùng Đông Nam Bộ với diện tích trồng đậu tương lớn nhất cả nước (chiếm 26,2% diện tích đậu tương của cả nước), khu vực miền núi Bắc Bộ (24,7%), vùng đồng bằng sông Hồng (17,5%), đồng bằng sông Cửu Long (12,4%). Tổng diện tích sáu vùng này chiếm 66,6% tổng diện tích đậu tương cả nước [8].

Theo Bộ Nông nghiệp và Phát triển nông thôn, đậu tương là một trong bốn loại cây trồng chủ lực, nhưng điều bất cập là diện tích gieo trồng ngày càng giảm. Trên thực tế, cây đậu tương đang bị cạnh tranh mạnh với những cây trồng khác vì năng suất ngày càng sụt giảm, đầu ra không ổn định, giá bán lại không cao. Hiện sản xuất đậu tương nội địa mới chỉ đủ cung cấp cho khoảng 8–10% nhu cầu chủ yếu để chế biến làm sữa đậu tương và các loại thực phẩm khác và phụ thuộc đến 90% nguồn nguyên liệu đậu tương từ các nguồn nhập khẩu, đa phần để chế biến thức ăn chăn nuôi. Đến năm 2023, diện tích trồng đậu tương trên cả nước chỉ còn 28 ngàn ha (giảm hơn 7,5 lần so với năm 2010), năng suất khoảng 1,6 tấn/ha, sản lượng đạt 45 ngàn tấn [8], [9]. Trong 10 năm trở lại đây, mỗi năm nước ta tiêu thụ trung bình gần 2 triệu tấn đậu tương. Khoảng 70% trong số này được sử dụng cho hoạt động ép dầu để sản xuất khô đậu tương - thành phần chính trong hỗn hợp thức ăn chăn nuôi. Phần lớn nhu cầu tiêu thụ đậu tương tại nước ta vẫn được đáp ứng bởi nguồn hàng nhập khẩu. Theo số liệu sơ bộ từ Tổng cục Hải quan Việt Nam, năm 2022, nước ta nhập khẩu 1,84 triệu tấn đậu tương, trong đó Brazil và Mỹ là hai nhà cung cấp lớn nhất. Mùa vụ 2023-2024, Việt Nam dự kiến nhập khẩu khoảng 2,6 triệu tấn

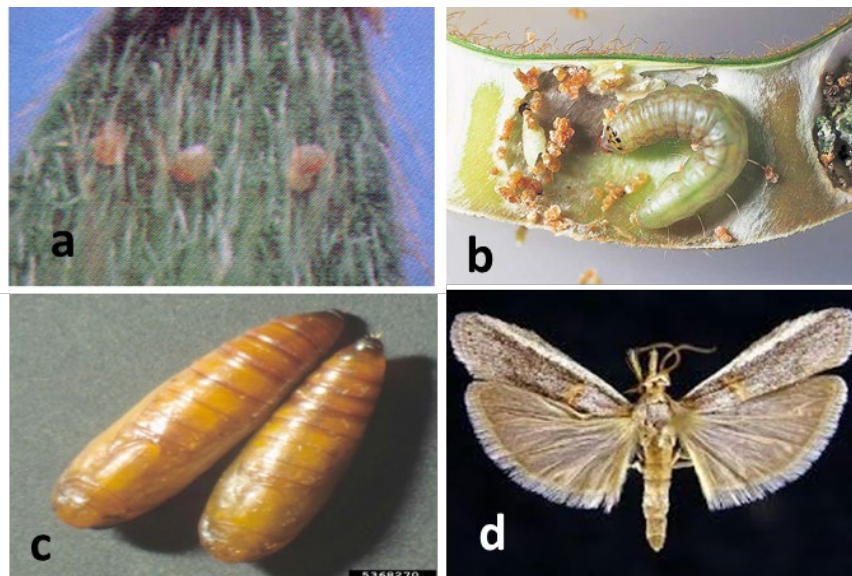
đậu tương, tăng mạnh 30% so với mùa vụ trước.

1.1.3. Sâu đục quả đậu tương *E. zinckenella* Treitschke và biện pháp phòng trừ

Trong sản xuất nông nghiệp, thiệt hại do côn trùng gây ra là một trong những yếu tố quan trọng nhất làm giảm năng suất của bất kỳ loài cây trồng nào [10]. Hầu như tất cả các bộ phận như rễ, thân, vỏ, chồi, lá, hoa và quả đều có thể bị sâu hại tấn công, phá hoại [11]. Giống như các loại cây trồng khác, cây đậu tương bị tấn công bởi nhiều loại sâu bệnh từ giai đoạn nảy mầm đến thu hoạch. Sâu hại đậu tương là nhóm đối tượng thường xuyên gây ra những thiệt hại đáng kể đối với sản xuất đậu tương. Sâu hại tất cả các bộ phận của cây, nhóm sâu hại lá, thân, quả, sâu hại trong đất. Có tới 360 loài sâu hại đậu tương ở các vùng trồng đậu tương trên thế giới. Tuy nhiên, các loài sâu hại và mức độ gây hại ở các vùng có sự biến thiên. Các loài sâu hại này được chia thành 3 nhóm chính là sâu ăn lá (sâu xanh *Helicoverpa armigera*, sâu cuốn lá đậu *Omiodes indicata*, sâu khoang *Spodoptera litura*, sâu xanh da láng *Spodoptera exigua*...), sâu đục quả (*E. zinckenella*, *Maruca vitrata*...) và sâu hại thân (sâu đục thân *Epinotia aporema*, dòi đục thân *Melanagromyza sojae*...). Trong số này, sâu đục quả đậu *E. zinckenella* Trietsche được ghi nhận là đối tượng sâu hại nghiêm trọng nhất cho đậu tương tại nhiều quốc gia, đặc biệt là vùng Đông Nam Á, gây thiệt hại nặng nề đến năng suất và giảm chất lượng hạt đáng kể nếu không phòng trừ kịp thời [12]. Ở Iran, thiệt hại năng suất do sâu đục quả đến 40%. Ở Đông Nam Á, thiệt hại do chúng gây ra ở Đài Loan là 10-15%, trong khi đó ở Indonesia lên tới 80%, ở Philippines là 57% [13]. Ở nước ta, sâu đục quả đậu tương *E. zinckenella* Treitschke là đối tượng gây hại nghiêm trọng và khó phòng trừ. Sâu đục quả phát sinh quanh năm. Ở các tỉnh phía Bắc sâu phát sinh và gây hại mạnh trên đậu tương vụ xuân (tháng 4 - 6) vụ hè (tháng 7 - 8) vụ hè thu (tháng 9 - 10). Vụ đông bị hại nhẹ hơn. Ở miền Nam sâu phát sinh và gây hại quanh năm, mạnh nhất vào tháng 4 - 5. Ngài đẻ trứng trên quả đậu là chính (73,2%), ở các bộ phận khác chỉ chiếm 11 - 20,7% số trứng. Thiệt hại do sâu đục quả tới năng suất lên đến 70-80% nếu không được phòng trừ hợp lý [14].

E. zinckenella Treitschke là loài côn trùng thuộc bộ cánh vảy Lepidoptera, họ Phycitidae, phân bố rộng rãi và gây hại trên nhiều loài ký chủ như: các cây họ đậu, cây đinh lăng, trong đó đậu tương là ký chủ ưa thích của chúng. Vòng đời của sâu

đục quả đậu tương dao động từ 32-51 ngày. Sâu phát triển qua 4 pha: trứng, sâu non, nhộng và trưởng thành (Hình 1.1). Sâu cái trưởng thành đẻ trứng trên cây sau đó sâu non đục vào vỏ và ăn hạt đang phát triển dẫn đến giảm cả chất lượng và năng suất. Triệu chứng gây hại của sâu đục quả rất rõ, thậm chí khi không có sự hiện diện của sâu non. Quả đậu xuất hiện đốm nâu, đây là vị trí mà sâu đục vào quả. Lỗ thoát ra của sâu non để vào nhộng trong đất cũng dễ dàng nhận thấy ở vỏ quả đậu. Trong quá trình gây hại sâu non thải ra phân làm cho quả đậu trở nên xốp và bị thối từng đám. Hạt bị ăn hết từng phần hoặc cả hạt, chỉ còn lại những màng như sợi tơ ở trong quả [13]. Thông thường, một sâu non có khả năng ăn hầu hết hạt trong một quả, sau đó tiếp tục đục để chui vào phá quả khác nên mỗi con có thể hại nhiều quả dẫn đến thiệt hại nghiêm trọng về năng suất [15].



Hình 1.1. Vòng đời của sâu đục quả đậu tương *Etiella zinckenella* Treischke

(<http://vitc.edu.vn/tudiennn/home/view/6091/Sau-duc-qua-dau-tuong->)

a. Trứng, b. Sâu non, c. Nhộng, d. Con trưởng thành

Do tập tính sống của sâu non là đục ăn vào trong quả nên các biện pháp kiểm soát côn trùng thông thường như phun thuốc trừ sâu thường không có hiệu quả với *E. zinckenella* Trietsche vì thuốc trừ sâu hầu như không thể tiếp cận ấu trùng bên trong vỏ quả bằng cách phun [16], [17]. Việc sử dụng các loại thuốc bảo vệ thực vật dạng nội hấp và lưu dẫn mạnh có khả năng diệt sâu đục thân, đục quả, tuy nhiên sẽ tác động xấu đến môi trường cũng như có hại đối với người phun cũng như người

tiêu thụ thực phẩm do lượng hóa chất còn tồn dư trong quả [18]. Ngoài ra, việc lạm dụng thuốc trừ sâu quá nhiều cũng dẫn đến xuất hiện các quần thể sâu hại có khả năng kháng thuốc, thậm chí kháng đa thuốc rất khó phòng trừ [19], [20].

Hiện nay, cách hiệu quả nhất để giảm thiệt hại do sâu đục quả *E. zinckenella* Trietsche gây ra cho đậu tương là sử dụng bẫy bả pheromone hoặc chọn các giống đậu tương kháng sâu bệnh [17], [21]. Dựa vào đặc tính sinh sản của hầu hết các loài bướm đêm là con cái tiết ra chất dẫn dụ giới tính (sex pheromone)- các chất dễ bay hơi để hấp dẫn và thu hút từ xa các con đực đến giao phối, người ta đã phát triển các loại bẫy pheromone nhằm mục đích dự báo tình hình và phòng trừ một số đối tượng sâu hại. Sau khi lắp đặt, bẫy pheromone sẽ giải phóng các chất mô phỏng pheromone sinh dục tự nhiên để thu hút con đực trưởng thành. Côn trùng sẽ bị mắc kẹt trong bẫy, dẫn đến giảm cơ hội giao phối và giảm số lượng cá thể trong quần thể ở thế hệ sau. Đối với loài *E. zinckenella* Trietsche, bốn chất thành phần đã được tìm thấy trong tuyến pheromone của bướm cái thuộc các quần thể ở Hungary và Ai Cập bao gồm: tetradecyl acetate, (Z)-9-tetradecenyl acetate (Z9-14:Ac), (Z)-11-tetradecenyl acetate (Z11-14:Ac) và (E)-11-tetradecenyl acetate. Các thử nghiệm bẫy pheromone phối trộn hai chất Z9-14:Ac và Z11-14:Ac cho kết quả khá tốt ở nhiều nước châu Âu, Bắc Phi và Ấn Độ. Tuy nhiên bẫy pheromone dùng môi nhử là hỗn hợp bốn chất tetradecyl/tetradecenyl acetate trên khi được thử nghiệm ở Đông Á và Úc thì hầu như không bắt được con bướm đực nào. Điều này cho thấy có sự khác biệt về thành phần các chất trong pheromone sinh dục giữa các quần thể sâu đục quả *E. zinckenella* Trietsche phân bố ở các vùng địa lý khác nhau. Một minh chứng cho giả thuyết này là kết quả nghiên cứu pheromone sinh dục của sâu đục quả *E. zinckenella* Trietsche thuộc các quần thể ở Nhật Bản. Ngoài bốn chất tetradecyl/tetradecenyl acetate nói trên, người ta đã phát hiện thêm chất decocenyl acetate E9-12:Ac, và chỉ cần bổ sung một lượng nhỏ chất này dạng tổng hợp vào môi nhử pheromone đã giúp tăng đáng kể khả năng dẫn dụ con đực [17]. Như vậy, mặc dù bẫy pheromone được đánh giá là an toàn, thân thiện với môi trường và con người nhưng để kiểm soát hiệu quả côn trùng gây hại, cần có các nghiên cứu sâu về quần thể côn trùng đích tác động, từ đó sử dụng các loại bẫy pheromone phù hợp.

Một phương pháp khác nhằm kiểm soát sâu hại mà không gây ô nhiễm môi

trường, đảm bảo quản lý bền vững hệ sinh thái là chọn lọc và phát triển các giống kháng sâu bệnh. Khả năng kháng côn trùng có thể được theo dõi thông qua việc lựa chọn các đặc điểm hình thái. Cấu trúc, chiều dài và mật độ lông tơ trên thân, lá và quả đậu được cho là ảnh hưởng rất nhiều đến khả năng kháng sâu của cây đậu tương. Theo đó, mật độ lông tơ càng cao thì mức độ cây bị sâu tấn công càng thấp và ngược lại [14]. Nhiều nghiên cứu trong và ngoài nước đã được thực hiện để đánh giá phản ứng của các giống đậu tương đối với sâu đục quả *E. zinckenella* Trietsche nhằm chọn giống có khả năng kháng với loài sâu hại này [14], [18], [22]. Tuy nhiên, chọn giống kháng sâu là công việc khó khăn và lâu dài, tốn nhiều chi phí và công sức. Các giống được chọn cần được tiếp tục nghiên cứu và khảo sát qua nhiều mùa vụ cũng như ở nhiều vùng địa lý khác nhau thì mới đưa ra được kết luận chính xác về mức độ kháng cũng như tính ổn định của việc duy trì tính kháng qua các thế hệ sau.

Cùng với những tiến bộ vượt bậc trong Công nghệ gen, cây trồng chuyển gen kháng sâu hiện đang là biện pháp kiểm soát côn trùng gây hại mùa màng được ưu tiên lựa chọn trong các hệ sinh thái nông nghiệp. Đặc trưng của chúng là có sự biểu hiện các gen mã hóa protein độc tố diệt côn trùng có nguồn gốc từ vi khuẩn *Bt*. Việc sử dụng cây trồng *Bt* trong nông nghiệp mang lại nhiều lợi ích, bao gồm việc kiểm soát côn trùng gây hại hiệu quả hơn, giảm thiểu sử dụng thuốc trừ sâu hóa học, tạo điều kiện và duy trì quần thể thiên địch trong các khu vực canh tác và cho phép thực hành nông nghiệp bền vững hơn [23]. Những nghiên cứu đầu tiên về chuyển gen *cry* có nguồn gốc từ *Bt* vào cây trồng được các nhà khoa học thực hiện vào giữa những năm 1980, và đến những năm 1990, khoai tây, ngô và bông chuyển gen *Bt* đã được phê duyệt thương mại hóa. Với đậu tương, mặc dù hiện nay diện tích canh tác các giống đậu tương biến đổi gen luôn chiếm vị trí cao so với các cây nông nghiệp khác nhưng lại chưa có nhiều giống chuyển gen *Bt* kháng sâu mà tập trung chủ yếu vào tính trạng kháng thuốc diệt cỏ. Ngoài ra, có rất ít công bố trong và ngoài nước liên quan đến nghiên cứu phát triển đậu tương chuyển gen *Bt* có khả năng kháng sâu đục quả *E. zinckenella* Trietsche. Gần đây, Phạm Lê Bích Hằng và cộng sự thuộc Viện nghiên cứu Hệ gen, Viện Khoa học và Công nghệ Việt Nam đã chứng minh sự phối hợp hai protein độc tố mã hóa bởi hai gen *cry1Aa4* và *cryIIa34* phân lập từ

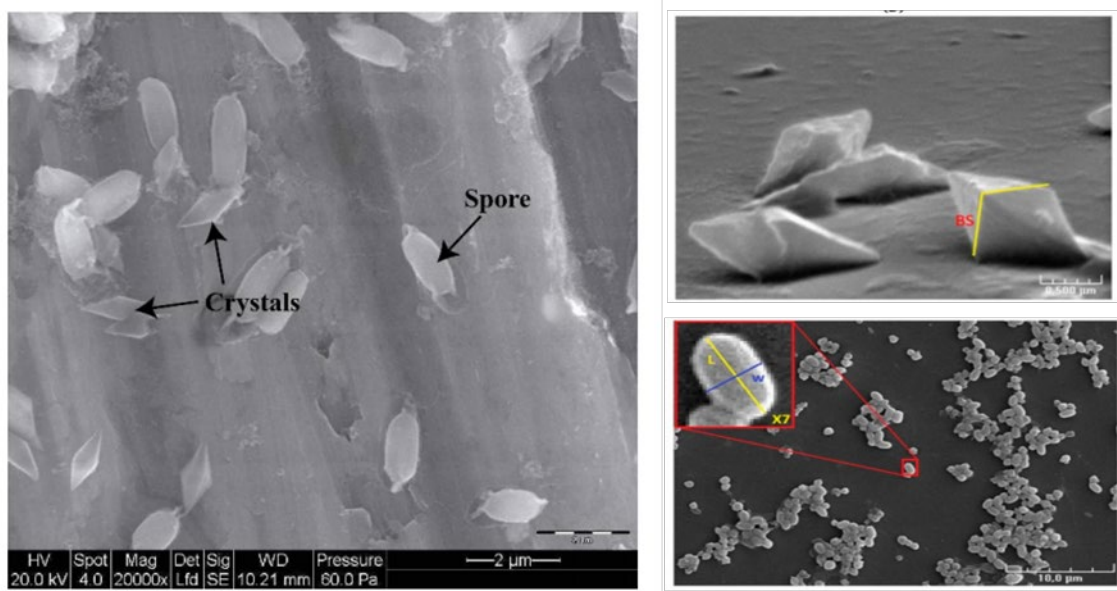
một chủng *Bt* bản địa có thể diệt hiệu quả loài *E. zinckenella* Trietsche [24]. Kết quả nghiên cứu cho thấy tiềm năng ứng dụng của hai protein độc tố này trong việc phát triển thuốc trừ sâu sinh học cũng như cây trồng chuyển gen *Bt* có khả năng kháng *E. Zinckenella*.

1.2. Protein độc tố diệt côn trùng có nguồn gốc từ *Bacillus thuringiensis*

1.2.1. Phân loại và danh pháp

B. thuringiensis (*Bt*) là vi khuẩn đất hình que, gram dương, hô hấp hiếu khí hoặc kỵ khí không bắt buộc, sinh bào tử, kích thước tế bào khoảng 0,8-1,4 μm x 2,5-10 μm , có phủ tiêm mao không dày, chuyển động được. Trong quá trình sinh trưởng và phát triển, *Bt* có khả năng sinh ra các loại protein độc tố diệt côn trùng. Chính vì điều này, vi khuẩn *Bt* thường được coi là mầm bệnh cơ hội đối với các loài sâu bọ. Việc sản sinh ra các protein độc tố diệt côn trùng được cho là mang lại lợi thế về mặt tiến hóa, cho phép tế bào vi khuẩn *Bt* có thể nảy mầm và sinh sôi nhanh chóng trong cơ thể vật chủ bị xâm nhiễm [25]. Nhân tố độc lực quan trọng nhất có khả năng diệt côn trùng ở các chủng *Bt* là các protein nội độc tố δ -endotoxin. Các độc tố này được sinh ra trong suốt pha tĩnh trong vòng đời sinh trưởng của *Bt*, phần lớn chúng tích tụ trong tế bào mẹ và kết cụm lại tạo thành các thể vùi dạng tinh thể có hình dạng không cố định có thể là hình lập phương, hình lưỡng tháp hoặc hình cầu. Các tinh thể này có thể chiếm tới 20-30% khối lượng khô của tế bào đã bào tử hóa. Khi ở trong tế bào sinh dưỡng thì bào tử và tinh thể thường nằm kề nhau, khi tế bào tan thì bào tử và tinh thể cùng thoát ra ngoài (Hình 1.2). Các nội độc tố *Bt* có thể được chia nhỏ hơn thành 2 lớp/họ là Cry (crystal protein - tinh thể độc) được mã hóa bởi các gen *cry* và Cyt (cytolytic toxin - độc tố phân giải tế bào) do gen *cyt* mã hóa [26]. Ngoài các protein nội độc tố được sinh ra trong quá trình bào tử hóa, trong suốt pha sinh trưởng sinh dưỡng, vi khuẩn *Bt* còn tổng hợp các loại độc tố khác không có cấu trúc tinh thể nhưng cũng có khả năng gây độc cho côn trùng là Vip (vegetative insecticidal protein) và Sip (secreted insecticidal proteins). Các độc tố này sau đó sẽ được tiết ra dưới dạng protein hòa tan trong môi trường nuôi cấy [26], [27].

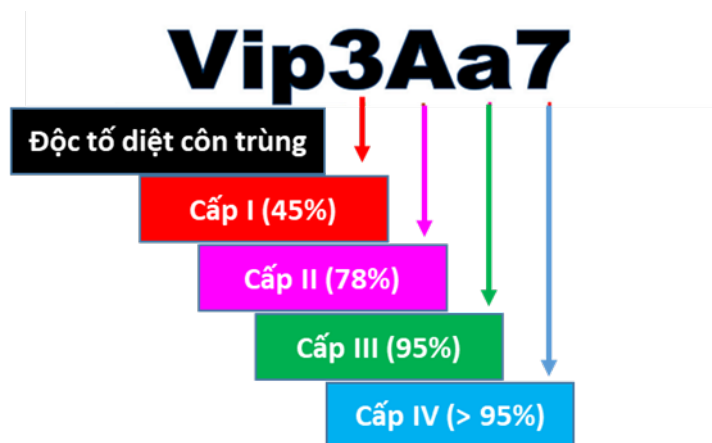
Trong giai đoạn đầu khi các gen mã hóa protein tinh thể diệt sâu *Bt* bắt đầu được phát hiện, phân lập và nghiên cứu, năm 1989, Hofte và Whiteley đã đề xuất khóa phân loại các gen *cry* thành 4 lớp chính dựa trên sự khác nhau về khối lượng protein và phổ tác dụng với côn trùng: (1) Gen *cryI* mã hóa protein có khối lượng phân tử 130-140 kDa gây độc đối với côn trùng bộ cánh vảy; (2) Gen *cryII* mã hóa tiền độc tố có khối lượng phân tử 69-71 kDa có hoạt lực với ấu trùng của bộ cánh vảy (Cry2B) và bộ hai cánh (Cry2A); (3) Gen *cryIII* mã hóa protein 73-74 kDa diệt côn trùng cánh cứng; (4) Gen *cryIV* mã hóa protein có khối lượng phân tử 135, 128, 74 và 72 kDa diệt ấu trùng thuộc bộ hai cánh [28].



Hình 1.2. Ảnh chụp SEM các tinh thể protein (hình lưỡng tháp) cùng các bào tử từ chủng *Bt* LIPMKA. [29]

Mặc dù khá hữu ích trong việc phân loại một cách có hệ thống các protein tinh thể độc đã được đặt tên tùy tiện trước đây, khóa phân loại của Hofte và Whiteley vẫn tồn tại nhiều hạn chế. Một trong số đó là việc các protein có trình tự tương đồng thường có tính đặc hiệu diệt côn trùng khác nhau, đòi hỏi chúng phải được xếp vào các nhóm phân loại chính khác nhau. Một hạn chế lớn khác là cần phải thực hiện toàn diện các thử nghiệm sinh học trước khi muốn phân loại một protein [27]. Để giải quyết những hạn chế trên, dựa trên khóa phân loại của Hofte và Whiteley, năm 1998, Crikmore và cộng sự đã lập một Hội đồng danh pháp các độc tố có nguồn gốc từ *Bt* và đưa ra một khóa phân loại mới hoàn thiện hơn dựa

trên độ tương đồng về trình tự axit amin của các protein độc tố để tiện cho việc bổ sung những độc tố *Bt* mới được phát hiện. Trong danh pháp này, tên protein độc tố sẽ được bắt đầu bằng chữ viết tắt cho họ độc tố (Cry/Cyt/Vip/Sip), theo sau là 4 kí tự đại diện cho 4 bậc phân cấp: số đếm (1,2,3...), chữ in hoa (A, B, C...), chữ thường (a, b, c...) và số đếm (1, 2, 3...) (Hình 1.3). Theo đó, các độc tố có trình tự amino axit tương đồng dưới 45% sẽ khác nhau ở bậc 1 (Ví dụ: Cry1, Cry2, Cry3), trong khi đó, các protein có độ tương đồng về trình tự nằm trong khoảng 45-78% sẽ khác nhau ở bậc 2 (Ví dụ: Cry1A, Cry1B). Bậc 3 (Ví dụ: Cry1Ab, Cry1Ac) giúp phân biệt các độc tố có độ tương đồng từ 78%-95%, và bậc 4 (Ví dụ: Cry1Ac1, Cry1Ac3) để phân biệt các độc tố có trình tự amino axit giống nhau trên 95% [30]. Khóa định danh 4 bậc này đã thể hiện tính ưu việt và được sử dụng làm quy chuẩn cho đến nay để đặt tên một protein độc tố *Bt* mới. Tính đến nay, các nghiên cứu đã phát hiện và định danh được khoảng hơn 700 loại độc tố Cry được phân thành 78 họ (Cry1-Cry78) và 40 loại độc tố Cyt thuộc 3 nhóm Cyt1-Cyt3 (www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/index.html). Đối với các độc tố pha sinh dưỡng, 15 protein Vip1, 20 protein Vip2, 111 protein Vip3 và 5 protein Vip4 đã được công bố và đặt tên [26].



Hình 1.3. Sơ đồ khái quát về hệ thống danh pháp 4 bậc để định danh các nội độc tố (Cry và Cyt) và các độc tố dạng tiết (Vip và Sip). [26]

Tuy thuận tiện trong việc định danh một lượng lớn các độc tố *Bt* đã và mới được phát hiện qua các nghiên cứu, người ta nhận thấy vẫn có điểm bất cập với cách phân loại này. Ví dụ, một số protein có độ tương đồng về trình tự axit amin rất thấp vẫn được xếp vào họ độc tố Cry như Cry6, Cry15 và Cry22. Khi ngày càng

nhiều các protein “ngoại lai” này được phát hiện và mô tả, chúng được xác định là thành viên thuộc các nhóm họ protein (Pfam) khác nhau (Ví dụ: nhóm họ protein Bin-like/ Toxin_10-like hoặc nhóm họ protein ETX/Mtx2-like là những protein có liên quan đến các độc tố kép – binary toxin (Bin) hoặc các độc tố diệt ấu trùng muỗi - mosquitocidal toxin (Mtx) sinh ra bởi vi khuẩn *Lysinibacillus sphaericus*) mặc dù đều đang được gọi là độc tố Cry. Sự gia tăng của các dự án giải trình tự hệ gen cùng những tiến bộ trong nghiên cứu cấu trúc protein đã cho thấy một thực tế là có rất nhiều loại protein diệt côn trùng có nguồn gốc từ vi khuẩn và hệ thống danh pháp hiện tại không thể hiện được sự đa dạng của chúng. Lý do này dẫn tới nhu cầu về một hệ thống phân loại phản ánh tốt hơn sự khác biệt về cấu trúc giữa các loại độc tố có nguồn gốc từ vi khuẩn đã được đặt ra. Trong công bố mới nhất vào năm 2020, Hội đồng danh pháp các độc tố có nguồn gốc từ *Bt* (nay là Bacterial Pesticidal Protein Resource Center – <https://www.bpprc.org/>) đã xây dựng nền móng của một hệ thống phân loại gồm 16 lớp dựa trên cơ sở cấu trúc của các protein độc tố (Bảng 1.1) [31]. Protein thuộc mỗi lớp phân loại trong hệ thống này đều có dạng cấu trúc điển hình riêng biệt đặc trưng cho lớp đó (Ví dụ: protein độc tố có cấu trúc gồm 3 vùng chức năng, protein trong cấu trúc có mang vùng chức năng Toxin_10/Bin-like, ETX/Mtx2-like...). Danh pháp mới được hội đồng đề xuất trong báo cáo này bao gồm: 3 chữ cái kí hiệu của lớp độc tố theo sau là 4 bậc phân loại được biểu thị bằng kí tự chữ và số.

Bảng 1.1. Các nhóm phân loại trong hệ thống danh pháp các protein độc tố diệt côn trùng đã sửa đổi. [31]

STT	Lớp độc tố diệt côn trùng	Phân loại cũ	Các vùng chức năng bảo thủ	Mô tả
1	Cry	Cry	pfam03945, pfam00555, cd04085	Các protein tinh thể phân lập từ <i>Bt</i> mà cấu trúc dạng hoạt hóa gồm 3 vùng chức năng. Ví dụ: Cry1Aa, Cry3Aa
2	Cyt	Cyt	pfam01338	Các độc tố phân giải tế bào, cấu trúc thường gồm 1 vùng chức năng đơn. Ví dụ: Cyt2Aa
3	Vip	Vip3	pfam12495, pfam02018	Các protein có cấu trúc đa vùng chức năng, ban đầu được xác định là các protein pha sinh dưỡng có hoạt tính

				diệt côn trùng. Ví dụ: Vip3Bc
4	Tpp	Cry, Bin	pfam05431	Các protein diệt sâu có chứa vùng chức năng Toxin_10 (Bin-like). Ví dụ: Tpp35Aa (tên cũ là Cry35Aa)
5	Mpp	Cry, Mtx2, Sip	pfam03318	Các protein diệt sâu thuộc họ protein ETX/Mtx2. Ví dụ: Mpp51Aa (tên cũ là Cry51Aa)
6	Gpp	Cry	pfam06355	Các độc tố diệt sâu giống Aegerolysin. Ví dụ: Gpp34Aa (tên cũ là Cry34Aa)
7	App	Cry, Pax, Xax, Yax		Các protein diệt sâu mà cấu trúc chủ yếu gồm các xoắn alpha. Ví dụ: App6Aa (tên cũ là Cry6Aa)
8	Spp		pfam01289, pfam17440	Các protein diệt sâu giống sphaericolysin
9	Mcf		pfam12920	Các protein liên quan đến độc tố diệt sâu bướm được mô tả ban đầu từ trực khuẩn phát quang <i>Photobacterium</i>
10	Mtx	Mtx1		Các protein liên quan đến độc tố Mtx1 (2VSE) phân lập được từ vi khuẩn <i>Lysinibacillus sphaericus</i>
11	Vpa	Vip2	cd00233	Các protein liên quan đến thành phần hoạt tính ADP-ribosyltransferase của các độc tố kép
12	Vpb	Vip1, Vip4	pfam07691, pfam03495, pfam17475, pfam17476	Các protein liên quan đến thành phần gắn của các độc tố kép
13	Pra	PirA		Các protein giống với thành phần độc tố A diệt côn trùng của trực khuẩn <i>Photobacterium</i>
14	Prb	PirB	pfam03945	Các protein giống với thành phần độc tố B diệt côn trùng của trực khuẩn <i>Photobacterium</i>
15	Mpf	PluMA CPF GNIP	pfam01823	Các độc tố diệt sâu thuộc siêu họ protein tạo lỗ màng hoặc phức hệ tấn công màng
16	Xpp			Protein diệt sâu hiện chưa mô tả được các đặc điểm cấu trúc

Bảng 1.1 là danh sách 16 lớp độc tố bước đầu được xác định trong hệ thống phân loại mới. Ba lớp trong số này (Cry, Cyt và Vip3), danh pháp của các độc tố *Bt*

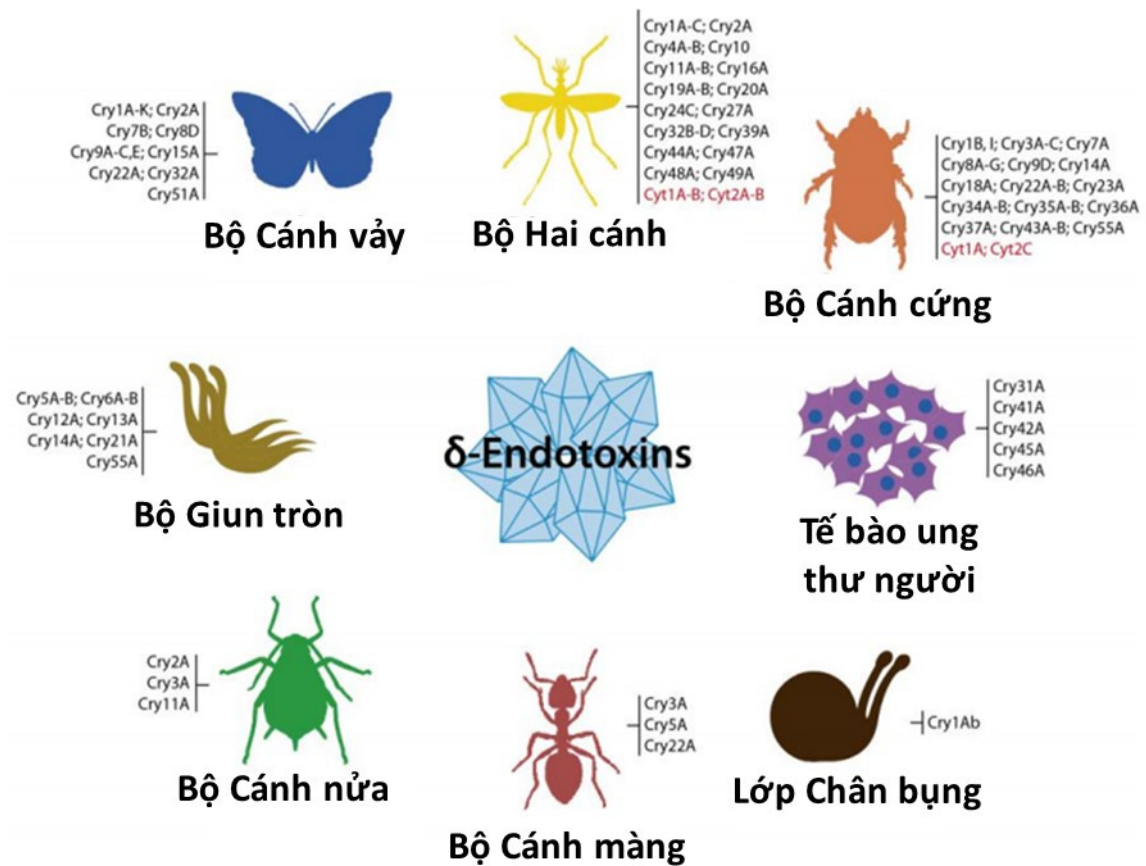
không thay đổi so với tên gọi cũ. Tuy nhiên, lớp độc tố Cry trong hệ thống phân loại mới này chỉ bao gồm các protein có cấu trúc 3 vùng chức năng. Các lớp độc tố còn lại đại diện cho các protein trước đây được gọi là độc tố Cry nhưng không có cấu trúc 3 vùng chức năng điển hình và các độc tố diệt côn trùng mới có nguồn gốc từ vi khuẩn khác không phải *Bt*. Các lớp này được phân ra dựa trên các vùng chức năng Pfam, các dạng cấu trúc đã biết và các thông tin có sẵn khác. Kí hiệu tên lớp 3 chữ cái được lựa chọn để phản ánh loại protein độc tố diệt côn trùng mà chúng đại diện (Ví dụ Mpp: lớp độc tố có chứa vùng chức năng Mtx2-like, Tpp: lớp độc tố có chứa vùng chức năng Toxin_10-like). Trường hợp các protein chưa có đủ thông tin để phân bổ chúng vào một lớp cụ thể thì sẽ được xếp vào 1 lớp tạm thời là Xpp [31].

1.2.2. Hoạt tính và phổ diệt côn trùng

Các protein tinh thể, được biết đến rộng rãi nhất là các δ -endotoxin được hình thành dưới dạng thể vùi dạng tinh thể cạnh trong quá trình phát sinh bào tử của *Bt*, bao gồm cả độc tố Cry và Cyt. Như đã đề cập ở trên, độc tố tinh thể Cry có thể được chia thành nhiều nhóm dựa trên cấu trúc phân tử của chúng. Nhóm lớn nhất được biết đến rộng rãi bao gồm các protein Cry có cấu trúc 3 vùng chức năng. Các độc tố Cry khác thuộc các họ protein khác nhau như họ protein ETX_MTX2, các độc tố kép Bin... Về phổ hoạt tính, các nghiên cứu đã ghi nhận một cách rõ ràng độc tố Cry có khả năng gây độc cho côn trùng thuộc bộ cánh vảy, bộ cánh cứng, bộ cánh nửa, bộ hai cánh, động vật không xương thuộc ngành giun tròn (bộ giun tròn Rhabditida ký sinh ở người và động vật hoặc sống tự do) và một số loài ốc [32], [33], [34], [35]. Một số protein Cry được chứng minh không gây độc cho động vật không xương sống nhưng có khả năng giết các dòng tế bào ung thư ở người một cách mạnh mẽ và đặc hiệu (khi được hoạt hóa bởi các protease) (Hình 1.4). Một số protein này thuộc nhóm độc tố Cry có cấu trúc 3 vùng chức năng (Cry31, Cry41, Cry63), số khác thuộc họ protein ETX-MTX2 (Cry45, Cry46, Cry64) [36]. Do số lượng các độc tố loại này được phát hiện ngày càng nhiều và hoạt tính gây độc tế bào đáng kể nên vào năm 2006 người ta đã đưa ra thuật ngữ riêng là “parasporin” để chỉ các protein tinh thể hình thành xung quanh bào tử ở vi khuẩn *Bt* và các vi

khuẩn khác mà không gây độc cho côn trùng nhưng có khả năng diệt tế bào ung thư [37].

Độc tố Cyt cũng hình thành thể vùi như độc tố Cry, nhưng tính tương đồng về trình tự axit amin giữa độc tố Cyt và độc tố Cry hầu như không đáng kể. Độc tố Cyt đầu tiên được Waalwijk phát hiện là Cyt1Aa1 vào năm 1985. Hiện nay người ta đã phát hiện được hơn 40 độc tố Cyt thuộc 3 nhóm là Cyt1, Cyt2 và Cyt3, các độc tố này có khối lượng 25-30 kDa. Các độc tố Cyt được phát hiện chủ yếu thuộc các



Hình 1.4. Phổ vật chủ của các nội độc tố dạng tinh thể δ -endotoxin có nguồn gốc từ *Bt*. [27]

phân loài *B. thuringiensis israelensis*, *morrisoni*, *medellin*, *neolonensis*, *kyushuensis*, *damstradiensis*, *fuokukaensis*, *tenebrionis*. Độc tố Cyt tạo thành một nhóm nhỏ các protein dạng tinh thể riêng biệt với hoạt tính diệt một số loại ấu trùng bộ hai cánh, đặc biệt là muỗi và ruồi đen [38], [39], [40]. Tuy nhiên một số chủng thuộc nhiều phân loài *Bt* như *Bt morrisoni* có mang các gen *cyt* có hoạt tính gây độc cho một dải rộng côn trùng bao gồm Bộ hai cánh, Bộ cánh vảy và Bộ cánh cứng

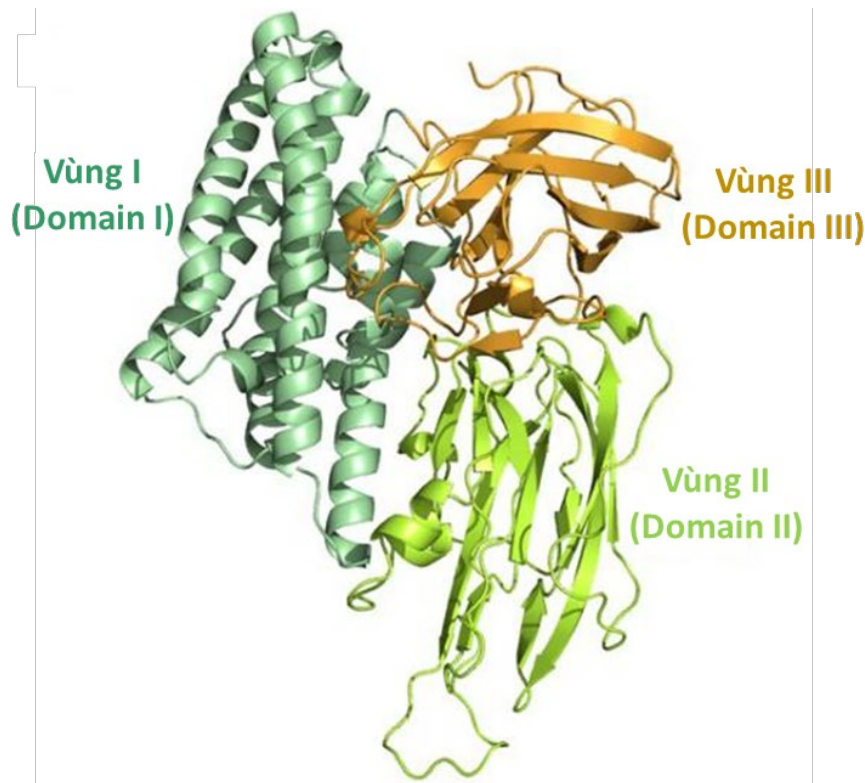
[41] (Hình 1.4). Một điểm đặc biệt khác của các protein Cyt là khả năng hiệp lực với các độc tố Bt khác bao gồm Cry và Vip3 trong hoạt động diệt sâu, đồng thời làm giảm mức độ kháng của một số loài côn trùng đối với một vài độc tố Cry [42], [43]. Ví dụ: Cyt1Aa có thể ngăn chặn khả năng kháng lại độc tố Cry11Aa ở ấu trùng quần thể muỗi *Culex quinquefasciatus* trong phòng thí nghiệm. Sự liên kết độc tố Cry11Aa với Cyt1Aa tạo điều kiện thuận lợi cho cả quá trình oligome hóa độc tố Cry11Aa lẫn việc hình thành lỗ màng, và đây được đề xuất là cơ chế của sự hiệp lực [42]. Ngoài ra, các độc tố Cyt1Ab và Cyt2Ba sinh ra bởi *Bt medellin* và *Bt israelensis* giúp tăng cường hoạt tính diệt ấu trùng muỗi *Aedes aegypti* và *C. quinquefasciatus* của *L. sphaericus* [44].

Các protein Vip phân bố rộng rãi giữa các loài *Bacillus*. Các protein này không được sinh ra ở dạng thể vùi tinh thể mà được tiết vào môi trường nuôi cấy trong giai đoạn phát triển sinh dưỡng của vi khuẩn. Giống như các độc tố Bt khác, thể nguyên bản của các protein Vip cũng là dạng bất hoạt, chúng chỉ được hoạt hóa sau khi được tiết ra ở màng tế bào ruột giữa của côn trùng thông qua hoạt động của các enzyme. Hiện tại, Hội đồng Danh pháp độc tố Bt đã xác định và phân loại các protein Vip thành bốn họ khác nhau là Vip1, Vip2, Vip3 và một họ protein Vip mới là Vip4. Protein Vip1 và Vip2 tạo thành một độc tố kép có hoạt tính cao trong diệt một số côn trùng hại thuộc bộ Cánh cứng và loài côn trùng chích hút nhựa cây *Aphis gossypii* thuộc bộ Cánh nửa. Trái lại, protein Vip3 là các độc tố đơn chuỗi (không phải độc tố kép) có hoạt tính diệt phổ rộng nhiều côn trùng thuộc bộ Cánh vẩy. Đặc biệt Vip3 được chứng minh có khả năng gây độc cho một số loài côn trùng ít mẫn cảm với các protein Cry1A (Ví dụ: *Agrotis ipsilon*, *Spodoptera exigua* và *S. frugiperda*). Riêng phân họ Vip4 đến nay vẫn chưa được xác định rõ phổ vật chủ [26], [27].

1.2.3. Cấu trúc protein Cry

Trong số các protein diệt côn trùng có nguồn gốc từ Bt, độc tố Cry có cấu trúc 3 vùng chức năng là nhóm độc tố quan trọng nhất, được phát hiện sớm và nghiên cứu chi tiết nhất. Những tiến bộ kỹ thuật trong sinh học phân tử, tinh thể học tia X và cộng hưởng từ hạt nhân (NMR) đã tạo ra một cuộc cách mạng trong nghiên cứu cấu trúc protein ở cấp độ phân tử giúp làm sáng tỏ các chức năng của chúng.

Cấu trúc của protein tinh thể đã được làm sáng tỏ nhờ kỹ thuật nghiên cứu tinh thể bằng tia X với những nghiên cứu đầu tiên về cấu trúc Cry3A và Cry1Aa, sau đó được mở rộng cho các độc tố khác. Protein tinh thể độc tố thường tồn tại ở dạng tiền độc tố và có cấu trúc đặc thù. Đặc điểm điển hình trong cấu trúc của protein tinh thể độc tố là có các vùng bảo thủ nằm xen kẽ các vùng biến đổi. Các vùng bảo thủ này đóng vai trò quan trọng về mặt cấu trúc và chức năng. Cấu trúc chung của protein tinh thể độc tố Cry gồm 3 vùng (Domain) riêng biệt [45], [46] (Hình 1.5).

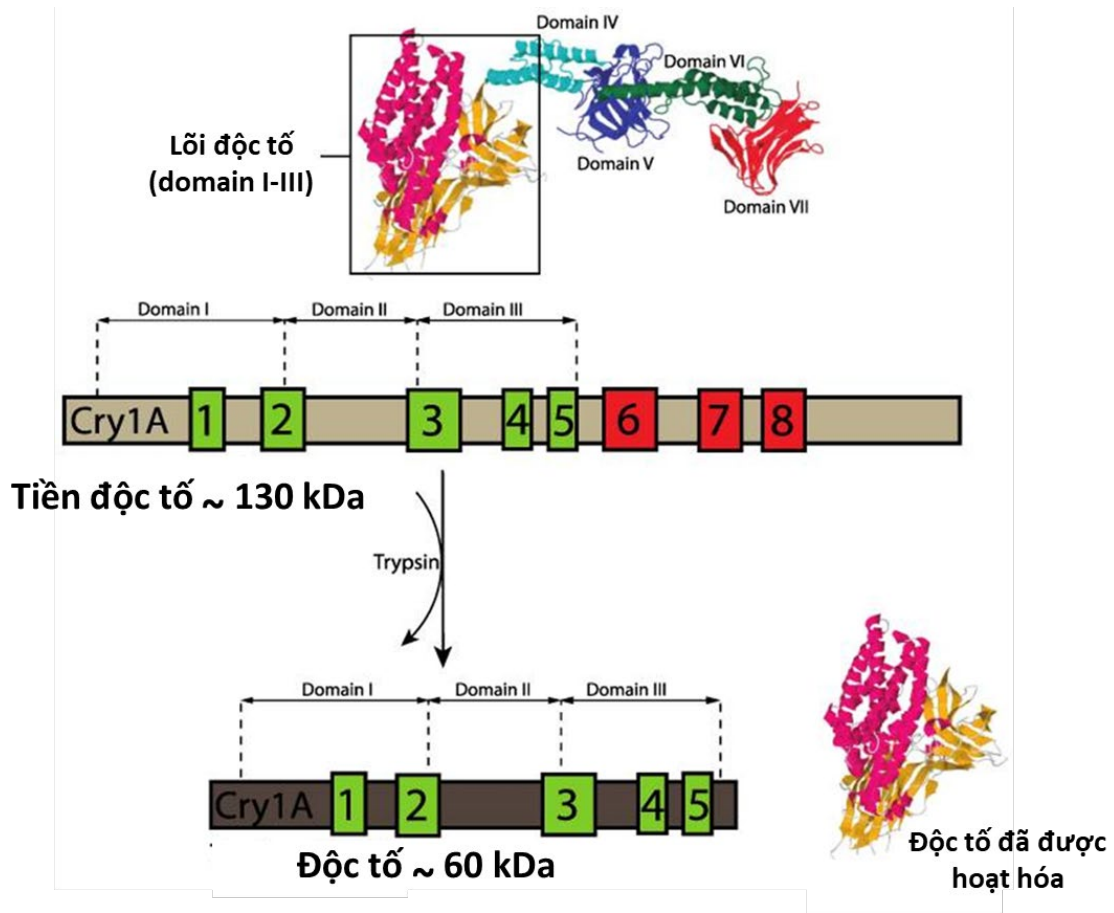


Hình 1.5. Mô hình cấu trúc không gian ba chiều của protein độc tố Cry1A.

[37]

Thứ tự các vùng được tính từ đầu N đến đầu C của chuỗi polypeptid. Vùng I, hay vùng tạo lỗ, nằm ở đầu N có độ bảo thủ cao, được tạo thành bởi 1 bó gồm 7 chuỗi xoắn α đối song song (α antiparallel) trong đó chuỗi số 5 được bao bọc bởi các chuỗi còn lại. Trong quá trình hoạt hóa của tất cả các độc tố Cry có cấu trúc 3 vùng chức năng thì vùng I sẽ chịu sự phân cắt bởi các protease và được cho là có vai trò chèn vào màng và tạo lỗ màng [37], [47]. Vùng II (còn gọi là vùng trung tâm hay vùng giữa) là vùng siêu biến bao gồm 3 tấm β đối song song tạo thành một dạng hình học topo điển hình gọi là “Greek key”, kiểu sắp xếp này cũng được gọi là

cuộn lăng kính β (β prism fold). Vùng II được cho là đóng vai trò quan trọng trong tương tác độc tố - thụ thể [37], [48]. Cuối cùng, vùng III (còn gọi là vùng gắn galactose) nằm ở đầu C, có tính bảo thủ không hoàn toàn, gồm các tấm β đối song song tạo thành một kẹp β (β sandwich) mà hình học topo gọi là “jelly roll”. Về chức năng, vùng này liên quan đến việc gắn thụ thể và hình thành lỗ màng [37].



Hình 1.6. Sự thay đổi cấu trúc không gian 3 chiều trong quá trình phân cắt hoạt hóa tiền độc tố Cry1Ac [27]

(Phần lõi độc bao gồm 3 vùng chức năng I, II và III (phần được đóng khung). 5 hộp màu xanh lá cây biểu thị 5 vùng trình tự axit amin bảo thủ nằm trong lõi độc tố được mô tả bởi Hofte và Whiteley, 3 hộp màu đỏ biểu thị 3 vùng trình tự nằm ngoài phần lõi độc được phát hiện bởi Schnef và cộng sự)

Các nghiên cứu về cấu trúc của tiền độc tố Cry1Ac cho thấy phần lõi độc tố cuộn gấp lại thành 3 vùng trong khi phần mở rộng (pro-region) tạo thành thêm 4 vùng chức năng nữa. Vùng IV và vùng VI có dạng xoắn α trong khi vùng V và VII

có cấu trúc cuộn xoắn β . Trong cấu trúc tạo tinh thể, vùng lõi độc của một phân tử độc tố này được nâng đỡ bởi vùng mở rộng của một phân tử khác. Thành phần axit amin ở vùng mở rộng này rất giàu cysteine giúp tạo các cầu nối disulphide liên phân tử, hình thành cấu trúc thể vùi dạng tinh thể. Tuy nhiên, các liên kết disulphide liên phân tử này kém bền hơn dạng liên kết nội phân tử, và đặc biệt kém bền trong điều kiện axit hoặc bazơ. Điều này giải thích cho việc các tinh thể protein độc Bt sẽ bị hòa tan trong môi trường ở ruột côn trùng [27].

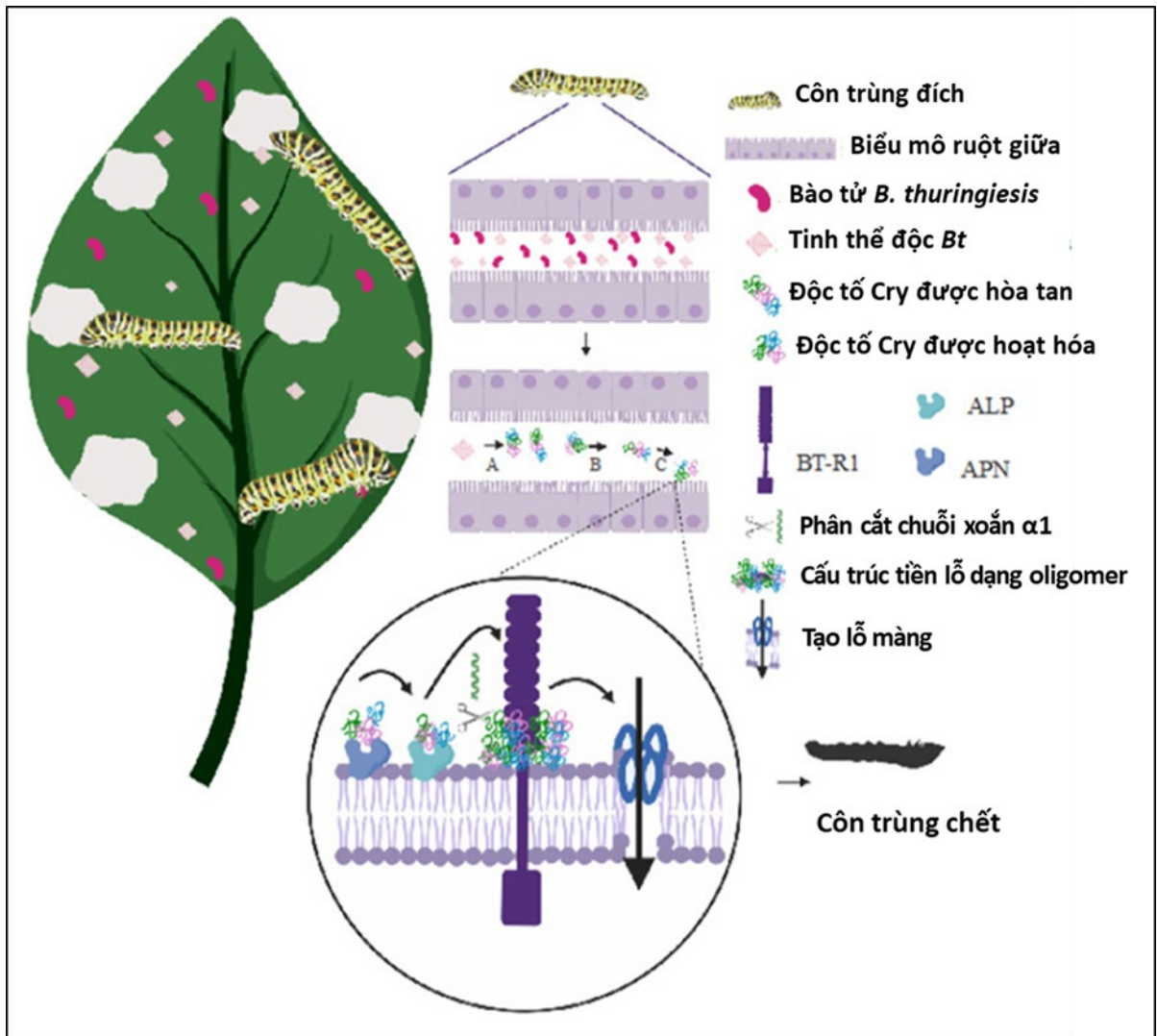
Kết quả so sánh nhiều trình tự của các độc tố Cry khác nhau cho thấy có nhiều nhất 5 khối trình tự bảo thủ nằm trong phần lõi độc có hoạt tính (3 vùng chức năng I, II và III) của các tiền độc tố và thêm 3 vùng axit amin bảo thủ nằm ngoài phần lõi độc về phía tận cùng đầu C của protein tiền độc tố [49]. Phần lõi độc của các tiền độc tố sẽ được giải phóng nhờ sự phân cắt của các protease trong ruột côn trùng tạo thành độc tố dạng hoạt động, được gọi là quá trình hoạt hóa protein (Hình 1.6). Một số ít tiền độc tố (Ví dụ: Cry3 và Cry11) cấu trúc không có vùng đầu C mở rộng này, do vậy chúng có kích thước ngắn hơn, xấp xỉ 70 kDa

1.2.4. Cơ chế diệt côn trùng

Trong số các protein tinh thể Cry đã được phát hiện và công bố, độc tố Cry có cấu trúc 3 vùng chức năng (3d-Cry) chiếm phần lớn và được nghiên cứu, mô tả sâu rộng nhất. Các tiền độc tố 3d-Cry được sản sinh ra có 2 dạng: tiền độc tố lớn có khối lượng phân tử khoảng 130 kDa (như Cry1Aa) và tiền độc tố ngắn trong khoảng 65-70 kDa (như Cry11Aa). Khi vào ruột giữa của côn trùng, tiền độc tố lớn sẽ được xử lý ở cả đầu C và đầu N bởi các protease, trong khi các tiền độc tố ngắn chỉ bị phân cắt ở đầu N. Ở cả hai trường hợp, kết quả xử lý của protease đều tạo ra độc tố Cry dạng hoạt hóa (còn gọi là lõi kháng protease) vẫn giữ nguyên cấu trúc 3 vùng chức năng với kích thước xấp xỉ 60-79 kDa. Đây chính là phân đoạn chịu trách nhiệm gây độc cho ấu trùng côn trùng. Tuy nhiên, quá trình hoạt hóa diễn ra không chính xác hoặc độc tố bị phân hủy bởi các protease khác sẽ làm giảm độc tính của các protein Cry [50].

Mặc dù cơ chế diệt côn trùng của các độc tố Cry đã được nghiên cứu rộng rãi, vẫn còn tồn tại nhiều tranh cãi. Cơ chế diệt ấu trùng côn trùng của các độc tố Cry được chấp nhận rộng rãi nhất cho tới nay là mô hình tạo lỗ màng được tóm lược trong

hình 1.7. Một khi ấu trùng ăn phải tinh thể độc tố, những tinh thể này hòa tan ở độ pH cực cao và bị phân giải bởi các protease trong ruột giữa ấu trùng tạo thành độc tố dạng hoạt hóa khoảng 60-70 kDa. Độc tố đã được kích hoạt vượt qua phức mạc tiếp xúc với các vi nhung mao ở ruột giữa, nhận diện và liên kết với các thụ thể trên màng để hình thành lỗ màng; do đó, tính đặc hiệu là rất quan trọng đối với độc tính của các protein Cry [39].



Hình 1.7. Cơ chế hoạt động của độc tố Cry theo mô hình tạo lỗ màng. [50]
(Khi ấu trùng ăn độc tố, các tinh thể tiền độc tố được hòa tan và phân cắt trong môi trường ruột giữa. Độc tố Cry nhận diện các thụ thể trên màng như APN, ALP và EC12/BT-R1. Độc tố Cry bị phân cắt chuỗi xoắn $\alpha 1$, dẫn đến hình thành cấu trúc oligomer tiền lỗ. Sau đó, cấu trúc oligomer được chèn vào màng tế bào tạo ra lỗ rỗng ion dẫn đến mất cân bằng thẩm thấu, ly giải tế bào và côn trùng bị chết. A, Sự hòa tan. B, Hoạt hóa bằng cách phân giải protein. C, Nhận diện thụ thể trên màng.)

Trong mô hình tạo lỗ màng, Vùng II được cho là đóng vai trò chính trong quá trình nhận diện và liên kết thụ thể, trong khi đó vùng I có nhiệm vụ chèn vào màng và hình thành lỗ màng. Cụ thể, các vòng (loop) có tính biến đổi cao lộ ra ở vùng II tham gia vào việc liên kết với các thụ thể thông qua 2 bước: (i) nhận biết các thụ thể aminopeptidase N (APN) và alkaline phosphatase (ALP) và sự hình thành liên kết yếu với độc tố Cry, tạo ra một phản ứng thuận nghịch [27], [51]. (ii) Nhận diện vùng ngoại bào (ectodomain) gồm 12 axit amin (EC12) của thụ thể cadherin (BT-R1) và hình thành một liên kết không đảo nghịch (Kd 19 nM). Các motif trình tự bảo thủ gần các đầu N và C của vùng ngoại bào này được chứng minh là rất quan trọng cho việc liên kết với độc tố trong tế bào côn trùng [51].

Sự tương tác giữa độc tố Cry và thụ thể BT-R1 tạo điều kiện thuận lợi cho sự phân cắt chuỗi xoắn $\alpha 1$ nằm ở tận cùng đầu N của vùng I, dẫn đến kích hoạt quá trình oligomer hóa hình thành cấu trúc oligomer tiền lỗ (pre-pore oligomer structure) có ái lực rất cao với các thụ thể màng APN và ALP [52], [53], [54]. Cấu trúc oligomer sau đó chèn vào màng tế bào tạo ra một lỗ rò ion dẫn đến tình trạng mất cân bằng thẩm thấu, ly giải tế bào ruột và côn trùng bị chết [45].

Trong quá trình nhận diện và gắn kết thụ thể, vùng III đóng vai trò quan trọng trong việc ổn định độc tố. Vùng này liên kết với N-acetylgalactosamine (GalNAc) trong thụ thể APN [55], [56]. Ngoài các thụ thể nêu trên, một số kênh protein vận chuyển cần ATP (ATP-binding cassette transporters- ABC proteins), đặc biệt là ABCC2- một thành viên thuộc phân họ C, cũng được chứng minh có liên quan đến độc tính của độc tố Cry. Các báo cáo cho thấy những protein này có thể giúp độc tố Cry1A liên kết với các thụ thể và chèn cấu trúc oligomer vào màng tế bào [57].

1.3. Một số thành tựu về tạo cây đậu tương chuyển gen kháng sâu

1.3.1. Tình hình nghiên cứu và lưu hành các giống đậu tương biến đổi gen kháng sâu hại trên thế giới

Việc ứng dụng công nghệ sinh học để tạo ra cây trồng biến đổi gen, mang gen kháng côn trùng có nguồn gốc từ *Bt* góp phần tăng năng suất, hạn chế việc sử dụng thuốc trừ sâu hóa học, tạo điều kiện canh tác an toàn bền vững. Những thập kỷ đầu tiên của công nghệ tạo cây trồng biến đổi gen đã khẳng định hiệu quả và sự cần

thiết của công nghệ đối với an ninh lương thực và bảo vệ môi trường. Trong số các cây trồng biến đổi gen, đậu tương luôn chiếm vị trí cao nhất về diện tích canh tác, với 50% tổng diện tích cây trồng biến đổi gen toàn cầu vào năm 2017 [58]. Mặc dù vậy, các giống đậu tương biến đổi gen có khả năng kháng sâu chưa nhiều mà tập trung chủ yếu vào tính trạng kháng thuốc diệt cỏ. Thực tế cho thấy trong 94,1 triệu ha trồng đậu tương biến đổi gen, có 69,7 triệu ha trồng đậu tương có khả năng kháng thuốc diệt cỏ và chỉ có 24,4 triệu ha dành cho đậu tương có tính trạng kết hợp kháng sâu và kháng thuốc diệt cỏ [59]. Đậu tương có tính trạng kết hợp này đã được triển khai thành công ở Argentina, Paraguay và Uruguay. Ở Brazil, từ năm 2013 đến 2017, chính phủ đã phê duyệt 11 giống đậu tương biến đổi gen được phép sử dụng cho thực phẩm, thức ăn chăn nuôi và canh tác, trong đó có 2 giống mang tính trạng kháng sâu (bộ Cánh vảy) kết hợp với tính kháng thuốc diệt cỏ (glyphosate) [58].

Cho đến nay, hầu hết các giống đậu tương kháng sâu trên thế giới được tạo ra thông qua con đường chuyển gen. Về đậu tương chuyển gen kháng sâu, hiện nay chỉ có hai công ty sản xuất giống cây trồng biến đổi gen là Monsanto và Dow AgroSciences tạo ra giống đậu tương có tính kháng sâu đặc hiệu được phép canh tác trên thế giới (Bảng 1.2). Cho tới nay, các gen mã hóa độc tố Cry1Ac, Cry1F, Cry1A105 và Cry2Ab2 đã được nghiên cứu chuyển vào đậu tương. Monsanto đã phát triển giống đậu tương biến đổi gen MON87701 (biểu hiện Cry1Ac, có nguồn gốc từ chủng *B. thuringiensis kurstaki* HD73) được cấp phép lưu hành thương mại lần đầu ở Brazil trong mùa vụ 2013–2014. Giống đậu tương này cho thấy khả năng kháng đáng kể với sâu bướm *H. armigera* trong suốt cả mùa sinh trưởng khi được canh tác lần đầu tiên. Ấu trùng *H. armigera* có tỷ lệ sống từ 5,4% đến 24,4%, thấp hơn đáng kể so với nhóm đối chứng ăn lá cây không chuyển gen *Bt* (71–94,9%). Ngoài ra, giống này cũng làm giảm tỉ lệ sống sót, khối lượng của ấu trùng sâu keo và khả năng sinh sản của loài bướm đêm *Spodoptera litura* [60]. Bằng phương pháp lai tạo, Monsanto tiếp tục phát triển giống đậu tương chuyển gen *Bt* có tên thương mại Intacta™ Roundup Ready™ 2 Pro, là sự kết hợp hai giống đậu tương chuyển gen MON87701 và MON89788 (mang gen kháng thuốc diệt cỏ 5-enolpyruvylshikimate-3-phosphate synthase- EPSPS, có nguồn gốc từ vi khuẩn

Agrobacterium spp.). Giống này có khả năng kháng lại một số loài sâu hại thuộc bộ Cánh vảy như bướm nhung *Anticarsia gemmatalis* Hübner và *Chrysodeixis includens* Walker [61], [62]. Gần đây, giống đậu tương chuyển gen MON87751 mang hai gen *Bt* giống như ngô chuyển gen kháng sâu MON89034 là gen tái tổ hợp *cry1A.105* và gen *cry2Ab2* giúp cây kháng lại một số loài côn trùng Cánh vảy, được cấp phép lưu hành vào năm 2016.

Bảng 1.2. Các giống đậu tương chuyển gen kháng sâu đã thương mại hóa [63]

ST T	Sự kiện chuyển gen đậu tương kháng sâu	Gen chuyển	Nguồn gen	Chức năng
1	DAS81419	<i>cry1Ac</i>	<i>B. thuringiensis</i> subsp. <i>kurstaki</i> strain HD73	Kháng côn trùng bộ Cánh vảy
		<i>cry1F</i>	<i>B. thuringiensis</i> var. <i>aizawai</i>	
		<i>pat</i>	<i>Streptomyces viridochromogenes</i>	Kháng thuốc diệt cỏ
2	DAS81419 x DAS44406	<i>cry1Ac</i>	<i>B. thuringiensis</i> subsp. <i>kurstaki</i> strain HD73	Kháng côn trùng bộ Cánh vảy
		<i>cry1F</i>	<i>B. thuringiensis</i> var. <i>aizawai</i>	
		<i>pat</i>	<i>Streptomyces viridochromogenes</i>	Kháng thuốc diệt cỏ
		<i>add-12</i>	<i>Delftia acidovorans</i>	
		<i>2mepsps</i>	<i>Zea mays</i>	
3	MON87701	<i>cry1Ac</i>	<i>B. thuringiensis</i> subsp. <i>kurstaki</i> strain HD73	Kháng côn trùng bộ Cánh vảy
4	MON87701 x MON89788	<i>cry1Ac</i>	<i>B. thuringiensis</i> subsp. <i>kurstaki</i> strain HD73	Kháng côn trùng bộ Cánh vảy
		<i>cp4 epsps</i>	<i>Agrobacterium tumefaciens</i> strain CP4	Kháng thuốc diệt cỏ
5	MON87751	<i>cry1A.105</i>	<i>B. thuringiensis</i> subsp. <i>kumamotoensis</i>	Kháng côn trùng bộ Cánh vảy
		<i>cry2Ab2</i>		
6	MON87751 x MON87701	<i>cry1A.105</i>	<i>B. thuringiensis</i> subsp. <i>kumamotoensis</i>	Kháng côn trùng bộ Cánh vảy
		<i>cry2Ab2</i>		
	MON87701 x MON87708	<i>cry1Ac</i>	<i>B. thuringiensis</i> subsp. <i>kurstaki</i> strain HD73	
	MON87708 x MON89788	<i>dmo</i>	<i>Stenotrophomonas maltophilia</i> strain DI-6	Kháng thuốc diệt cỏ
		<i>cp4 epsps</i>	<i>Agrobacterium tumefaciens</i> strain CP4	

Tiếp nối sự phát triển của ba giống đậu tương trên, Công ty Monsanto đã tạo chọn tạo ra giống MON 87751 × MON 87701 × MON 87708 × MON 89788 mang tất cả các gen *Bt* của các giống đậu tương thành phần, và được 2 quốc gia là Đài Loan và Hàn Quốc cho phép canh tác [64]. Tương tự, giống đậu tương biến đổi gen DAS81419 và DAS81419 x DAS44406 đã được công ty Dow AgroSciences pháp triển sử dụng phương pháp chuyển gen thông qua vi khuẩn *Agrobacterium* để biểu hiện các protein Cry1Ac, Cry1F và phosphinothricin acetyltransferase (PAT) có nguồn gốc tương ứng từ *B. thuringiensis* subsp. *kurstaki*, *B. thuringiensis* subsp. *aizawai* và *Streptomyces viridochromogenes*, giúp bảo vệ cây khỏi thiệt hại do một số loài sâu bọ Cánh vẩy gây ra [65]. Bên cạnh đó, một số quốc gia phát triển cũng đẩy mạnh các nghiên cứu sử dụng gen mã hóa độc tố *Bt* mới để chọn tạo giống đậu tương có khả năng kháng nhiều loài sâu khác nhau. Qin và đồng tác giả (2019) đã tiến hành chuyển gen *cry8* có nguồn gốc từ chủng *Bt* HBF-18 vào giống đậu tương Jinong 28 giúp cây kháng được sâu non bọ cánh cứng *Holotrichia parallela* [66].

1.3.2. Nghiên cứu tạo giống đậu tương kháng sâu ở Việt nam

Nghiên cứu tạo giống cây trồng biến đổi gen là lĩnh vực được quan tâm ở Việt Nam. Về mảng tạo giống cây trồng kháng sâu bệnh bằng công nghệ gen, nhiều dự án nghiên cứu đã được tiến hành và thu được những thành tựu nhất định trên một số loại cây hoa màu và lương thực như khoai lang, ngô, bông và đậu tương. Khoai lang chuyển gen *cry8Db* trên giống nền là khoai lang KB1 của Việt Nam có sức chống chịu tốt với bọ hà đã được tạo thành công tại Viện Công nghệ Sinh học [67]. Để đối phó với sâu hại bông và cỏ dại, Tập đoàn Dệt May Việt Nam, Bộ Công Thương đã tiến hành nghiên cứu và đưa ra 17 dòng bông chuyển gen *vip3A* và *bar* có khả năng kháng sâu và chịu thuốc diệt cỏ có triển vọng [68], [69], [70]. Ngô chuyển gen kháng sâu cũng có bước đột phá với thành công của Công ty Syngenta Việt Nam trong việc tạo ra 2 giống ngô biến đổi gen NK66 BT và NK66 Bt/GT mang gen *cry1Ab* kháng sâu đục thân trên giống nền NK66 (<https://www.syngenta.com.vn/thanh-tuu-cong-nghe>). Ở Việt Nam, từ sau năm 2000, nghiên cứu chuyển gen vào đậu tương đã được thực hiện thành công ở nhiều trung tâm, viện nghiên cứu, tuy nhiên phần lớn đều tập trung vào các gen liên quan đến khả năng chịu hạn và kháng thuốc diệt cỏ [71], [72]. Về đậu tương chuyển gen

Bt kháng sâu, dự án “Tạo dòng đậu tương biến đổi gen kháng sâu và chịu hạn” và “Nghiên cứu tạo giống đậu tương biến đổi gen kháng ruồi đục thân và sâu đục quả” được thực hiện bởi Viện lúa Đồng Bằng Sông Cửu Long vào các giai đoạn 2006-2010 và 2011-2016, tập trung vào việc chuyển các gen *cry1Ac*, *cry2Aa*, *cry4A* và *vip3A* vào giống đậu tương nhập nội Williams 82, Maverick và giống đậu tương Việt Nam MTĐ176 [73], [74]. Gần đây nhất, từ 2017-2020, PGS. TS Nguyễn Văn Đồng, Viện Di truyền Nông nghiệp đã chủ trì nhiệm vụ “Phân lập thiết kế gen kháng sâu tạo giống đậu tương biến đổi gen”. Kết quả đề tài đã xây dựng được bộ chủng *Bt* bản địa cùng các gen mã hóa protein kháng sâu tiềm năng có hoạt tính diệt sâu hại quả đậu tương [24]. Các gen tiềm năng này đã được phân lập, cải biến mã và thiết kế cấu trúc biểu hiện trong thực vật, đặt tiền đề cho các nghiên cứu tạo cây đậu tương biến đổi gen có năng suất cao, chất lượng tốt, kháng sâu có thể áp dụng được cho sản xuất, góp phần đáp ứng nhu cầu thực tiễn, đồng thời giảm ảnh hưởng của việc sử dụng thuốc trừ sâu đến môi trường và sức khỏe con người. Việt Nam được công nhận là một trong những quốc gia có tính đa dạng sinh học cao nhất thế giới, vì thế việc khai thác nghiên cứu sàng lọc và lựa chọn các chủng *Bt* bản địa mang các gen đặc hiệu diệt côn trùng đích có ý nghĩa khoa học, tính khả thi và triển vọng ứng dụng thực tiễn lớn.

Tóm lại, đậu tương là một trong những cây trồng quan trọng nhất trên toàn cầu nhờ giá trị to lớn đối với ngành chế biến thức ăn chăn nuôi và thực phẩm. Tuy vậy, tại Việt Nam, diện tích canh tác đậu tương ngày càng thu hẹp, cộng thêm năng suất thấp dẫn đến sản lượng đậu tương nội địa mới chỉ cung cấp được một phần rất nhỏ cho nhu cầu tiêu thụ trong nước, còn lại phải phụ thuộc hoàn toàn vào nguồn đậu tương nhập khẩu từ nước ngoài. Thiệt hại do côn trùng gây ra, trong đó có sâu đục quả *E. zinckenella* Treitschke là nguyên nhân hàng đầu dẫn đến việc giảm năng suất gieo trồng đậu tương. Việc phòng trừ *E. zinckenella* Treitschke bằng thuốc trừ sâu hóa học và thuốc bảo vệ thực vật thường đem lại hiệu quả không cao, trái lại còn gây ô nhiễm môi trường và có hại cho con người do sự tồn dư hóa chất độc hại trong quả. Do vậy, việc nghiên cứu, phát triển thuốc trừ sâu sinh học có nguồn gốc *Bt* hoặc tạo cây đậu tương chuyển gen *Bt* có khả năng kháng sâu đục quả là một yêu cầu bức thiết do thực tiễn đặt ra. Tuy nhiên, nguồn gen *Bt* có tiềm năng làm cơ sở

để thực hiện các nghiên cứu này mới được xác định và công bố rất ít. Bên cạnh đó, việc tìm ra các gen độc tố mới có hoạt lực cao trong việc diệt sâu đục quả đậu tương *E. zinckenella* Treitschke mang tầm chiến lược lâu dài, góp phần giải quyết tình trạng kháng thuốc xuất hiện trong các quần thể sâu trên đồng ruộng. Việc tìm ra các gen *Bt* mới với hoạt tính mong muốn được chúng tôi kỳ vọng đạt được bằng cách khai thác Bộ sưu tập *Bt* bản địa, bao gồm đa dạng các chủng *Bt* thu thập tại nhiều tỉnh, thành của Việt Nam. Sau khi sàng lọc được các chủng *Bt* tiềm năng có hoạt lực cao trong diệt sâu đục quả đậu, thay vì sử dụng phương pháp PCR với bộ mồi phát hiện gen độc tố được sử dụng truyền thống trước đây giúp dự đoán thành phần gen mã hóa protein diệt côn trùng trong mỗi chủng *Bt*, chúng tôi tiếp cận bằng phương pháp giải trình tự gen thế hệ mới kết hợp với việc phân tích, tìm kiếm dữ liệu sử dụng công cụ tin sinh học thông lượng cao nhằm xác định chính xác chủng *Bt* mang gen mã hóa độc tố *Bt* mới cũng như trình tự của các gen ứng viên. Sau khi được phân lập và giải trình tự Sanger để kiểm chứng một cách chắc chắn, protein tái tổ hợp mã hóa bởi các gen độc tố *Bt* này được biểu hiện và tinh sạch để thử khả năng diệt ấu trùng sâu đục quả *E. zinckenella* Treitschke. Độc tố *Bt* mới có hoạt lực mạnh nhất sẽ tiếp tục được nghiên cứu sâu hơn về khả năng biểu hiện của gen mã hóa trong hệ biểu hiện thực vật thông qua các chiến lược như cải biến mã di truyền, lựa chọn promoter cũng như tối ưu các điều kiện biểu hiện. Kết quả của nghiên cứu cung cấp dữ liệu về các chủng *Bt* bản địa của Việt Nam cùng các gen mã hóa độc tố *Bt* mới có hoạt lực cao trong diệt sâu *E. zinckenella* Treitschke. Đây là cơ sở quan trọng, hữu ích cho những nghiên cứu tiếp theo về sản xuất thuốc trừ sâu sinh học *Bt* cũng như tạo cây đậu tương chuyển gen *Bt* kháng loại sâu đục quả khó phòng trừ này.

Chương 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. ĐỐI TƯỢNG NGHIÊN CỨU

2.1.1. Vật liệu

2.1.1.1. *Vật liệu thực vật và côn trùng*

Cây thuốc lá *N. benthamiana* 4 tuần tuổi được trồng thủy canh trong buồng sinh trưởng với điều kiện nhiệt độ là 21°C và quang chu kỳ sáng/tối là 16/8 h do Phòng Công nghệ Tế bào Thực vật, Viện Công nghệ sinh học- Viện Hàn Lâm Khoa học và Công nghệ Việt Nam cung cấp.

Ấu trùng sâu đục quả đậu tương *E. zinckenella Treitschke* tuổi 2-3 và thức ăn nhân tạo được cung cấp bởi Viện Bảo vệ Thực vật.

2.1.1.2. *Vi sinh vật*

- 221 chủng vi khuẩn *Bt* bản địa lấy từ Bộ sưu tập *Bt* Việt Nam do Trung tâm Giống và Bảo tồn nguồn gen Vi sinh vật, Viện Công nghệ sinh học, Viện Khoa học và Công nghệ Việt Nam cung cấp, có nguồn gốc từ 4 tỉnh miền núi phía Bắc Việt Nam bao gồm Lào Cai, Tuyên Quang, Hà Giang, Điện Biên (Bảng phụ lục 1).
- Chủng *E. coli* DH5 α (Invitrogen, Mỹ) được sử dụng làm tế bào chủ để nhân dòng gen.
- Chủng *E. coli* BL21 (DE3) (Novagen, Mỹ) được sử dụng làm vật chủ để biểu hiện protein tái tổ hợp.
- Chủng vi khuẩn *Agrobacterium tumefaciens* AGL1 (Nova Lifetech, Hongkong) được sử dụng cho thí nghiệm chuyển gen thông qua *Agrobacterium* và biểu hiện tạm thời.

2.1.1.3. *Vector*

- Vector pJET1.2/blunt (Thermo Fisher Scientific, Mỹ) được sử dụng để nhân dòng gen (Hình phụ lục 1a).
- Vector pET32a(+) (Addgene, Mỹ) được thiết kế để biểu hiện protein tái tổ hợp (Hình phụ lục 1b).

- Vector pFGC5941 (Addgene, Mỹ) được sử dụng để thiết kế vector chuyển gen vào thực vật (Hình phụ lục 1c).
- Vector pBI121/35S_HcPro, vector biểu hiện chứa gen mã hóa cho HcPro- một proteinase của virus PVY (cung cấp bởi Phòng Công nghệ DNA Ứng dụng, Viện Công nghệ Sinh học), được dùng để đồng biểu hiện trong thí nghiệm biểu hiện tạm thời với vector tái tổ hợp chứa gen đích cần biểu hiện.

2.1.1.4. Kháng thể

- Kháng thể đơn dòng chuột kháng histidine (monoclonal anti-6X His tag antibody) của Sigma, Mỹ.
- Kháng thể cộng hợp peroxidase kháng IgG của chuột được mua từ hãng Sigma, Mỹ.

2.1.1.5. Môi nhân gen

Bảng 2.1. Các cặp môi được sử dụng trong nghiên cứu

STT	Tên môi	Trình tự môi	Mục đích
1	pJET1.2_F	CGACTCACTATAGGGAGAGCG GC	Chọn lọc vector tách dòng pJET1.2/blunt có mang gen đích, giải trình tự gen đích được nhân dòng trong pJET1.2/blunt
2	pJET1.2_R	AAGAACATCGATTTTCCATGGC AG	
3	cry1Aa_F	GGATCC CATGGATAACAATCCG AACATCA	Nhân gen <i>cry1Aa</i> từ chủng Đ6.2 (3,5 kb)
4	cry1Aa_R	CTCGAGG CACTCATCAAAGGT ACCCAAT	
5	cry1Ac_F	GGATCC CATGGATAACAATCCG AACATCA	Nhân gen <i>cry1Ac</i> từ chủng Đ6.2 (3,5 kb)
6	cry1Ac_R	CTCGAGT TCTGGTGCCTATTA ATGTCTT	
7	cry1Be_F	GGATCCT GTAAGTACAACGTA AAAGGAGGAA	Nhân gen <i>cry1Be</i> từ chủng Đ6.1 (3,6 kb)
8	cry1Be_R	CTCGAGC AAGCAGAATTCATC CGATAAAC	
9	cry1Na_F	GGATCC CAGCCTCATTTTAAAA TTATTCCTTTTAAATG	Nhân gen <i>cry1Na</i> từ chủng Đ6.1 (1,8 kb)
10	cry1Na_R	CTCGAGT TCAACAGGAATAAA TTCAATTTTATCC	
11	cry2Ab_F	GGATCC CATGAATAGTGTATTG AATAGCGGAA	Nhân gen <i>cry2Ab</i> từ chủng SP14.2 (1,9 kb)
12	cry2Ab_R	CTCGAG AATAAAGTGGTGAAAT ATTAGTTGGTA	

13	cry2Ah_F	GGATCC ATGAATAATGTATTG AATAGCGGAAG	Nhân gen <i>cry2Ah</i> từ chủng BD8.2 (1,9 kb)
14	cry2Ah_R	CTCGAG AATAAAGTGGTGAAAT ATTAGTTGGAA	
15	cry1Be_F1	GGATCCT GTAAGTACAACGTA AAAGGAGGAA	Nhân phân đoạn gen <i>cry1Be</i> (2 kb)
16	cry1Be_R1	CAAGCAGAATTCATCCGATAA AC	
17	rbcS1A-pro_F	GAATTC CCCGTGGTCGAGATTG TG	Nhân đoạn promoter rbcS1A từ Arabidopsis (1,1 kb)
18	rbcS1A-pro_R	CTCGAG TGTTCTTCTTTACTCT TTGTGTGA	

Trình tự in đậm: vị trí nhận biết của enzyme cắt hạn chế *Bam*HI (GGATCC), *Xho*I (CTCGAG) và *Eco*RI (GAATTC)

2.1.2. Hóa chất và thiết bị

2.1.2.1. Các loại hóa chất

Thang DNA 1 kb (Thermo Scientific, USA); Master Mix 2X (Promega, USA); Kit tạo thư viện DNA (SMRTbell Template Prep Kits, PacBio, USA); Kit tinh sạch DNA (GeneJET Gel Extraction Kit, Thermo Scientific, USA); Kit tách chiết plasmid (Bioneer, Hàn Quốc); các loại enzyme cắt giới hạn *Bam*HI, *Xho*I, *Eco*RI, *Nco*I và *Xma*I (Thermo Scientific, USA); bovine serum albumin (Sigma-Aldrich); dung dịch Coomassie Plus Protein Assay Reagent (Thermo Scientific, USA), Coomassie Brilliant Blue R-250 (Merck, Đức); Ni-NTA resin (Cube Biotech, Đức).

2.1.2.2. Thiết bị, máy móc

Hệ thống giải trình tự thế hệ mới PacBio SEQUEL system (PacBio, USA); hệ thống điện di tự động được thiết lập để kiểm soát chất lượng DNA (Bioanalyzer system (Agilent 2100, USA); máy PCR system 97000 (Applied Biosystem, USA); máy điện di Powerpac300 (Bio-Rad, USA); bộ điện di protein (Bio-Rad); máy soi DNA (mini-transilluminator, Bio-Rad, USA), máy vortex (minishaker, IKA, Đức); máy ly tâm lạnh (Hitech), máy Nanodrop lite (Thermo Scientific), thiết bị QUBIT 3.0 (Thermo Scientific), máy xung điện Gene Pulser (Bio-Rad), thiết bị hấp khử trùng (Hirayama); máy đo pH (Horiba), cân phân tích

(Sartorius), tủ cấy vi sinh (Esco), box cấy vô trùng, bể ổn nhiệt... cùng các trang thiết bị khác.

2.2. PHƯƠNG PHÁP NGHIÊN CỨU

2.2.1. Sàng lọc các chủng *Bacillus thuringiensis* có hoạt tính diệt sâu đục quả đậu tương

2.2.1.1. Phương pháp hoạt hoá các chủng *Bt*

Các chủng *Bt* được giữ giống trong các ống chuyên dụng chứa glycerol 20% và bảo quản ở -80°C được lấy ra đặt trên đá để rã đông. Dùng que cấy vô trùng nhúng vào dịch giữ chủng khuẩn và cấy ria lên đĩa petri chứa môi trường LB đặc nhằm mục đích phân tách các khuẩn lạc riêng rẽ. Đĩa cấy sau đó được ủ ở tủ nuôi 28°C trong 24 giờ.

2.2.1.2. Phương pháp tăng sinh khối

Nhân giống cấp 1: Dùng đầu que cấy vô trùng lấy khuẩn lạc của các chủng vi khuẩn *Bt* từ đĩa thạch cấy vào 5 ml môi trường LB, nuôi lắc 200 vòng/phút ở 28°C trong 24 giờ, đến khi $\text{OD}_{600\text{ nm}} = 0,6 - 0,8$.

Cấy chuyển: Các chủng *Bt* sau khi được hoạt hóa, lấy 5 ml dịch chuyển sang nuôi ở 50 ml môi trường lỏng LB trong bình tam giác 250 ml. Nuôi lắc ở 28°C , 200 vòng/phút trong 72 giờ sao cho $\text{OD}_{600} \geq 2$. Tiến hành thu canh khuẩn và chứa trong các ống ficol, bảo quản ở tủ lạnh 4°C .

2.2.1.3. Xác định hình thái bào tử và protein tinh thể bằng kính hiển vi điện tử quét (SEM)

Dịch khuẩn *Bt* sau 72 giờ nuôi cấy để tạo bào tử và hình thành protein tinh thể được ly tâm thu sinh khối ở 13000 vòng/phút trong 5 phút, 4°C . Cặn khuẩn sau đó được rửa bằng đệm sodium cacodylate 0,1 M (pH = 7,2) trước khi được cố định lần lượt trong các dung dịch glutaraldehyde 4% (2 giờ) và osmium tetroxide 1% (1 giờ) ở 4°C . Sau khi rửa 3 lần bằng đệm cacodylate 0,1 M, mẫu được khử nước bằng cách cho qua rượu ethanol ở các nồng độ khác nhau (30, 50, 70, 90, 100%). Giai đoạn cuối cùng là gắn trên giá đỡ kính hiển vi điện tử quét và phủ hợp kim vàng-

palladium bằng thiết bị phún xạ anion. Các mẫu được kiểm tra và chụp ảnh bằng kính hiển vi điện tử quét LEO440 ở chùm tia 20 kV với độ phóng đại 5000x.

2.2.1.4. Phương pháp đánh giá khả năng diệt sâu của các chủng *Bt*

Chuẩn bị môi trường thức ăn nhân tạo cho sâu non *E. zinckenella* Treitschke: Thành phần của thức ăn nhân tạo gồm: bột đậu nành 125 g/L, bột ngô 20 g/L, ascorbic acid 4 g/L, cao chiết nấm men 40 g/L, mầm lúa mạch 4 g/L, acetic acid 2 g/L, choline chloride 2 g/L, casein 25 g/L, sorbic acid 1,5 g/L, vitamin tổng hợp 5 ml/L, dohyfral (0.5 g/L). Cân 200 g thức ăn nhân tạo, thêm 625 ml nước cất và cho vào máy xay. Đun nóng 20 g thạch agar trong 375 ml nước cất rồi đổ vào hỗn hợp thức ăn đã xay nhuyễn, trộn đều lại một lần nữa bằng máy xay. Đổ thức ăn vào các hộp nhựa kích thước 9 x 6 x 3,5 cm, mỗi hộp 20 g, chờ nguội rồi dùng dao rạch thành các khối vuông nhỏ (2 x 2 mm) hoặc tạo các lỗ nhỏ sâu từ bề mặt đến đáy khối thức ăn.

Đánh giá độc tính của các chủng *Bt* trên sâu đục quả *E. zinckenella* Treitschke: Các chủng vi khuẩn *Bt* sau khi hoạt hóa, nhân nuôi (sau 72 giờ) thu sinh khối được đánh giá khả năng diệt sâu non *Etiella zinckenella* Treitschke tuổi 2-3. Bỏ sung 2 ml dịch canh khuẩn vào hộp môi trường (rải đều), để qua đêm cho dịch canh khuẩn thấm đều vào môi trường. Tiến hành bắt sâu non vào các hộp nhựa chứa môi trường đã chuẩn bị (mỗi hộp 10 con, mỗi công thức lặp lại 3 lần). Song song là đối chứng âm bỏ sung 2 ml nước cất vô trùng + 10 sâu tuổi 2-3/hộp môi trường, 3 lần nhắc lại. Nuôi sâu ở điều kiện nhiệt độ phòng, sau 24 giờ nuôi theo dõi số lượng sâu chết trong 5 ngày.

Hiệu lực diệt sâu được tính theo công thức Abbott (1925):

$$A = (C - T) / C \times 100$$

A: % sâu chết

C: Số sâu sống ở mẫu đối chứng

T: Số sâu sống ở mẫu thí nghiệm

2.2.2. Giải trình tự hệ gen vi khuẩn *Bacillus thuringiensis* bằng công nghệ SMRT Sequencing của PacBio và khai thác cơ sở dữ liệu hệ gen *Bt* để tìm các gen mã hóa protein độc tố diệt sâu bằng công cụ tin sinh học

2.2.2.1. Phương pháp tách chiết DNA tổng số từ vi khuẩn *Bt*

DNA tổng số của *Bt* được tách chiết theo phương pháp của Mayjonade và cộng sự [75]. Thu tế bào vi khuẩn từ 2 ml dịch nuôi cấy qua đêm bằng cách ly tâm ở 5000 g trong 10 phút ở nhiệt độ phòng. Bổ sung 600 μ l đệm chiết (PVP 1%, Na₂S₂O₅ 1%, NaCl 0,5 M, Tris 100 mM, EDTA 50 mM, SDS 1,25%) đã được làm ấm tới 65°C và 4 μ l RNase A (100 mg/ml QIAGEN), lập tức trộn đều bằng cách hút lên xuống 10 lần bằng pipet với đầu tip to, tránh sử dụng vortex vì sẽ làm đứt gãy DNA. Ủ mẫu ở 50°C trong 30 phút, cứ 10 phút đảo ống mẫu 20 lần giúp đồng nhất dịch phá tế bào. Bổ sung 200 μ l potassium acetate 5M, đảo ống 20 lần, ủ mẫu ở 4°C để tủa hoàn toàn protein. Ly tâm 5000g trong 10 phút ở 4°C, chuyển dịch nổi sang ống 2 ml mới. Bổ sung 1 lượng thể thích tương đương đệm gắn (PEG8000 20%, NaCl 3M) và 1/18 thể tích dung dịch hạt từ Serapure (hạt từ Sera-Mag SpeedBead đã được carboxyl hóa- Fisher Scientific, Mỹ) 2%, PEG8000 18%, NaCl 1M, Tris-HCl 10 mM pH 8.0, EDTA 1 mM pH 8.0, Tween 20 0,05%), trộn đều bằng cách đảo ống 20 lần, ủ mẫu ở nhiệt độ phòng trong 10 phút. Ly tâm nhanh trong 1 giây, đặt ống trên rack từ 3 phút đến khi dịch trong ống trong trở nên suốt. Loại bỏ dịch mà không phá vỡ tủa hạt từ. Bổ sung 1 ml đệm rửa (Ethanol 75%), lấy ống khỏi rack từ và đảo ống 20 lần. Ly tâm nhanh trong 1 giây, đặt ống trên rack từ và đợi 30 giây đến khi dịch trong ống trong trở nên suốt, hút bỏ đệm rửa. Lặp lại bước rửa này thêm 1 lần nữa. Sau đó làm khô tủa hạt từ trong 1 phút. Bổ sung 150 μ l đệm EB đã được làm ấm 50°C, bung thành ống để hạt từ phân tán đều trong đệm. Ly tâm nhanh 1 giây và đặt ống trên rack từ, chờ đến khi dịch trong ống trở nên trong suốt, hút dịch gDNA chuyển sang ống mới. Chất lượng DNA được kiểm tra bằng phương pháp điện di trên gel agarose 1%; đo nồng độ tương đối cùng các chỉ số độ sạch A260/A280, A260/A230 bằng máy Nanophotometer (IMPLEN) và đo chính xác nồng độ DNA bằng bộ kit Qubit™ dsDNA HS Assay Kits cùng với thiết bị QUBIT 3.0.

2.2.2.2. Phương pháp chuẩn bị thư viện

DNA tổng số đạt chuẩn (DNA toàn vẹn không bị đứt gãy trong quá trình tách chiết, DNA có giá trị A280/A260 đạt 1,8-2,0, A260/A230 đạt 1,8-2,0, tổng lượng DNA thu được khoảng 2-3 µg) được sử dụng làm nguyên liệu tạo thư viện DNA. Chúng tôi sử dụng 1 µg DNA genome pha loãng trong 150 µl dung dịch đệm EB để cắt bằng thiết bị g-TUBE™ (Covaris) theo quy trình cắt DNA genome thành đoạn nhỏ 10 kb. DNA trước sau khi cắt được kiểm tra bằng thiết bị Bioanalyzer Agilent 2100 với bộ kit Agilent DNA 12000 Kit (5067-1508 - Agilent). DNA sau khi cắt được chuẩn bị thư viện bằng bộ kit SMRTbell Damage Repair Kit – SPv3 (100-992-200 – PacBio). DNA thư viện được kiểm tra lại một lần nữa bằng Bioanalyzer Agilent 2100 trước khi chuyển sang bước tiếp theo.

2.2.2.3. Phương pháp giải trình tự genome

Thư viện sau khi chuẩn bị được gắn với polymerase và tinh sạch bằng bộ kit Sequel Binding and Internal Ctrl Kit 2.0 (101-400-900 – PacBio) và SMRTbell Clean Up Column v2 Kit-Dif (101-184-100 – PacBio) theo quy trình được tạo bởi phần mềm Sample Setup có trong SMRTLink portal phiên bản 5.1. Thư viện DNA được đưa vào Sample Plate (000-448-888 – PacBio) với lượng DNA được tính toán để nồng độ thư viện load lên chip là 20 pM. Thư viện được giải trình tự bằng thiết bị giải trình tự đoạn dài PacBio SEQUEL sử dụng chip giải trình tự Sequel SMRT Cell 1M v2 Tray (101-008-000 – PacBio) cùng bộ kit giải trình tự Sequel Sequencing Kit 2.1 (101-310-500 - PacBio). Thời gian giải trình tự là 600 phút.

2.2.2.4. Phương pháp lắp ráp *denovo* hệ gen

Tín hiệu giải trình tự được xử lý, đánh giá bằng máy chủ PrimaryAnalysis Computer và chuyển thành dữ liệu thô (raw read). Toàn bộ dữ liệu này được tự động chuyển sang hệ máy chủ SecondaryAnalysis Server thông qua mạng nội bộ. Dữ liệu giải trình tự được lắp ráp *denovo* bằng phần mềm HGAP4 tích hợp trong cổng thông tin SMRTLink portal phiên bản 5.1 trên SecondaryAnalysis Server (<https://www.pacb.com/support/software-downloads/> - mới nhất tính đến thời điểm 07/2018). Khi lắp ráp, các thông số đầu vào của HGAP4 được thiết lập như sau: độ

dài genome là 8.000.000 bp (8 Mb), độ bao phủ tối thiểu là 30, độ dài cutoff của seed read là tự động, giá trị filter dataset là 0,7

Chất lượng lắp ráp hệ gen được đánh giá bằng 2 phần mềm QUAST phiên bản 4.6.3 (<http://bioinf.spbau.ru/quast>) [76] và BUSCO phiên bản 3 (<https://busco.ezlab.org/>) [77], [78]. QUAST được sử dụng để đánh giá các chỉ số về contig: số lượng contigs, độ dài lớn nhất của contig, tổng độ dài của các contig, N50, L50, phần trăm bao phủ của các contig trong genome lắp ráp. BUSCO định lượng độ hoàn chỉnh của bộ gen lắp ráp dựa trên tập hợp các orthologs từ OrthoDB v9.

2.2.2.5. Phương pháp chú giải nhanh hệ gen lắp ráp *denovo*

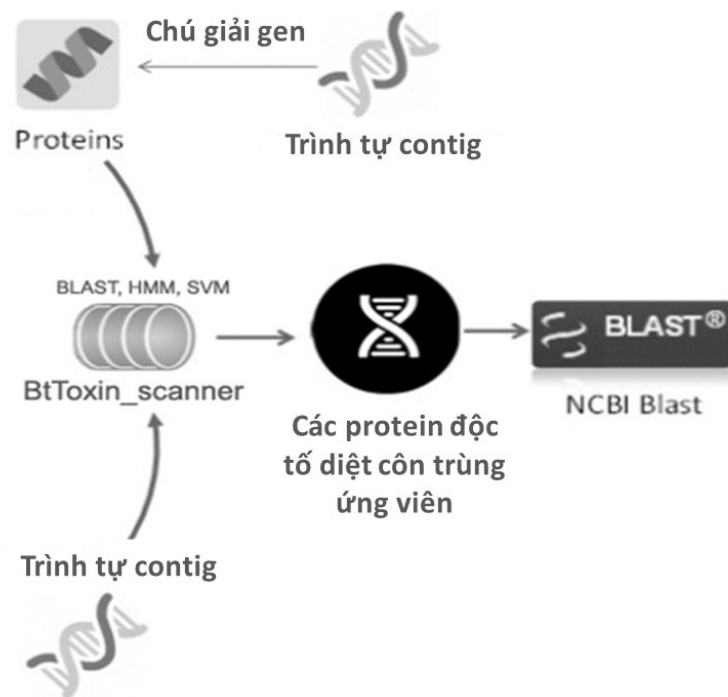
Hệ gen lắp ráp *denovo* của *Bt* được phân tích và chú giải nhanh bằng gói dịch vụ chú giải hệ gen trên hệ server PATRIC (The Pathosystems Resource Integration Center - <https://www.patricbrc.org/>) [79]. Gói dịch vụ chú giải hệ gen của PATRIC bao gồm quy trình chú giải hệ gen khép kín sử dụng RASTtk (RAST tool kit -Rapid Annotation using Subsystem Technology là bộ công cụ chú giải hệ gen được phát triển từ năm 2008 dành cho chú giải các bộ gen vi khuẩn [80] và nhiều cơ sở dữ liệu liên kết từ NCBI, CDC, KEGGs...

2.2.2.6. Phương pháp dự đoán, xác định các gen độc tố *Bt* từ cơ sở dữ liệu trình tự hệ gen các chủng *Bt* bằng công cụ *BtToxin_scanner*

Với cơ sở dữ liệu hệ gen đã được lắp ráp của các chủng *Bt*, việc tìm kiếm và tiên đoán gen mã hóa protein độc tố *Bt* được tiến hành nhờ công cụ *BtToxin_scanner*. Đây là một công cụ khai thác trên web (http://bcam.hzau.edu.cn/BtToxin_scanner/index.php) tích hợp 3 module với tính năng dự đoán là BLAST (Basic Local Alignment Search Tool), HMM (Hidden Markov Model) và SVM (Supporting Vector Machine). Cơ sở dữ liệu BLAST được xây dựng sử dụng bộ dữ liệu các gen độc tố *Bt* được duy trì và cập nhật bởi Crickmore và cs. (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/) với giá trị ngưỡng kỳ vọng là $1e-25$. Mô hình HMM (Cry.hmm, Cyt.hmm và Vip.hmm) được tạo dựng bằng ứng dụng HMMER3.0 với giá trị ngưỡng kỳ vọng là $1e-10$. Bộ phân loại SVM cho protein Cry được phát triển với thành phần axit amin và dipeptide sử dụng

LIBSVM 2.91. Bất kỳ trình tự nào vượt qua được 1 trong 2 module dự đoán này đều được coi là protein Cry ứng viên. Trong khi đó, các trình tự protein Vip và protein Cyt chỉ được dự đoán bởi module HMM với giá trị ngưỡng kỳ vọng là $1e-10$.

Trong quá trình xử lý trình tự đầu vào, dựa trên cơ sở các dạng trình tự khác nhau mà các module tương thích được lệnh chuyển đổi sang trình tự protein. Với trình tự nucleotit, module dịch mã 6 khung đọc được sử dụng để thu được tất cả trình tự protein giả định, hạn chế việc bỏ sót các trình tự độc tố tiềm năng. Để kiểm định các trình tự ứng viên nhằm loại bỏ trình tự dương tính giả, thuật toán NCBI Blastp/Blastx được thực thi trên cơ sở dữ liệu non-redundant protein với các thông số mặc định. Trình tự ứng viên sẽ được chấp nhận dựa trên kết quả blast và sự có mặt của các vùng chức năng độc tố được xác thực bởi CDD (Conserve Domain Database) của NCBI (Hình 2.1).



Hình 2.1. Quy trình xác định các gen độc tố diệt côn trùng từ dữ liệu giải trình tự hệ gen các chủng *Bt*. [41]

2.2.3. Phương pháp phân lập, tách dòng gen

2.2.3.1. Khuếch đại gen đích bằng phương pháp PCR

Từ DNA hệ gen các chủng *Bt* đã lựa chọn, gen *cry1Aa*, *cry1Ac*, *cry1Na*, *cry1Be*, *cry2Ab* và *cry2Ah* được nhân lên bằng phương pháp PCR sử dụng các cặp mồi đặc hiệu được trình bày trong bảng 2.1. Tại đầu 5' của các mồi xuôi và ngược ứng với mỗi gen được thiết kế thêm điểm nhận biết của enzyme hạn chế *Bam*HI và *Xho*I để thuận tiện cho quá trình thiết kế vector sau này. Để thực hiện phản ứng nhân gen, chúng tôi sử dụng Phusion DNA polymerase ngoài hoạt tính 5'-3' polymerase xúc tác cho quá trình polymer hóa các deoxiribonucleotide giúp khuếch đại một trình tự DNA theo hàm mũ lên đến hàng triệu bản copy, enzyme này còn có khả năng đọc sửa (proofreading) nhờ hoạt tính 3'-5' exonuclease giúp hạn chế các sai hỏng trên sợi DNA được tổng hợp mới. Phản ứng PCR được tiến hành trong tổng thể tích 25 μ l có chứa dung dịch đệm (10X) cho Phusion DNA polymerase, hỗn hợp 4 loại deoxynucleoside triphosphate (dNTPs, 1 mM mỗi loại), 1 pmol mồi mỗi loại xuôi và ngược cho từng gen, 1 U Phusion DNA polymerase và khoảng 50 ng DNA tổng số dùng làm khuôn để khuếch đại gen quan tâm. Chu trình nhiệt của phản ứng PCR gồm các bước: (1) biến tính DNA khuôn từ dạng sợi kép thành dạng sợi đơn bằng cách nâng nhiệt độ lên 98°C trong vòng 30 s; (2) Tổng hợp sợi polynucleotide được tiến hành trong 30 chu kỳ, mỗi chu kỳ gồm 3 bước: biến tính DNA sợi kép thành sợi đơn ở 98°C trong 10 s; gắn mồi ở 60°C trong 10 s; kéo dài sợi polynucleotide ở 72°C trong 1 phút; (3) Duy trì nhiệt độ 72°C trong khoảng 3 phút để các sợi được tổng hợp một cách hoàn thiện. Sản phẩm PCR được kiểm tra bằng điện di trên gel agarose 0,8% và giữ ở -20°C.

2.2.3.2. Phương pháp tách dòng gen

Sản phẩm PCR được làm sạch theo bộ Kit QIAquick Gel Extraction và gắn vào vector tách dòng pJET1.2/blunt. Phản ứng nối ghép gen được xúc tác bởi T4-DNA ligase trong hỗn hợp 10 μ l gồm các thành phần như sau: dung dịch đệm (10X) cho T4-DNA ligase, 10 ng vector pJET1.2/blunt, 10 – 20 ng DNA, 1 U T4 DNA-ligase. Hỗn hợp được ủ ở 22°C trong 60 phút và sau đó được biến nạp vào tế bào khả biến *E.coli* DH5 α .

2.2.3.3. *Biến nạp DNA plasmid vào tế bào E. coli bằng phương pháp sốc nhiệt*

Chuẩn bị tế bào khả biến:

Tế bào *E.coli* DH5 α được cấy ria trên môi trường LB đặc và nuôi qua đêm ở 37°C. Chọn một khuẩn lạc nuôi cấy trong 5 ml LB lỏng, lắc 200 v/p, ở 37°C, qua đêm. Hút 1% dịch khuẩn ở trên cho vào 50 ml LB lỏng, lắc 200 v/p, ở 37°C, trong 2-3 giờ, đo OD₆₀₀ đạt 0,4 - 0,6 thì chuyển dịch nuôi cấy sang ống ly tâm, để 15 phút ở trong đá (hoặc ở 4°C). Ly tâm 4000 v/p, ở 4°C, trong 5 phút. Thu cạn và hoà tan trong 30 ml CaCl₂ 0,1M, để trong đá 15 phút, sau đó ly tâm 4000 v/p, ở 4°C, trong 5 phút (lặp lại ba lần, để trong đá lần thứ hai 30 phút, lần thứ ba 60 phút). Đổ dịch CaCl₂, hoà tan trong 1ml CaCl₂/100 ml LB ban đầu. Bổ sung 15-20% glycerol. Chia 50 μ l dịch tế bào khả biến vào mỗi ống eppendorf 1,5 ml. Bảo quản tế bào khả biến trong tủ lạnh -80°C.

Biến nạp:

Tế bào khả biến được lấy ra từ tủ -80°C và được làm tan trên đá. Bổ sung 5 μ l vectơ đã gắn gen vào ống đựng tế bào khả biến và trộn nhẹ. Hỗn hợp được đặt trong đá 10-15 phút, rồi được chuyển sang bể ổn nhiệt có nhiệt độ 42°C trong 1 phút 30 giây. Sau đó mẫu được lấy ra và đặt ngay lên đá lạnh trong 5 phút. Bổ sung 500 μ l môi trường LB lỏng, hỗn hợp được nuôi ở 37°C, lắc 200 v/p trong 1 giờ. Sử dụng que cấy trải, trải 100-200 μ l dịch khuẩn lên đĩa môi trường LB đặc có chứa 50 mg/L carbenicillin. Ủ đĩa ở 37°C qua đêm.

2.2.3.4. *Tách chiết plasmid từ tế bào E. coli*

Tế bào vi khuẩn sau khi nuôi cấy lỏng trong môi trường LB được thu lại bằng cách ly tâm 5000 v/p trong 5 phút và được hòa vào 100 μ l dung dịch I (50 mM Tris-HCl, pH 8; 10 mM EDTA, pH 8; 50 mM glucose). Thêm 200 μ l dung dịch II (0,2 M NaOH; 1% SDS), đảo ống nhẹ nhàng. Thêm 150 μ l dung dịch III lạnh (3 M potassium acetate; 11,5% acetic acid) đảo nhanh và đặt trên đá trong khoảng 5 phút. Lúc này, xác tế bào và DNA hệ gen bị kết tủa lại thành mảng trắng được loại bỏ bằng cách ly tâm 13000 v/p trong 5 phút. Dung dịch pha tan được hút sang ống mới và hòa theo tỷ lệ 1:1 (v/v) với hỗn hợp chloroform:isoamyl alcohol 24:1 (v/v), lắc

đều và ly tâm 10 phút với tốc độ 13000 v/p nhằm loại bỏ triệt để các protein bám DNA. Dung dịch pha trên chứa DNA plasmid được thu lại. DNA plasmid được rửa bằng một lần thể tích isopropanol ở -20°C trong 1 giờ. Tiến hành ly tâm thu rửa DNA ở 13000 v/p trong 10 phút. Rửa DNA được rửa lại bằng cồn 70%, làm khô và hòa vào 40 µl đệm TE có chứa 0,1 mg RNase/ml giúp loại bỏ RNA tạp nhiễm. DNA plasmid được kiểm tra trên gel agarose 0,8% và được cất giữ ở -20°C.

2.2.3.5. Phương pháp xác định trình tự nucleotide của Sanger

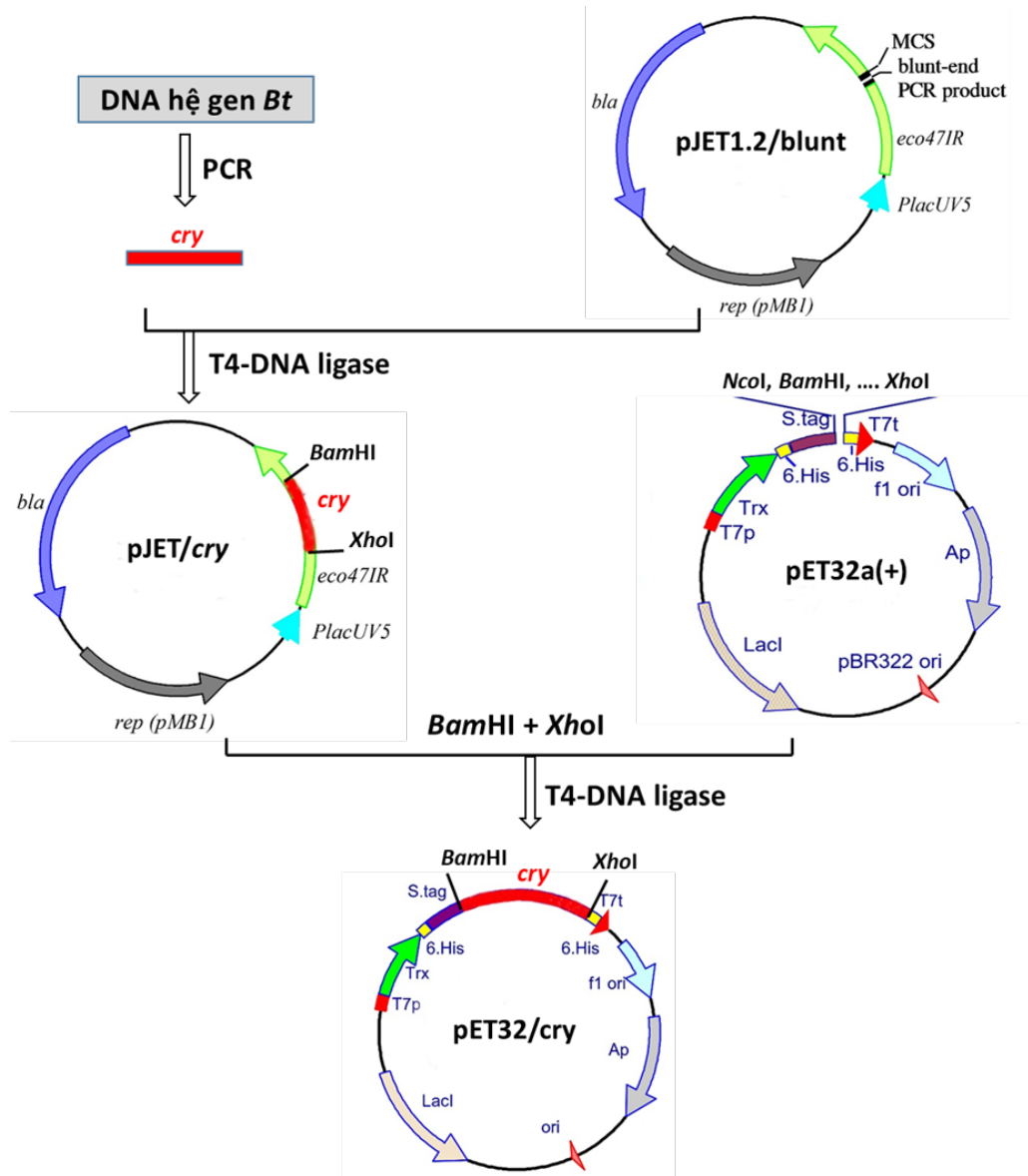
Đoạn gen quan tâm gắn trên vector tách dòng được xác định trình tự trên máy đọc tự động ABI PRIM[®] 3100 Avant Genetic Analyzer bằng cách sử dụng bộ hoá chất sinh chuẩn bigDye[®] Terminator v3.1 Cycle Sequencing. Phản ứng được thực hiện trong tổng thể tích 15 µl hỗn hợp gồm: dung dịch đệm (5X), 1 pmol mỗi, 3 µl bigDye và 100 ng plasmid. Chu trình nhiệt của phản ứng PCR đọc trình tự trên máy luân nhiệt GeneAmp[®] PCR System 9700 như sau: 96°C trong 1 phút, 25 chu kỳ (96°C trong 10 s, 55°C trong 5 s, 60°C trong 2 phút), 72°C trong 10 phút.

Sản phẩm PCR được tinh sạch bằng cách bổ sung 5 µl EDTA 125 mM, 60 µl cồn 100% và ủ ở nhiệt độ phòng trong 15 phút. Tiếp đó ly tâm 12.000 v/p trong 15 phút để kết tủa các đoạn DNA, sau đó loại bỏ cồn. Bổ sung 60 µl ethanol 70% và ly tâm 10.000 v/p trong 10 phút. Làm khô kết tủa DNA, bổ sung 10 µl Hi-Di[™] Formamide và biến tính ở 95°C trong 5 phút. Các mẫu được tra vào các giếng của khay đựng mẫu, sau đó điện di trong ống mao quản 80 cm x 50 µl với polymer POP-4[™] của hãng ABI, Mỹ. Kết quả được xử lý bằng phần mềm BioEdit v7.2.5.

2.2.4. Biểu hiện gen mã hóa protein độc tố diệt côn trùng trong hệ biểu hiện *E. coli* BL21 và đánh giá hoạt lực diệt sâu của protein tái tổ hợp

2.2.4.1. Thiết kế vector biểu hiện pET32a(+) mang gen *cry1Na*, *cry1Be*, *cry2Ab*, *cry2Ah*

Các gen *cry1Na*, *cry2Ab*, *cry2Ah* được nhân dòng trong vector pJET1.2/blunt, 2 đầu có mang điểm nhận biết của cặp enzyme *Bam*HI và *Xho*I. Vùng lõi độc bao gồm 3 vùng chức năng (domain I-III) của gen *cry1Be* được khuếch đại nhờ PCR sử dụng cặp mồi đặc hiệu cũng được thiết kế thêm điểm nhận



Hình 2.2. Sơ đồ quá trình nhân dòng và thiết kế các vector biểu hiện pET32a(+) mang các gen *cry*

biết của *Bam*HI và *Xho*I (bảng 2.1). Phân đoạn lõi độc *cry*/Be cùng các plasmid pJET/cry1Na, pJET/cry2Ab và pJET/cry2Ah được xử lý đồng thời bằng *Bam*HI và *Xho*I. Các đoạn gen quan tâm mang đầu so le nhờ xử lý bởi hai enzyme trên được phân tách trên gel agarose 0,8% và được tinh chế bằng bộ Kit QIAquick Gel Extraction, sau đó được nối ghép vào vector pET32a(+) cũng đã được cắt mở vòng bằng cùng loại enzyme tạo thành các vector biểu hiện pET32/cry1Na, pET32/cry1Be, pET32/cry2Ab và pET32/cry2Ah. Các vector tái tổ hợp này sau khi cắt kiểm tra bằng *Bam*HI và *Xho*I được giải trình tự để đảm bảo các gen đích được nối ghép đảm bảo đúng khung đọc và không xảy ra sai sót. Cuối cùng, các vector

được chuyển vào tế bào *E. coli* BL21 để phục vụ thí nghiệm biểu hiện gen. Quá trình thiết kế các vector biểu hiện pET32a(+) mang các gen cry được mô tả chi tiết ở hình 2.2.

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

Một khuẩn lạc đơn của chủng *E. coli* BL21 mang vector biểu hiện được nuôi trong 5 ml môi trường LB có bổ sung 100 µg/ml kháng sinh ampicillin ở tủ lạnh 37°C qua đêm. Dịch khuẩn sau đó được cấy chuyển theo tỉ lệ 1/100 vào môi trường LB mới và tiếp tục nuôi cấy ở 37°C đến khi OD₆₀₀ đạt khoảng 0,4 - 0,6. Để kích hoạt tế bào vi khuẩn sinh tổng hợp protein ngoại lai, dịch nuôi cấy được bổ sung IPTG đến nồng độ 0,1 mM. Sau 4 giờ nuôi cấy trong môi trường cảm ứng, tế bào vi khuẩn được thu lại bằng cách ly tâm 6000 v/p trong 5 phút ở 4°C và bảo quản ở -80°C để thực hiện các phân tích tiếp theo. Tế bào sau khi được giã đông được hòa vào đệm PBS (137 mM NaCl; 2,7 mM KCl; 10 mM Na₂HPO₄; 1,8 mM KH₂PO₄; pH = 7,4), sau đó bổ sung đệm xử lý mẫu 6X (375 mM Tris-HCl, 9% SDS (w/v), 50% Glycerol (v/v), 9% β-mercaptoethanol (v/v), 0.03% Bromophenol blue (w/v)) và ủ ở 95°C trong 15 phút. Dịch xử lý mẫu chứa protein tổng số của mỗi chủng biểu hiện sau đó được loại bỏ cặn tế bào bằng ly tâm 13000 vòng/phút trong 5 phút, sau đó được điện di kiểm tra trên SDS-PAGE.

Để kiểm tra tính tan của protein tái tổ hợp thu được ở các thể biến nạp *E. coli* BL21, tế bào vi khuẩn sau quá trình nuôi cấy cảm ứng được hòa trong đệm PBS lạnh và tiến hành siêu âm (đầu dò hoạt động ở mức 60 kHz trong 30s, nghỉ 1 phút, lặp lại 15 chu kỳ) đến khi dung dịch tế bào trở nên trong và không nhớt. Dịch phá tế bào sau đó được ly tâm 13000 v/p trong 15 phút ở 4°C. Phần dịch nổi được tách riêng khỏi cặn tế bào được gọi là pha protein tan. Cặn tế bào được hòa lại vào một lượng đệm PBS tương đương, được gọi là pha cặn chứa protein không tan. Cả hai pha protein này sẽ được kiểm tra bằng SDS-PAGE.

2.2.4.3. Điện di protein trên SDS-PAGE

Phương pháp điện di kiểm tra protein trên SDS-PAGE được tiến hành theo Laemmli (1970) gồm có các hóa chất sau: dung dịch polyacrylamide 30%; đệm Tris-HCl 1,5 M, pH 8,8; đệm Tris-HCl 0,5 M, pH 6,8; SDS 10%; APS 10%, TEMED, đệm

chạy điện di (Tris-HCl 0,05 M; glycine 0,192 M; SDS 0,1%); đệm xử lý mẫu trước khi điện di 5X (Tris-HCl 250 mM; SDS 10%; glycerol 50%; bromophenol blue 0,05%; DTT 500 mM; pH 6,8); dung dịch nhuộm gel (methanol 30%; acetic acid 10%; coomassie brilliant blue 0,025%); dung dịch rửa gel (methanol 30%; acetic acid 10%). Dịch chứa protein trước khi điện di được bổ sung đệm xử lý mẫu và ủ ở 100°C trong 15 phút. Quá trình điện di được thực hiện trên bộ điện di đứng của hãng Bio-Rad sử dụng bản gel polyacrylamide gồm 2 phần: phần trên là gel 5% để cô mẫu, phần dưới là gel 10% để phân tách protein theo khối lượng phân tử. Bản gel sau đó được nhuộm với dung dịch nhuộm trong 30 phút và được rửa với dung dịch rửa cho đến khi các băng vạch protein trở nên rõ nét.

2.2.4.4. Phương pháp lai miễn dịch Western blot

Hóa chất và dung dịch: đệm chuyển màng (Tris 25 mM, glycine 250 mM, SDS 0,2% (w/v), methanol 20%); đệm TBS (Tris 10 mM, NaCl 150 mM, pH 7,5); đệm TTBS (Tris 10 mM, NaCl 150 mM, Tween 20 0,05%, pH 7,5); dung dịch đệm phủ màng (blocking buffer) (5% sữa tách béo pha trong đệm TBS); dung dịch kháng thể 1: kháng thể 1 (anti-6X His tag) pha loãng 1000X trong dung dịch sữa tách béo 5%; dung dịch kháng thể 2: kháng thể 2 (anti-mouse IgG cộng hợp HRP) pha loãng 10000X trong dung dịch sữa tách béo 5%; dung dịch hiện màu (50 mg DAB (3,3'-diaminobenzidine) hòa tan trong 100 ml đệm TBS có bổ sung 10 μ l 30% H₂O₂).

Qui trình: Protein sau khi được phân tách trên gel polyacrylamide bằng điện di SDS-PAGE được thấm chuyển lên màng nitrocellulose bằng đệm chuyển màng ở hiệu điện thế 200 V trong 2 giờ. Tráng màng vài lần với nước khử ion, sau đó ngâm vào dung dịch đệm phủ màng, lắc qua đêm ở điều kiện lạnh nhằm khóa toàn bộ các vùng trên màng không có protein gắn vào để ngăn chặn bất kỳ sự liên kết không đặc hiệu nào giữa kháng thể với bề mặt của màng. Màng sau đó được ủ với dung dịch kháng thể 1, lắc nhẹ trong 3 giờ. Rửa màng bằng đệm TTBS trong 5 phút, lặp lại 3 lần. Màng tiếp tục được ủ với dung dịch kháng thể 2 và lắc nhẹ ở nhiệt độ phòng. Sau 1 giờ, màng được rửa như đã mô tả ở trên trước khi được ngâm trong dung dịch hiện màu trong 10 phút. Dừng phản ứng bằng cách rửa màng nhiều lần với nước khử ion. Chụp ảnh kết quả bằng máy ảnh thông thường.

2.2.4.5. Tinh sạch protein tái tổ hợp bằng phương pháp sắc ký ái lực

Protein dạng dung hợp Trx-His-Stag-Cry2Ab biểu hiện trong pha tan được tinh sạch bằng phương pháp sắc ký ái lực ion kim loại cố định sử dụng cột Ni-NTA. Tế bào vi khuẩn sau khi được nuôi cấy trong điều kiện cảm ứng với IPTG được thu lại bằng ly tâm và hòa trong đệm gắn (binding buffer) (20 mM NaH₂PO₄; 0,5 M NaCl; 20 mM imidazole; pH 8,0). Tiến hành siêu âm phá tế bào, sau đó ly tâm ở 13000 v/p trong 30 phút để loại bỏ xác tế bào cùng các protein ở pha cặn. Dịch protein pha tan được ủ với Ni-NTA, trộn nhẹ nhàng ở 4°C trong 1 giờ, cho phép protein tái tổ hợp Trx-His-Stag-Cry2Ab39 có chứa 6X His tag liên kết ái lực với hạt Ni-NTA. Phức hợp protein/Ni-NTA sau đó được đưa lên cột lọc sắc ký và được rửa với dung dịch đệm rửa (washing buffer) (20 mM NaH₂PO₄; 0,5 M NaCl, 40 mM imidazole; pH 8,0) để loại bỏ các protein tạp nhiễm liên kết không đặc hiệu. Cuối cùng, protein dung hợp Trx-His-Stag-Cry2Ab được tách khỏi cột bằng đệm elution buffer (20 mM NaH₂PO₄; 0,5 M NaCl; 500 mM imidazole; pH 8,0). Protein được thu theo phân đoạn, mỗi phân đoạn 1 ml được tách riêng vào các ống riêng biệt và được kiểm tra bằng điện di SDS-PAGE. Mức độ tinh sạch của protein trong các phân đoạn được đánh giá bằng hệ thống VersaDoc Imaging. Các phân đoạn chứa protein tái tổ hợp ít bị tạp nhiễm được gộp chung để tiến hành loại muối cùng imidazole và cô đặc bằng cột cô đặc protein 3K MWCO (hãng Thermo Scientific). Nồng độ của protein tái tổ hợp đã tinh chế được xác định bằng phương pháp Bradford [81] dựa trên đường chuẩn được xây dựng với protein chuẩn BSA. Protein thành phẩm được bảo quản trong glycerol 50% (v/v) ở -20°C phục vụ cho thí nghiệm đánh giá hoạt lực diệt sâu đục quả *E. zinckenella* Treitschke.

2.2.4.6. Đánh giá khả năng diệt ấu trùng sâu đục quả *E. zinckenella* Treitschke của protein tái tổ hợp

*Thử nghiệm khả năng diệt sâu *E. zinckenella* của protein tái tổ hợp:* các chủng *E. coli* BL21 biểu hiện các protein độc tố tái tổ hợp Cry1Be, Cry1Na, Cry2Ab và Cry2Ah được nuôi cấy ở điều kiện cảm ứng và thu sinh khối của 150 ml dịch khuẩn sau 4 giờ. Tế bào *E. coli* sau đó được hòa trong 10 ml dung dịch EB và siêu âm phá tế bào để thu dịch protein tổng số có chứa các protein mong muốn. Bổ sung 2 ml dịch protein tổng số vào cốc chứa 20 g môi trường thức ăn nhân tạo. Mỗi

cốc được thả 10 cá thể sâu non tuổi 2-3 để thử nghiệm. Song song với các công thức thí nghiệm, thực hiện tương tự với các công thức đối chứng sử dụng dịch protein tổng số của vi khuẩn *E. coli* BL21, dung dịch đệm EB và nước deion. Thí nghiệm được lặp lại 3 lần. Theo dõi và thống kê số lượng sâu chết sau 7 ngày.

Xác định giá trị liều gây chết trung bình (LC₅₀) của protein tái tổ hợp dạng tinh sạch: Dịch protein Trx-His-Stag-Cry2Ab được pha loãng bằng đệm PBS và nhỏ đều lên phần thức ăn nhân tạo chứa trong mỗi hộp thử nghiệm sao cho đạt đến nồng độ cuối cùng lần lượt là 0,5; 1,0; 1,5; 2,0; 3,0; 3,5 và 4 µg protein/g thức ăn, để khô tự nhiên ở nhiệt độ phòng. Ở lô đối chứng dương và âm, sử dụng dịch canh khuẩn *Bt* SP14.2 và đệm PBS, theo thứ tự tương ứng, để thay thế cho dịch protein tinh sạch. Với mỗi hộp thử nghiệm, 10 ấu trùng *E. zinckenella* Treitschke tuổi 2 được thả vào để ăn thức ăn nhân tạo. Các hộp được giữ 28°C ± 2°C trong 7 ngày. Quan sát đĩa hàng ngày để thống kê số sâu chết. Thí nghiệm được lặp lại 3 lần. Liều gây chết trung bình (LC₅₀) sau 7 ngày của protein tái tổ hợp với ấu trùng sâu đục quả và giới hạn tin cậy (fiducial limits) được xác định bằng phân tích Probit [82] sử dụng phần mềm thống kê SPSS của IBM (phiên bản 20).

2.2.5. Cải biến mã di truyền và biểu hiện gen *cry2Ab39* trong lá cây thuốc lá *N. benthamiana* bằng công nghệ biểu hiện gen tạm thời thông qua *Agrobacterium tumefaciens*

2.2.5.1. Tối ưu mã gen độc tố *cry2Ab39* để thích hợp biểu hiện trong thuốc lá *N. benthamiana*

Căn cứ vào bảng tần suất sử dụng các codon của *N. benthamiana* trên ngân hàng dữ liệu mã di truyền Codon Usage Database (<https://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=4100>) [83], chúng tôi loại bỏ các codon hiếm (tần suất sử dụng dưới 50%) trong trình tự của gen *cry2Ab39* kiểu dại và thay thế bằng các codon được loài thực vật này ưu tiên sử dụng mà không làm thay đổi trình tự axit amin. Công cụ web Seqool (<http://www.biossc.de/seqool/index.html>) được sử dụng để tìm các vùng giàu AT, các trình tự tín hiệu polyadenyl hóa tiềm năng ở thực vật cũng như các mô típ trình tự có khả năng gây ra sự bất ổn cho RNA ở thực vật [84]. Các trình tự này sau đó

được thay thế thủ công bằng các codon thoái hóa. Ngoài ra, phần mềm NUPACK66 (<http://www.nupack.org>) được sử dụng để tiên đoán và loại trừ các khả năng hình thành cấu trúc thứ cấp dạng vòng ổn định của phân tử mRNA. Với mục đích giữ lại protein tái tổ hợp ở mạng lưới nội chất (ER) của tế bào thực vật chủ, chúng tôi đã sử dụng đoạn trình tự gồm 69 nucleotide mã hóa cho chuỗi peptide tín hiệu của gen LeB4 (quy định protein legumin type B ở cây đậu răng ngựa *Vicia faba*) và đoạn mã hóa trình tự KDEL (tín hiệu lưu giữ protein trên ER) để dung hợp với gen *cry2Ab39* sau khi đã tối ưu mã tại vị trí đầu 5' và 3', theo thứ tự tương ứng. Sau vị trí KDEL, trình tự mã hóa 6X His tag cũng được thêm vào để tạo thuận lợi cho việc phát hiện cũng như tinh sạch protein tái tổ hợp. Mặt khác, để đơn giản hóa việc nối ghép gen trong quá trình thiết kế vector chuyển gen vào thực vật, hai vị trí nhận biết của *NcoI* và *XmaI* được thêm vào tận cùng đầu 5' và 3' của gen *cry2Ab39*. Gen đích sau quá trình cải biến được được kí hiệu là *cry2Ab39opt* và được tổng hợp nhân tạo bởi GenScript (Hong Kông), đồng thời được nhân dòng trong vector pUC57.

2.2.5.2. Thiết kế vector chuyển gen pFGC/35S-*cry2Ab39opt*

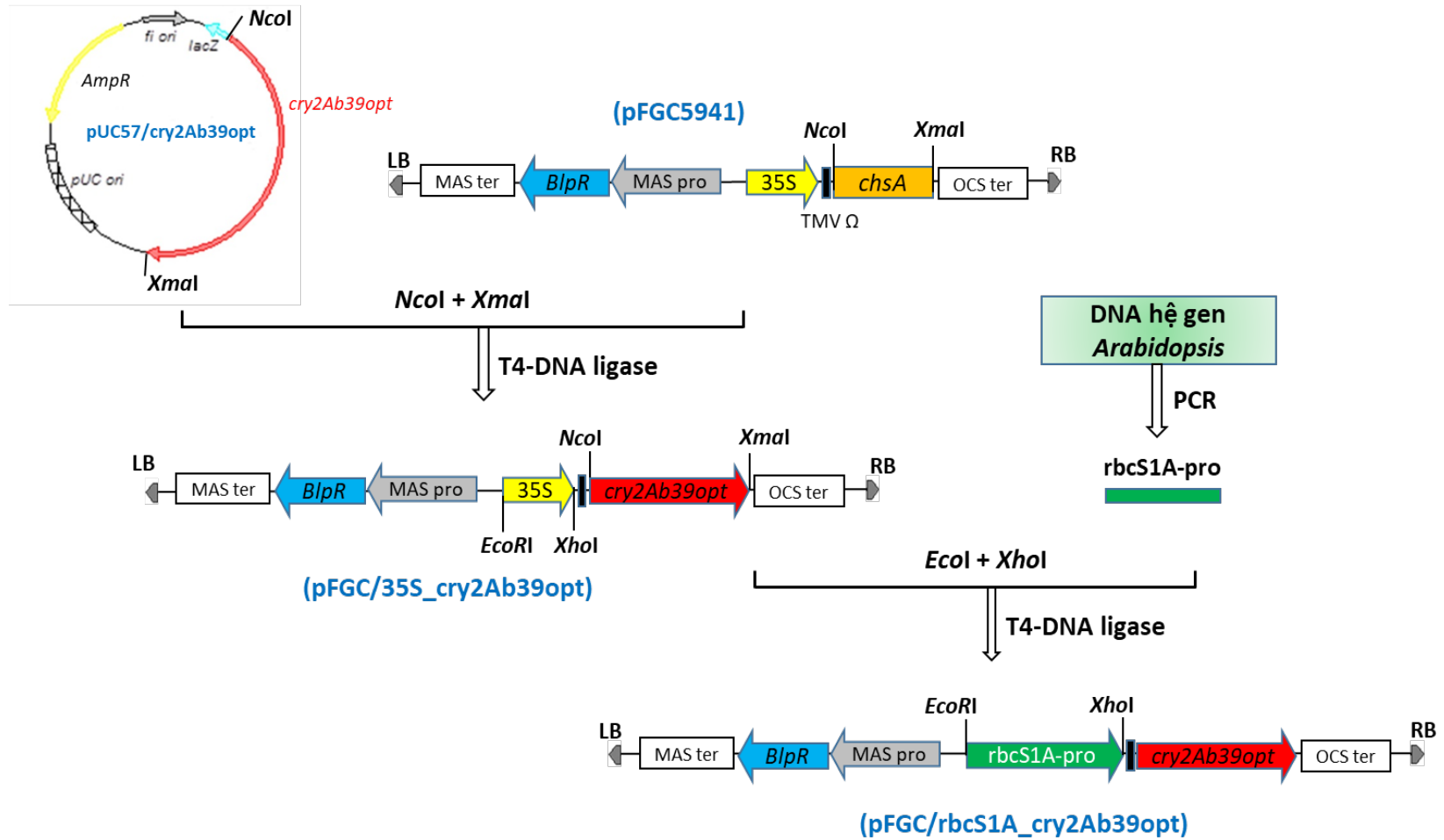
Plasmid pUC57/*cry2Ab39opt* mang gen nhân tạo *cry2Ab39opt* và vector pFGC5941 đồng thời được xử lý bằng cặp enzyme hạn chế *NcoI* và *XmaI*. Sản phẩm tinh sạch của đoạn gen *cry2Ab39opt* mang đầu so le và vector pFGC5941 đã mở vòng được sử dụng làm nguyên liệu cho phản ứng nối ghép dưới sự xúc tác của T4-DNA ligase tạo thành vector pFGC/35S-*cry2Ab39opt*. Vector tái tổ hợp này sau khi được kiểm tra bằng phản ứng PCR nhân gen *cry2Ab39opt* và cắt bằng cặp enzyme hạn chế *NcoI/XmaI* được chuyển vào tế bào *A. tumefaciens AGL1* phục vụ thí nghiệm biểu hiện tạm thời gen đích trên cây thuốc lá. Quá trình thiết kế vector chuyển gen pFGC/35S-*cry2Ab39opt* được mô tả ở hình 2.3.

2.2.5.3. Thiết kế vector chuyển gen pFGC/*rbcS1A-pro-cry2Ab39opt*

Từ DNA hệ gen của *Arabidopsis thaliana*, đoạn promoter của gen *rbcS1A* mã hóa cho ribulose-1,5-bisphosphate carboxylase-oxygenase (RuBisCO) được khuếch đại bằng kỹ thuật PCR với cặp môi đặc hiệu (Bảng 2.1). Đoạn promoter này

sau đó được tách dòng bằng cách gắn vào pJET1.2/blunt để tạo thành vector pJET/rbcS1A-pro và được kiểm tra bằng phương pháp giải trình tự của Sanger.

Vector tách dòng pJET/rbcS1A-pro và vector pFGC/35S_cry2Ab39opt sau đó được xử lý đồng thời bằng hai enzyme hạn chế *EcoRI* và *XhoI*. Đoạn rbcS1A-pro mang đầu so le và vector pFGC/35S_cry2Ab39opt đã cắt bỏ phân đoạn promoter 35S được phân tách trên gel agarose 0,8%, tinh sạch và nối ghép dưới sự xúc tác của T4-DNA ligase để tạo thành vector pFGC/rbcS1A-pro_cry2Ab39opt. Vector tái tổ hợp này sau khi được kiểm tra bằng phản ứng PCR nhân phân đoạn rbcS1A-pro và cắt bằng cặp enzyme hạn chế *EcoRI/XhoI* được chuyển vào tế bào *A. tumefaciens* AGL1 phục vụ thí nghiệm biểu hiện tạm thời gen đích trên cây thuốc lá. Quá trình thiết kế vector chuyển gen pFGC/rbcS1A-pro_cry2Ab39opt được mô tả ở hình 2.3.



Hình 2.3. Sơ đồ quá trình thiết kế các vector chuyển gen pFGC/35S_cry2Ab39opt và pFGC/rbcS1A-pro_cry2Ab39opt

2.2.5.4. Tạo chủng vi khuẩn *Agrobacterium* mang vector chuyển gen

Các plasmid pFGC/35S_cry2Ab39opt và pFGC/rbcS1A-pro_cry2Ab39opt được biến nạp vào chủng vi khuẩn *Agrobacterium tumefaciens* AGL1 bằng phương pháp xung điện [85]. Những dòng khuẩn *A. tumefaciens* cho kết quả dương tính với phản ứng PCR khuếch đại gen *cry2Ab39opt* sẽ được nuôi cấy trong 5 ml môi trường LB có bổ sung kháng sinh chọn lọc qua đêm ở tủ lạnh 28°C. Lưu giữ dịch vi khuẩn trong glycerol 20% và đặt trong tủ lạnh sâu -80°C.

2.2.5.5. Xâm nhiễm vi khuẩn *Agrobacterium* vào cây thuốc lá bằng phương pháp hút chân không.

Chủng *Agrobacterium tumefaciens* AGL1 mang vector nhị thể pFGC/35S_cry2Ab39opt hoặc pFGC/rbcS1A-pro_cry2Ab39opt được dùng để chuyển gen vào thực vật. Chủng *A. tumefaciens* AGL1 chứa vector pBI121/35S_HcPro mang gen mã hóa cho protein HcPro - một proteinase của virus PVY, có chức năng ức chế cơ chế câm lặng RNA ở thực vật bằng cách ngăn cản sự phân giải các mRNA và RNA sợi đôi từ đó giảm lượng siRNA hoặc làm hỏng chức năng cắt của enzyme cắt sợi đôi Dicer và RISC [86] – được dùng để đồng biến nạp, giúp tăng cường sự biểu hiện của protein Cry2Ab39. Các chủng này được nuôi trên môi trường LB đặc có bổ sung kháng sinh kanamycin 50 mg/L và rifamycin 50 mg/L trong tủ ấm 28°C trong 2 ngày để tạo khuẩn lạc. 1 khuẩn lạc của mỗi chủng được nuôi riêng rẽ trong 10 ml LB có chứa kháng sinh chọn lọc, đặt trong tủ lạnh 28°C qua đêm. Ngày tiếp theo, toàn bộ dịch khuẩn được chuyển sang bình nuôi cấy chứa 250 ml môi trường LB và tiếp tục được nuôi trong tủ lạnh 28°C thêm 14-16 giờ cho đến khi OD₆₀₀ đạt 1,8. Tiến hành thu tế bào vi khuẩn bằng cách ly tâm 5000 vòng/phút ở 4°C trong 15 phút. Sau khi loại bỏ dịch nổi, cặn khuẩn *A. tumefaciens* chứa vector pFGC/35S_cry2Ab39opt hoặc cặn khuẩn chứa vector pFGC/rbcS1A-pro_cry2Ab39opt được trộn lẫn với cặn khuẩn mang cấu trúc biểu hiện protein HcPro, sau đó được hòa trong đệm MES (10 mM MgCl₂; 10 mM 2-(N-morpholino) ethanesulfonic acid (MES); pH = 5,6) để đạt OD₆₀₀ = 0,5.

Hỗn hợp dịch khuẩn biến nạp được đặt trong một bình hút chân không bằng thủy tinh. Cây thuốc lá *N. benthamiana* trồng thủy canh 4 tuần tuổi sau khi loại bỏ

những lá già, lá non, chỉ giữ lại lá bánh tẻ được úp ngược ngập nhúng hoàn toàn vào dịch biến nạp, phần rễ cây được ngăn cách với dịch biến nạp nhờ một tấm chắn. Sau đó, nắp bình hút chân không được đặt lên, van khí đóng lại và máy hút chân không được bật với thời gian hút chân không là 1 phút 30 giây. Cây sau khi biến nạp được đặt trở lại hệ thống nuôi cây thủy canh. Các lá biến nạp được thu hoạch ở 5 ngày sau biến nạp và cất giữ ở tủ -80°C phục vụ cho thí nghiệm phân tích sự biểu hiện của gen đích.

Để tìm ra điều kiện tối ưu nhằm nâng cao mức độ biểu hiện tạm thời của protein Cry2Ab39 tái tổ hợp trong lá thuốc lá *N. benthamiana*, sự tác động của các yếu tố khác nhau bao gồm cấu trúc chuyển gen (pFGC/35S_cry2Ab39opt, pFGC/rbcS1A-pro_cry2Ab39opt), mật độ khuẩn của hỗn hợp dịch khuẩn biến nạp ($OD_{600} = 0,1; 0,3; 0,5; 0,8$) và thời gian thu hoạch lá (1, 2, 3, 4, 5, 6, 7 ngày sau biến nạp) đã được xem xét, so sánh và đánh giá. Mức độ biểu hiện của protein mục tiêu được phân tích bằng kỹ thuật Western blot và quy trình biểu hiện tối ưu sau khi xác định sẽ được áp dụng trên quy mô lớn hơn để cung cấp nguyên liệu lá cho thí nghiệm tinh chế protein.

2.2.5.6. Tách chiết và kiểm tra sự biểu hiện tạm thời của protein Cry2Ab39 tái tổ hợp trong lá thuốc lá *N. benthamiana* đã gây nhiễm.

Nghiên mịn 1 g mẫu lá trong nitor lỏng, bổ sung đệm PBS 1X, pH 7,4 theo tỉ lệ 3 ml đệm/ 1 g mẫu. Dịch chiết protein tổng số sau đó được ly tâm ở 13000 v/p trong 30 phút ở 4°C để loại bỏ xác tế bào. Nồng độ protein tổng số trong dịch chiết được định lượng bằng phương pháp Bradford [81] dựa trên đường chuẩn được xây dựng với protein chuẩn BSA. 30 μ g protein tổng số mỗi mẫu lá được phân tách trên gel polyacrylamide 10% bằng điện di SDS-PAGE (đã trình bày trong mục 2.2.3.9), sau đó được chuyển lên màng nitrocellulose để kiểm tra sự có mặt của protein tái tổ hợp bằng kỹ thuật lai Western blot sử dụng kháng thể kháng His-tag (đã trình bày trong mục 2.2.3.10). Mức độ biểu hiện của protein tái tổ hợp ở mỗi mẫu lá được bán định lượng dựa vào độ đậm nhạt của băng vạch hiện trên màng nitrocellulose qua phân tích sử dụng phần mềm ImageJ.

2.2.5.7. Tinh sạch protein Cry2Ab39 tái tổ hợp biểu hiện trong lá thuốc lá

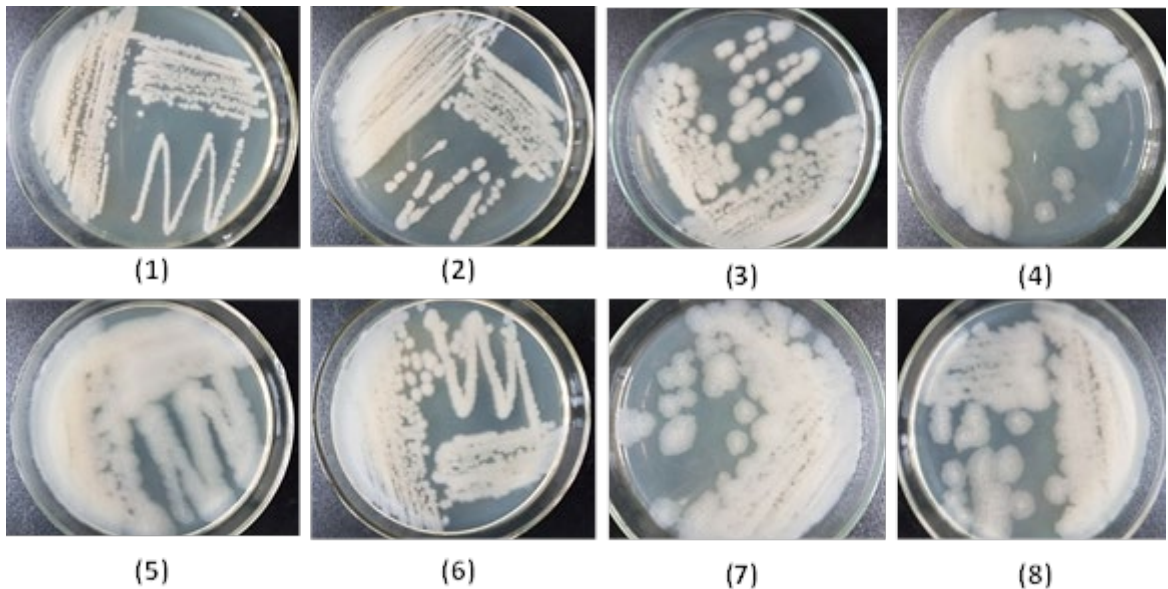
Protein Cry2Ab39 biểu hiện trong lá cây *N. benthamiana* được tinh chế bằng phương pháp sắc ký ái lực ion kim loại cố định (IMAC) sử dụng cột Ni-NTA, áp dụng với các protein tái tổ hợp có chứa trình tự 6xHis. 100 gam lá thuốc lá được nghiền trong nitor lỏng và trộn trong 300 ml dung dịch đệm gắn (NaH_2PO_4 20 mM, NaCl 0,5 M, imidazole 20 mM, pH 8.0). Dịch chiết thô được ly tâm hai lần ở 13.000 vòng/phút, 4°C trong 30 phút để loại bỏ xác tế bào. Phần dịch nổi thu được sau đó được trộn với 10 ml nhựa Ni-NTA, trộn nhẹ nhàng ở 4°C trong 1 giờ, cho phép protein tái tổ hợp Cry2Ab39 có chứa 6xHis liên kết ái lực với hạt Ni-NTA. Các bước tiếp theo được thực hiện như quy trình đã được mô tả trong mục 2.2.4.5.

Chương 3. KẾT QUẢ NGHIÊN CỨU

3.1. Kết quả sàng lọc chủng *Bt* có hoạt lực diệt sâu đục quả đậu tương cao từ các chủng trong Bộ sưu tập *Bt* bản địa Việt Nam

3.1.1. Kết quả hoạt hóa và nhân nuôi thu sinh khối các chủng *Bt*

Tiến hành hoạt hóa 221 chủng *Bt* (phân lập từ 4 tỉnh miền núi phía Bắc bao gồm Lào Cai, Tuyên Quang, Hà Giang và Điện Biên) thuộc Bộ sưu tập vi khuẩn *Bt* bản địa Việt Nam của Trung tâm Giống và Bảo tồn nguồn gen Vi sinh vật, Viện Công nghệ Sinh học. Các chủng bảo quản trong glycerol 20% ở -80°C được rã đông trên đá và tiến hành cấy rìa trên đĩa môi trường thạch LB và nuôi ở 28°C trong 2 ngày. Kết quả quan sát cho thấy tất cả các chủng được hoạt hóa đều tạo khuẩn lạc riêng rẽ với các đặc điểm hình thái đặc trưng cho vi khuẩn *Bacillus* như khuẩn lạc tròn, rìa răng cưa không đều, màu vàng xám hoặc vàng nhạt... Các khuẩn lạc của mỗi chủng đều có sự đồng nhất về hình thái cho thấy độ thuần của chủng, không bị lẫn các chủng tạp nhiễm khác (Hình 3.1).



Hình 3.1. Hình thái khuẩn lạc đặc trưng của một số chủng *Bt* được hoạt hóa trên môi trường thạch LB.

1-8: Các chủng *Bt* Đ6.1, SP14.2, TC1.3, BD8.2, Đ6.2, TC9.1, SP12.1, ĐC2.3

Từ các đĩa khuẩn *Bt* đã được hoạt hóa, với mỗi chủng, chọn 1 khuẩn lạc nuôi nhân giống cấp 1 trong 5 ml môi trường LB. Sau 24 giờ nuôi cấy, dịch nhân giống được đo OD ở bước sóng 600 nm để xác định mật độ tế bào. Kết quả thu được các chủng có OD₆₀₀ đều đạt từ 0,6 – 0,8. Tiến hành cấy chuyển 5 ml dịch nhân giống cấp 1 sang 50 ml môi trường LB lỏng, nhân nuôi để thu sinh khối. Sau 72 giờ, thu dịch canh khuẩn và tiến hành xác định mật độ tế bào bằng phương pháp đo OD₆₀₀. Kết quả phép đo cho thấy, dịch nuôi các chủng *Bt* đều có OD₆₀₀ ≥ 2 (tương đương mật độ tế bào từ 10⁷ – 10⁹ cFu/ml), đảm bảo phục vụ cho nghiên cứu thử hoạt tính diệt sâu đục quả đậu tương tiếp theo (Bảng 3.1, Bảng phụ lục 1). Dịch canh khuẩn chứa sinh khối các chủng *Bt* được bảo quản trong tủ lạnh 4°C để sử dụng cho các nghiên cứu tiếp theo.

Bảng 3.1. Mật độ tế bào của một số chủng *Bt* nghiên cứu sau 72 giờ nuôi cấy

STT	Tên chủng	Địa điểm/Nguồn phân lập	OD ₆₀₀ (72 giờ)
1	SP12.1	Sa Pa (Lào Cai)	2,770
2	SP14.2	Sa Pa (Lào Cai)	3,846
3	TC3.6	Lào Cai	2,550
4	LVT2.4	Lào Cai	3,247
5	Đ6.1	Tuyên Quang	2,325
6	Đ6.2	Tuyên Quang	3,490
7	ĐC1.6	Tuyên Quang	3,406
8	ĐC2.7	Tuyên Quang	3,421
9	DHG15.1	Hà Giang	2,837
10	DHG7.1	Hà Giang	2,831
11	LHG10.3	Hà Giang	4,746
12	LHG3.2	Hà Giang	3,240
13	BD23.7	Điện Biên	2,846
14	ĐB4.3	Điện Biên	3,864
15	ĐB9.1	Điện Biên	4,012
16	BD8.2	Điện Biên	3,465

3.1.2. Kết quả thử khả năng diệt sâu đục quả đậu tương của các chủng *Bt* nghiên cứu.

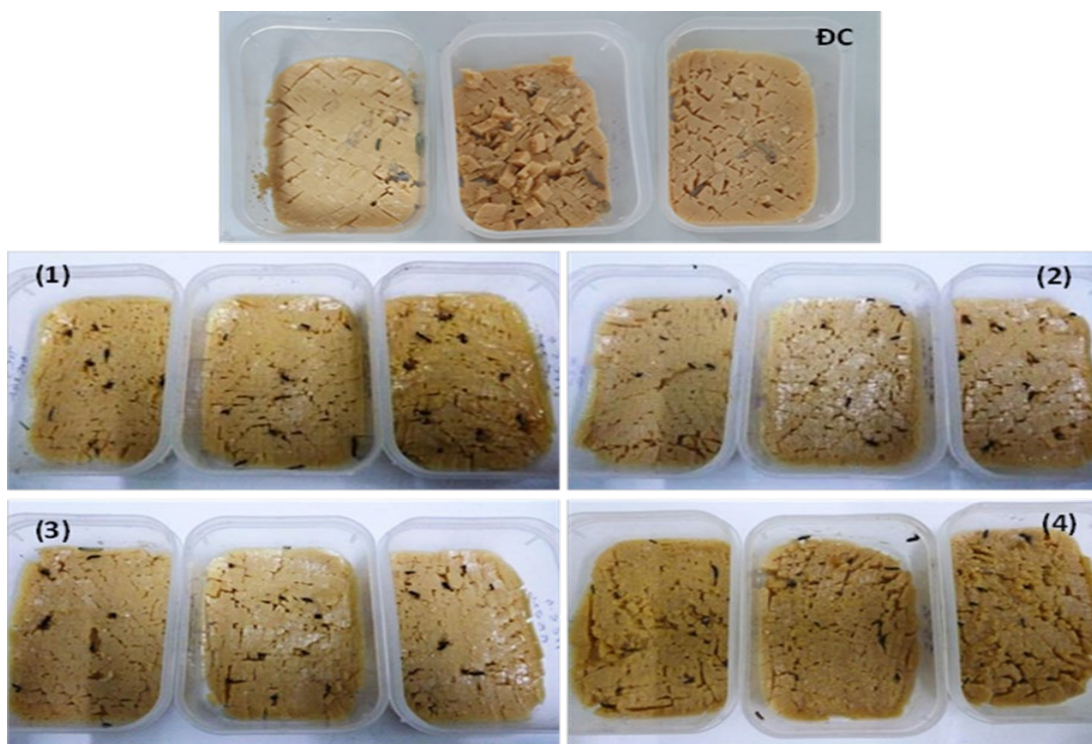
Sử dụng môi trường thức ăn nhân tạo để thử hoạt tính diệt sâu đục quả đậu tương của dịch sinh khối 221 chủng *Bt* nghiên cứu. Tiến hành xác định tỉ lệ phối trộn thể tích dịch canh khuẩn với khối lượng thức ăn. Kết quả sau khi thử các tỉ lệ

phối trộn khác nhau như 1%, 5%, 10% và 15%, chúng tôi nhận thấy với tỉ lệ 10% (2 ml dịch vi khuẩn/20 g thức ăn) là thích hợp nhất vì lượng dịch và mật độ tế bào vi khuẩn vừa đủ để thấm đều vào thức ăn, vừa tạo độ ẩm thích hợp (60 - 70%) cho sâu non phát triển bình thường. Với các tỉ lệ phối trộn thấp hơn (1% và 5%), lượng dịch chưa đủ để thấm đều vào thức ăn, ảnh hưởng đến tính chính xác của thí nghiệm. Trong khi đó, ở tỉ lệ phối trộn cao hơn (15%), độ ẩm của thức ăn nuôi sâu rất cao, không phù hợp cho sự phát triển của sâu non khi đánh giá, dẫn đến tỉ lệ sâu chết cơ học cao.

Tiến hành thử hoạt tính diệt sâu non đục quả đậu tương tuổi 2-3 của các chủng *Bt* với công thức 10% dịch khuẩn/thức ăn. Kết quả thống kê số sâu chết do ăn phải độc tố *Bt* sau 7 ngày thử nghiệm (Bảng 3.2, Bảng phụ lục 1) cho thấy: trong số 221 chủng *Bt* nghiên cứu, nhóm có hoạt tính diệt sâu từ 0-20% gồm 138 chủng, chiếm tỉ lệ cao nhất (62,44%), tiếp đó là các nhóm có hoạt tính diệt sâu từ 21-70% chiếm 31,22% với 69 chủng. Số chủng *Bt* có độc tính cao với sâu đục quả đậu tương sàng lọc được chiếm tỉ lệ khá thấp, trong đó 10 chủng có hoạt tính từ 71 - 84% (chiếm 4,52%) và chỉ có 4 chủng (chiếm 1,81%) cho tỉ lệ sâu chết trên 85%. Ở các đĩa thử nghiệm bổ sung dịch khuẩn *Bt*, trường hợp sâu chết do tác động của độc tố *Bt* thường có biểu hiện đặc trưng như: xác sâu màu đen, hình dạng sâu chết vẫn nguyên vẹn, không bị hoại mục, vỡ nát. Trong khi đó, ở các đĩa đối chứng sử dụng nước deion để thay thế dịch canh khuẩn, sâu non vẫn sống và phát triển bình thường. Sâu có xu hướng đục lỗ trên bề mặt môi trường thức ăn nhân tạo để chui sâu xuống dưới (Hình 3.2).

Bảng 3.2. Hoạt tính diệt sâu đục quả đậu tương của các chủng *Bt* nghiên cứu

Hoạt tính diệt sâu	Số chủng				Tổng số	Tỉ lệ (%)
	Lào Cai	Tuyên Quang	Hà Giang	Điện Biên		
0-20%	79	11	36	12	138	62,44
21-40%	25	3	10	2	40	18,10
41-60%	12	1	7	3	23	10,41
61-70%	2	1	1	2	6	2,71
71-84%	4	0	4	2	10	4,52
85-100%	1	2	0	1	4	1,81
Tổng số	123	18	58	22	221	100



Hình 3.2. Hình ảnh thử hoạt tính diệt sâu non đục quả đậu tương tuổi 2-3 của các chủng *Bt* có hoạt tính cao

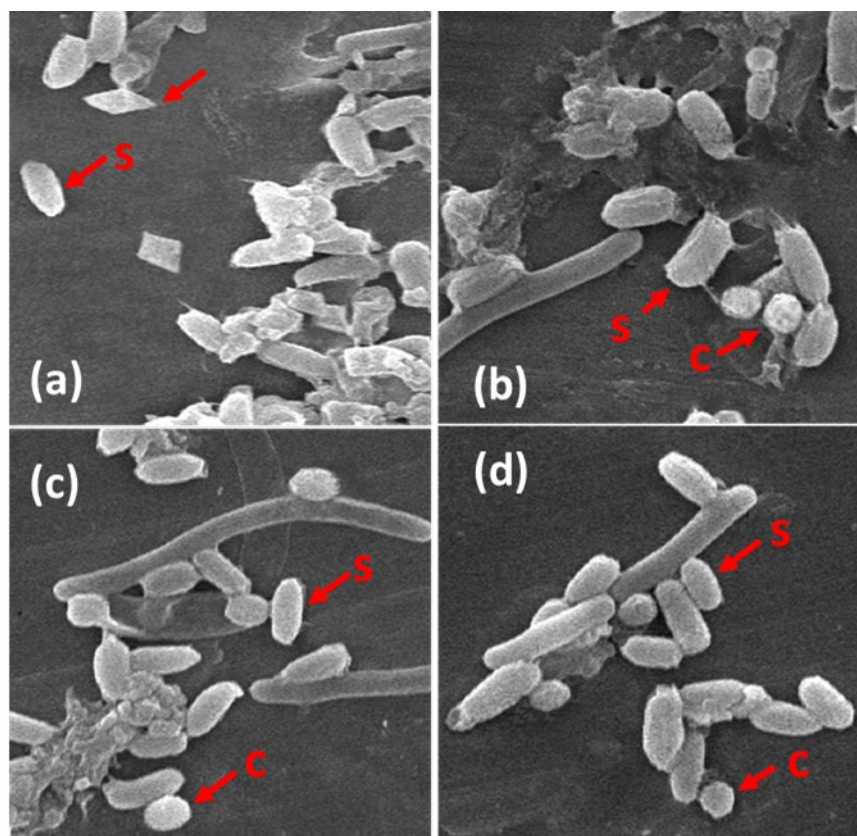
ĐC: nước deion thay thế cho dịch canh khuẩn. (1): chủng Đ6.1 (hoạt tính diệt sâu: 93,33%); (2): chủng Đ6.2 (hoạt tính diệt sâu: 100%); (3): chủng BD8.2 (hoạt tính diệt sâu: 100%); (4): chủng SP14.2 (hoạt tính diệt sâu: 93,33%)

Như vậy, từ 221 chủng sau khi sàng lọc đã tuyển chọn được 4 chủng có hoạt tính diệt sâu đục quả đậu tương $\geq 85\%$, bao gồm chủng SP14.2 phân lập từ tỉnh Lào Cai (tỉ lệ diệt sâu là 93,33%); hai chủng Đ6.1 và Đ6.2 phân lập từ Tuyên Quang (tỉ lệ diệt sâu lần lượt là 93,33% và 100%); và chủng BD8.2 phân lập từ Điện Biên (tỉ lệ diệt sâu là 100%). Bốn chủng này tiếp tục được sử dụng để nghiên cứu đánh giá đặc điểm sinh tinh thể độc tố và động thái sinh trưởng, cũng như sàng lọc gen gây chết sâu.

3.1.3. Kết quả xác định đặc điểm tinh thể độc tố và động thái sinh trưởng gây chết sâu của các chủng vi khuẩn *Bt* có hoạt tính diệt sâu đục quả đậu tương $\geq 85\%$

Một đặc điểm đặc trưng của các chủng *Bt* là khả năng tạo bào tử và sinh tinh thể độc tố. Thông thường các chủng *Bt* sau khi nuôi cấy 48 - 72 giờ thì hình thành bào tử và tinh thể. Mật độ bào tử trung bình có thể đạt 10^8 bào tử/ml, lượng tinh thể

giai đoạn này có thể chiếm khoảng 75% tổng tinh thể tạo thành. Sau 72 - 96 giờ, lượng tinh thể đạt cực đại, bào tử không tăng, tế bào giảm. Do vậy, sự hình thành bào tử và đặc điểm tinh thể của 4 chủng *Bt* có hoạt lực diệt sâu đục quả đậu tương cao đã được xác định bằng kính hiển vi điện tử quét. Kết quả chụp SEM (Hình 3.3) cho thấy ba chủng Đ6.1, Đ6.2 và BD8.2 đều sinh tinh thể hình cầu. Trong khi đó, tinh thể của chủng SP14.2 sinh ra có dạng hình lưỡng tháp.

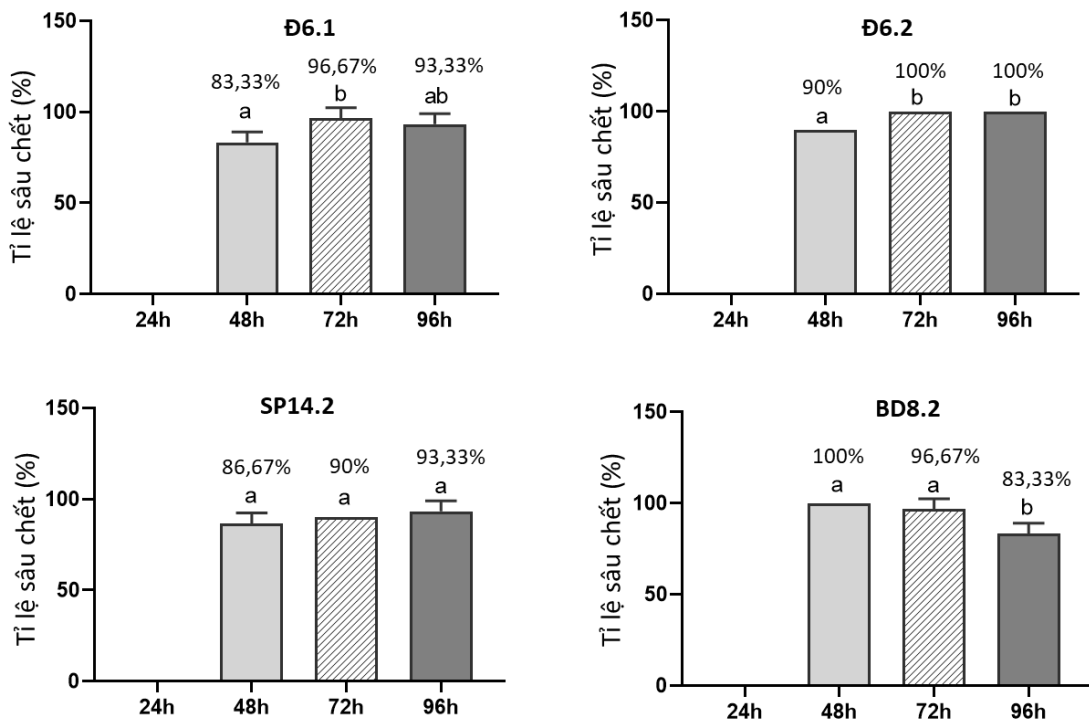


Hình 3.3. Hình thái bào tử và tinh thể của các chủng *Bt* trên kính hiển vi điện tử quét với độ phóng đại 5000x.

(a): tinh thể dạng lưỡng tháp của chủng SP14.2; (b, c, d): tinh thể hình cầu của các chủng Đ6.1, Đ6.2 và BD8.2. Hình mũi tên chỉ bào tử (s-spore) và protein tinh thể (c-crystal).

Để xác định tại pha sinh trưởng nào các chủng SP14.2, Đ6.1, Đ6.2 và BD8.2 có thể gây chết sâu thí nghiệm, từ đó đưa ra dự đoán nguyên nhân sâu chết do tác động của độc tố tinh thể *Bt* hay do các loại độc tố khác, chúng tôi tiến hành đánh giá hoạt tính diệt sâu đục quả đậu tương của các chủng *Bt* nghiên cứu tại các thời điểm 24 giờ, 48 giờ, 72 giờ và 96 giờ nuôi cấy. Kết quả sau 3 lần lặp lại thí nghiệm cho

thấy, dịch khuẩn sau 24 giờ nuôi cấy của cả 4 chủng *Bt* đều không có khả năng diệt sâu non đục quả đậu tương. Tuy nhiên, 4 chủng này sau khi nuôi cấy 48 giờ đều cho thấy hoạt tính diệt sâu ở mức cao (trên 80% -90%), đặc biệt chủng BD8.2 sau 48 giờ nuôi cấy có tỉ lệ diệt sâu đạt mức tối đa (100%). Ba chủng còn lại là Đ6.1, Đ6.2 và SP14.2, tỉ lệ sâu chết tăng dần và đều đạt cực đại (93,33% - 100%) sau 72 giờ và 96 giờ. Trong khi đó tại thời điểm 96 giờ, hoạt tính diệt sâu non đục quả đậu tương của chủng BD8.2 có sự giảm nhẹ (83,33%) (Hình 3. 4).



Hình 3.4. Đồ thị biểu diễn sự ảnh hưởng của động thái sinh trưởng đến hoạt tính diệt sâu đục quả đậu tương của 4 chủng *Bt* nghiên cứu

Động thái này hoàn toàn phù hợp với đặc điểm sinh tinh thể độc của các chủng *Bt*, chứng tỏ các chủng SP14.2, Đ6.1, Đ6.2 và BD8.2 có độc lực cao với sâu đục quả đậu tương. Theo các nghiên cứu trước đó, bốn chủng này đã được phân loại dưới loài bằng typ huyết thanh và được xác định là *B. thuringiensis canadensis* (SP14.2), *B. thuringiensis aizawai* (Đ6.1) và *B. thuringiensis kurstaki* (Đ6.2, BD8.2). Trong các nghiên cứu tiếp theo, chúng tôi sẽ tiến hành giải trình tự hệ gen của các chủng khuẩn này nhằm khai thác nguồn gen mã hóa các độc tố *Bt* tiềm năng trong việc diệt sâu đục quả đậu tương.

3.2. Phân tích, xác định các gen độc tố *Bt* mới tiềm năng từ cơ sở dữ liệu trình tự hệ gen của 4 chủng *Bt* có độc lực cao với sâu đục quả đậu tương

3.2.1. Kết quả giải trình tự hệ gen 4 chủng *Bt* bằng công nghệ SMRT Sequencing của PacBio

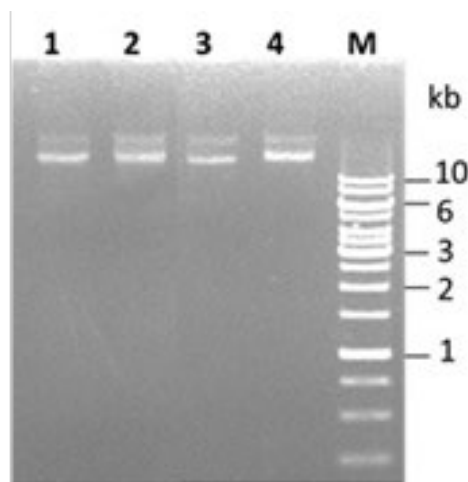
*3.2.1.1. Kết quả tách chiết DNA hệ gen có khối lượng phân tử cao từ các chủng *Bt**

Với mục đích tách chiết DNA tổng số cho giải trình tự đoạn dài yêu cầu đối với DNA thành phẩm khắt khe hơn rất nhiều so với tách chiết DNA tổng số cho những mục đích khác. Yêu cầu DNA tổng số cần phải toàn vẹn, hạn chế việc bị đứt gãy, phân mảnh và phải đảm bảo độ sạch cao. Hiện nay, công nghệ giải trình tự đoạn dài vẫn là một công nghệ mới, số thiết bị máy móc ít; vì vậy, mặc dù số nghiên cứu ứng dụng công nghệ giải trình tự thì có nhiều nhưng số nghiên cứu trực tiếp nâng cao chất lượng giải trình tự thì rất ít. Hiện nay chỉ có khoảng dưới 20 nghiên cứu và hướng dẫn dành cho việc tách chiết DNA phục vụ cho giải trình tự đoạn dài (theo thông kê của hãng Pacbio). Bên cạnh đó, cũng chưa có một bộ kit thương mại nào dành riêng cho việc tách chiết DNA tổng số dành cho mục đích giải trình tự đoạn dài.

Giống như mọi sinh vật nhân sơ, hệ gen *Bt* được tổ chức thành DNA nhiễm sắc thể xoắn kép dạng vòng và plasmid (hay còn gọi là DNA ngoài nhiễm sắc thể). Một chủng *Bt* có thể mang từ 0-12 plasmid với kích thước khác nhau từ 2 kb đến hơn 500 kb. Trong đó các gen quan trọng có tính thương mại như gen mã hóa các protein tinh thể diệt côn trùng hoặc protein sinh dưỡng diệt sâu hầu hết thường nằm trên các plasmid kích thước lớn (megaplasmid). Vì vậy, để khai thác được các gen này, việc chuẩn bị nguyên liệu DNA genome giàu plasmid là điều kiện tiên quyết. Hơn nữa, việc sử dụng phương pháp giải trình tự thế hệ thứ ba đòi hỏi DNA đầu vào phải có phân tử lượng cao và chất lượng tốt.

Trong nghiên cứu này, việc tách chiết DNA hệ gen 4 chủng *Bt* SP14.2, Đ6.1, Đ6.2 và BD8.2 được thực hiện theo Mayjonade và cs. (2018) sử dụng hạt từ đã được cacboxy hóa để thu và tinh sạch gDNA. Kết quả điện di trên gel agarose cho thấy DNA hệ gen tách chiết được từ các chủng *Bt* gồm các phân đoạn DNA phân tử

lượng cao, hạn chế tối đa sự đứt gãy, và nhiều khả năng bao gồm cả các plasmid kích thước lớn (Hình 3.5).



Hình 3.5. DNA tổng số tách chiết từ 4 chủng *Bt* nghiên cứu

1: chủng SP14.2; 2: chủng Đ6.1; 3: chủng Đ6.2; 4: chủng BD8.2; M: thang DNA 1 kb (Thermo Scientific)

Kết quả đo Nanophotometer cho thấy nồng độ DNA của các mẫu dao động trong khoảng 34 - 57 ng/μl, giá trị A260/A280 ở mức 1,78 - 1,85, giá trị A260/A230 là 1,90 - 2,01. Kết quả đo QUBIT với bộ kit Qubit™ dsDNA HS Assay Kits cho kết quả nồng độ DNA đạt từ 22 - 36 ng/μl. Với tổng thể tích mẫu thu được là 150 μl, chúng tôi có tổng số khoảng 3 - 5 μg DNA cho mỗi mẫu (Bảng 3.3). Như vậy, DNA genome thu được đã vượt qua QC, đủ số lượng và chất lượng để chuẩn bị làm thư viện.

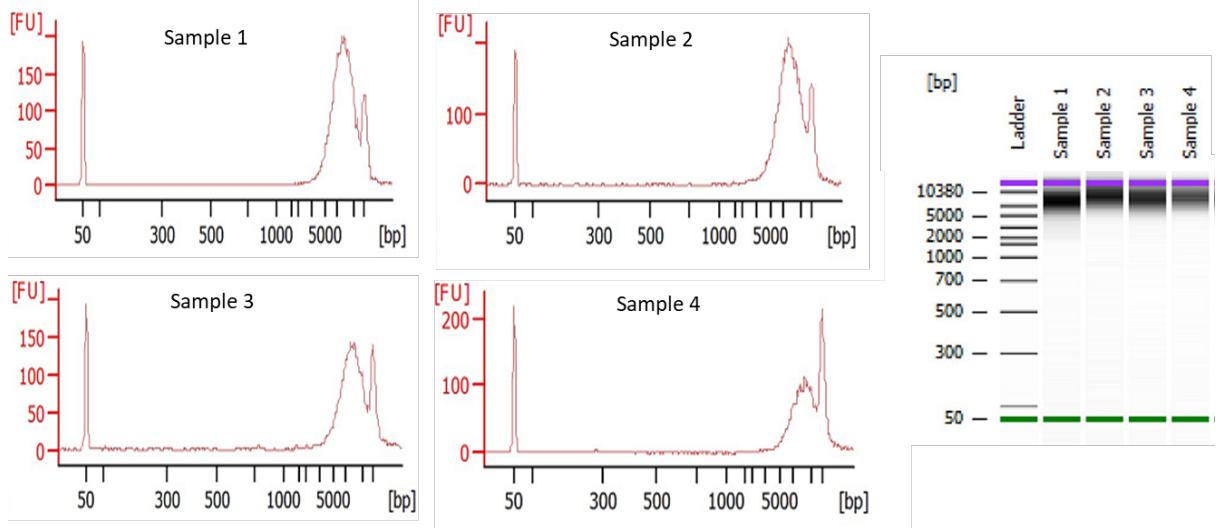
Bảng 3.3. Kết quả đo nanodrop và Qubit các mẫu DNA

STT	Tên mẫu	Nanodrop (ng/ul)	A260/280	A260/230	Qubit (ng/ul)	Qubit/Nanodrop (ratio)	Thể tích (ul)	Total (ng)
1	SP14.2	54	1,80	2,01	26	48.52	150	3900
2	Đ6.1	57	1,85	1,98	35	62.32	150	5250
3	Đ6.2	34	1,84	1,90	22	65.29	150	3300
4	BD8.2	45	1,78	2,01	36	81.11	150	5400

3.2.1.2. Kết quả tạo thư viện DNA

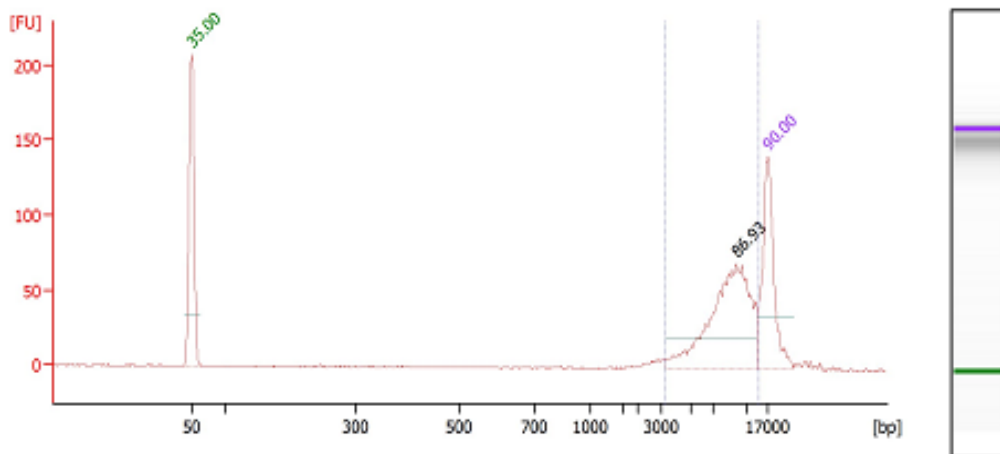
Với nồng độ DNA đo được bằng Qubit chúng tôi pha loãng DNA genome 4 chủng *Bt* với đệm EB để đạt nồng độ yêu cầu cho việc cắt DNA là khoảng 1 μg

DNA trong 150 μ l đệm sau đó chuyển DNA vào ống g-TUBE™ và ly tâm 6000 vòng trong 1 phút, đảo ngược ống và ly tâm lại với cùng tốc độ và thời gian. DNA sẽ được thu lại và chạy kiểm tra trên máy Bioanalyzer Agilent 2100. Kết quả điện di và phân tích trên bioanalyzer (Hình 3.6) cho thấy độ dài trung bình của các đoạn DNA đã cắt là khoảng 8000 bp.



Hình 3.6. Kết quả điện di và phân tích sản phẩm cắt DNA genome 4 chủng *Bt* bằng g-TUBE™ trên máy bioanalyzer

Sample 1: chủng SP14.1; Sample 2: chủng Đ6.1; Sample 3: chủng Đ6.2; Sample 4: chủng BD8.2



Hình 3.7. Kết quả phân tích thư viện DNA bằng máy bioanalyzer

Toàn bộ DNA sau cắt được đem đi làm thư viện với bộ kit SMRTbell Damage Repair Kit – SPv3. Sau khi hoàn thành sửa đuôi, gắn barcode, adapter, SMRTbell; DNA lại được kiểm tra bằng thiết bị Bioanalyzer Agilent 2100 và Qubit

một lần nữa trước khi chuyển sang bước tiếp theo. Kết quả phân tích khẳng định thao tác chuẩn bị thư viện hạn chế tối đa việc làm đứt gãy DNA vì độ dài trung bình của các đoạn DNA trước và sau khi làm thư viện là tương đương nhau (Hình 3.7).

3.2.1.3. Kết quả giải trình tự

DNA thư viện sau khi gắn polymerase sẽ được trộn với đệm cùng mẫu đối chứng nội kiểm (internal control) là một thư viện DNA chuẩn do hãng cung cấp và được nạp lên chip SMRTcell bằng phương pháp khuếch tán. Thời gian giải trình tự 1 chip là 600 phút (movie time = 600). Sau khi giải trình tự chúng tôi thu được kết quả như ở bảng 3.4.

Bảng 3.4. Kết quả giải trình tự các chủng vi khuẩn SP14.2, Đ6.1, Đ6.2 và BD8.2 bằng hệ máy PacBio SEQUEL

Chủng Bt	SP14.2	Đ6.1	Đ6.2	BD8.2
Số polymerase read	55.214	40.455	29.569	32.241
Tổng số Subread	275.957	164.431	140.734	162.421
Tổng số base	878.661.035	664.799.858	544.823.633	592.675.302
Độ dài polymerase read trung bình	31.773	26.838	26.396	26.540
Độ dài subread lớn nhất	38.537	40.694	34.171	39.658
Tỉ lệ Adapter Dimers (0-10 bp) %	0,05%	0,04%	0,05%	0,05%
Tỉ lệ đoạn đọc ngắn (11-100 bp) %	1,52%	1,36%	1,64%	1,51%
Tỉ lệ trùng khớp của mẫu nội kiểm	0,85	0,84	0,85	0,85

Dữ liệu bảng 3.4 cho thấy số lượng polymerase read là 29.569 – 55.214, số lượng subread là 140.734 – 275.957, tổng số base đọc được là trong khoảng 544.823.633 – 878.661.035 bp (544 -878 Mb); độ dài trung bình của một polymerase read là từ 26.396 – 31.773 bp; độ dài của subread dài nhất là 34.171 – 40.694 bp. Các thông số giúp đánh giá chất lượng một lần giải trình tự trên hệ máy PacBio SEQUEL bao gồm: tỉ lệ phần trăm adapter dimers (0-10 bp), phần trăm các đoạn đọc từ 11-100 bp (các đoạn DNA ngắn hoặc bị đứt gãy); tỉ lệ trùng khớp giữa trình tự thu được của mẫu nội kiểm với trình tự biết trước của chúng. Với tổng số phần trăm các đoạn đọc ngắn hơn 100 bp chiếm dưới 2% và mức độ trùng hợp

trung bình của mẫu nội kiểm trên 0,8 thì có thể khẳng định việc giải trình tự hệ gen 4 chủng *Bt* đạt kết quả tốt.

3.2.1.4. Kết quả lắp ráp *denovo* hệ gen

Theo thông kê thì độ lớn của hệ gen vi khuẩn *Bt* trung bình từ 6 đến 7 Mb, do đó giá trị độ dài genome ước đoán được lựa chọn để thiết đặt trong quá trình lắp ráp *denovo* bằng phần mềm HGAP4 là 7000000 bp (7 Mb). Số liệu kết quả của quá trình lắp ráp được thể hiện trong bảng 3.5. Theo đó, sau khi lắp ráp hoàn chỉnh, số contig thu được đối với mỗi chủng SP14.2, Đ6.1, Đ6.2 và BD8.2 theo thứ tự tương ứng là 57, 38, 61, 48; độ dài contigs lớn nhất là từ 1.320.742 - 1.802.229 bp, tổng độ dài các contig lắp ráp được ở mỗi chủng trong khoảng 6.868.840 bp – 7.025.211 bp, tỉ lệ G+C là gần 35%. Trong toàn bộ các contig không có một base N nào. Khi đánh giá chất lượng lắp ráp bằng BUSCO chúng tôi thu được kết quả độ hoàn thiện bộ gen đối với mỗi chủng *Bt* ước đoán từ 90,1 - 92,9%.

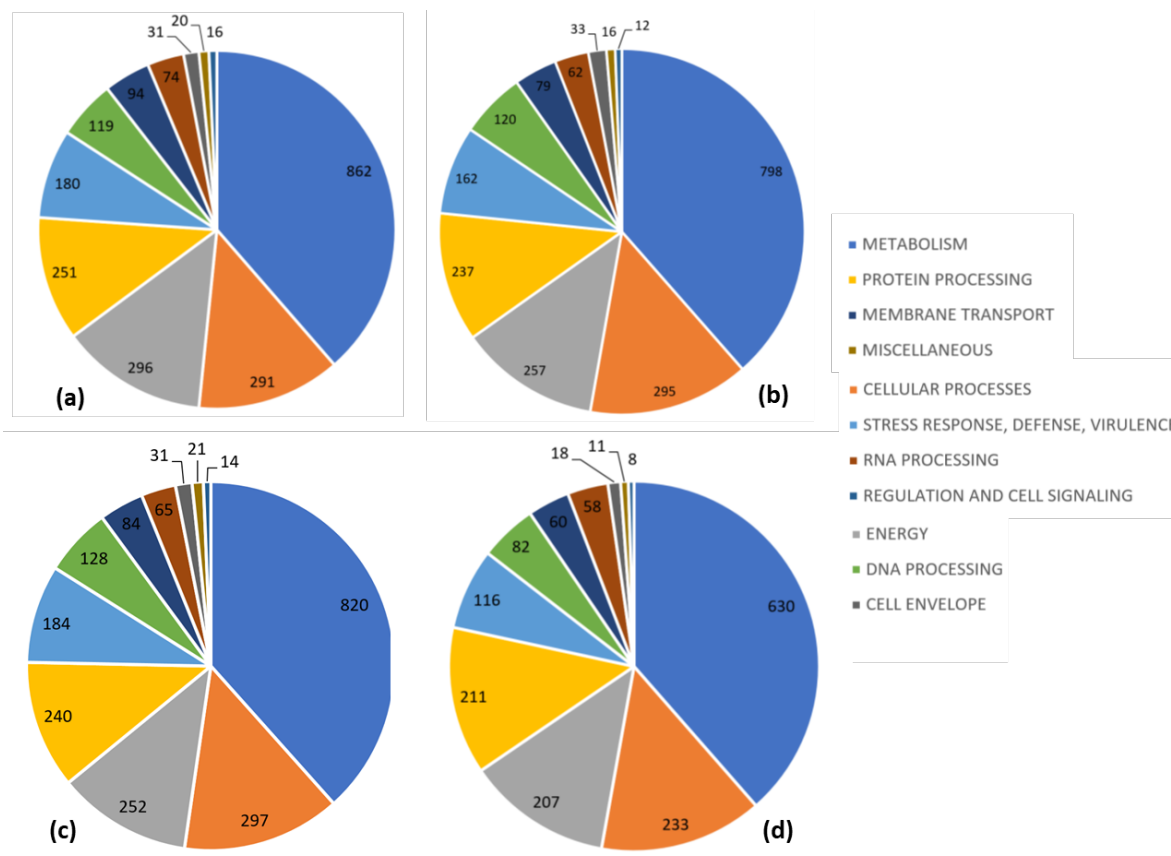
Bảng 3.5. Kết quả lắp ráp *denovo* các chủng vi khuẩn *Bt*

Chủng <i>Bt</i>	SP14.2	Đ6.1	Đ6.2	BD8.2
Số lượng contig	57	38	61	48
# contigs (≥ 0 bp)	57	38	61	48
# contigs (≥ 1000 bp)	57	38	61	48
# contigs (≥ 5000 bp)	52	36	55	43
# contigs (≥ 10000 bp)	38	31	46	39
# contigs (≥ 25000 bp)	28	19	31	25
# contigs (≥ 50000 bp)	27	15	25	23
Độ dài contig lớn nhất (bp)	1.709.205	1.802.229	1.330.436	1.320.742
Kích thước hệ gen (bp)	7.018.726	7.025.211	6.868.840	6.932.040
N50	330824	771356	512572	658180
GC (%)	34,98	34,94	34,91	34,94
# N's	0	0	0	0

3.2.1.5. Kết quả chú giải hệ gen lắp ráp *denovo* các chủng *Bt*

Dịch vụ chú giải hệ gen của PATRIC rất nhanh và toàn diện. Do sử dụng RAST toolkit [80] cùng với sức mạnh của server nên kết quả chú giải hệ gen trên PATRIC chỉ từ 15 đến 20 phút cho một hệ gen vi khuẩn. Kết quả chú giải tương đối đầy đủ: thông kê số lượng trình tự mã hóa (coding sequence), tRNA, rRNA, chú

giải chức năng các trình tự mã hóa, phân lớp chúng vào các hệ thống, các con đường chuyển hóa, các nhóm protein, sắp xếp trực quan vị trí của các gene chức năng trên từng contig, hình ảnh hóa kết quả chú giải dưới dạng genome dạng vòng, đặc biệt PATRIC hướng đến đối tượng nhà nghiên cứu quan tâm đến y sinh học vì vậy PATRIC cung cấp thêm thông tin các trình tự mã hóa có liên quan đến gây bệnh hoặc hướng thuốc.



Hình 3.8. Kết quả phân chia các trình tự mã hóa vào các hệ thống con (a) chủng SP14.2, (b) chủng Đ6.1, (c) chủng Đ6.2, (d) chủng BD8.2

Kết quả chú giải nhanh hệ gen lắp ráp *denovo* của 4 chủng SP14.2, Đ6.1, Đ6.2 và BD8.2 thu được số trình tự mã hóa theo thứ tự tương ứng là 8.329, 7.641, 7.980 và 7.781. Đồng thời, số trình tự tRNA và rRNA tương ứng là 108 và 26, 127 và 38, 112 và 28, 108 và 34. Trong số trình tự mã hóa khoảng hơn 2000 gen đã được phân bổ vào các hệ thống con đối với mỗi chủng (Hình 3.8). Trong đó, số trình tự thuộc nhóm hệ thống chuyển hóa trao đổi chất (metabolism) chiếm tỉ lệ cao

nhất, tiếp theo là hệ thống liên quan đến xử lý protein (Protein processing) và năng lượng (Energy).

Như vậy, hệ gen 4 chủng *Bt* SP14.2, Đ6.1, Đ6.2 và BD8.2 đã được giải trình tự và lắp ráp được gần đầy đủ với tỉ lệ hoàn thiện từ 90,1 - 92,9%, tổng độ dài contigs là 6.868.840 bp – 7.025.211 bp, đồng thời đã tiến hành chú giải nhanh hệ gen lắp ráp *denovo* của các chủng vi khuẩn này. Trong phân tích tiếp theo, cơ sở dữ liệu trình tự thu được sẽ được sử dụng để tìm kiếm, xác định các gen mã hóa protein độc tố diệt sâu của các chủng *Bt* nghiên cứu, từ đó dự đoán các gen mới tiềm năng trong diệt trừ sâu đục quả đậu tương *E. zinckenella* Treitschke.

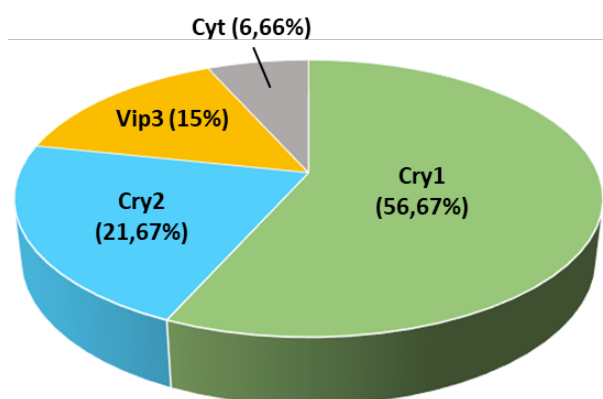
3.2.2. Kết quả phân tích các trình tự gen độc tố *Bt* từ cơ sở dữ liệu trình tự hệ gen 4 chủng *Bt* bằng công cụ *BtToxin_scanner* và xác định các gen mới tiềm năng trong diệt trừ sâu đục quả đậu tương

Với cơ sở dữ liệu hệ gen đã được lắp ráp *denovo* của chủng 4 *Bt* SP14.2, Đ6.1, Đ6.2 và BD8.2, việc tìm kiếm và tiên đoán gen mã hóa protein độc tố *Bt* được tiến hành nhờ công cụ *BtToxin_scanner*. Đây là một công cụ khai thác trên web (http://bcam.hzau.edu.cn/BtToxin_scanner/index.php) tích hợp 3 module với tính năng dự đoán là BLAST (Basic Local Alignment Search Tool), HMM (Hidden Markov Model) và SVM (Supporting Vector Machine). Cơ sở dữ liệu BLAST được xây dựng sử dụng bộ dữ liệu các gen độc tố *Bt* được duy trì và cập nhật hàng năm bởi Crickmore và cộng sự.

Kết quả phân tích dựa trên công cụ *BtToxin_scanner* cho thấy cả 4 chủng *Bt* SP14.2, Đ6.1, Đ6.2 và BD8.2 được lựa chọn đều chứa các trình tự được dự đoán mã hóa cho các protein độc tố *Bt*. Số lượng gen tìm thấy trong các chủng dao động từ 11-22 trình tự. Danh sách các gen độc tố tìm thấy ở mỗi chủng *Bt* được liệt kê cụ thể trong bảng 3.6. Về thành phần độc tố, phần lớn các trình tự được tìm thấy trong 4 chủng này mã hóa cho các protein độc tố thuộc nhóm tinh thể độc Cry (47 trình tự, chiếm 78,34%). Ngoài ra có 9 trình tự thuộc nhóm độc tố pha sinh dưỡng Vip3A (chiếm 15%) và 4 trình tự thuộc nhóm độc tố Cyt (chiếm 6,66%). Trong nhóm tinh thể độc Cry, Cry1 chiếm ưu thế với 34 trình tự, 13 trình tự còn lại thuộc phân họ Cry2 (Hình 3.9).

Bảng 3.6. Danh sách các trình tự được dự đoán mã hóa protein độc tố *Bt* khai thác được từ cơ sở dữ liệu trình tự hệ gen 4 chủng *Bt* bằng công cụ BtToxin_scanner

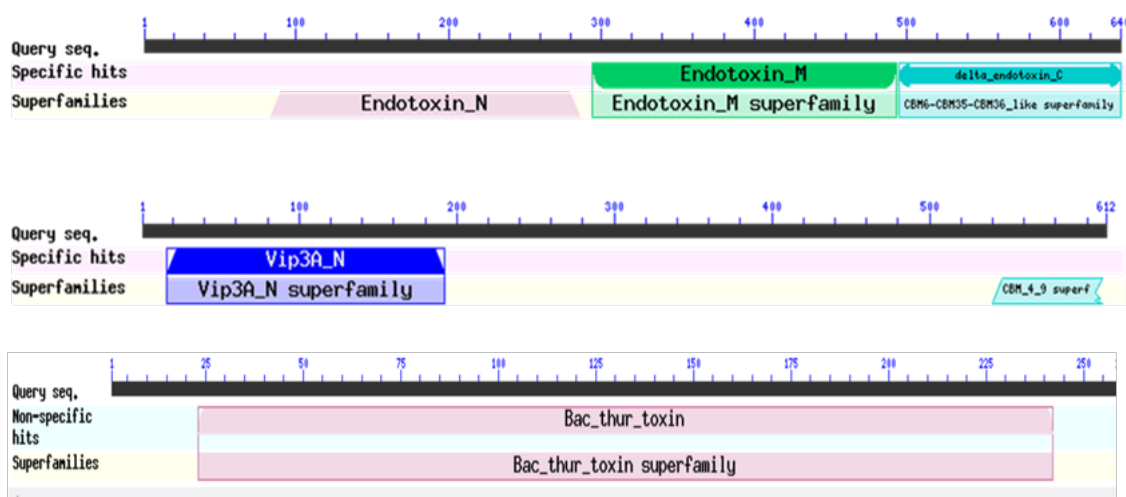
TT	Tên chủng	Các gen độc tố được dự đoán bởi BtToxin_scanner	Số lượng gen
1	SP14.2	<i>cry1Aa7, cry1Aa8, cry1Aa14, cry1Aa21, cry1Ab9, cry1Ia2, cry1Gc, cry2Aa9, cry2Ab, vip3Aa10, cyt</i>	11
2	Đ6.2	<i>cry1Aa, cry1Ab5, cry1Ab9, cry1Ab11, cry1Ac5, cry1Ac, cry1Ia2, cry2Aa9, cry2Ab3, vip3Ad5, vip3Aa21, cyt</i>	12
3	Đ6.1	<i>cry1Aa3, cry1Ah1, cry1Ah3, cry1Be1, cry1Be, cry1Ca5, cry1Ia2, cry1Ia31, cry1Ie4, cry1If1, cry1La1, cry2Aa9, cry2Ab3, cry2Ab29, cry2Ad5, cry2Ah, cry2Ah6, vip3Aa16, vip3Aa9, vip3Ab2, vip3Ad5, cyt</i>	22
4	BD8.2	<i>cry1Aa3, cry1Ah1, cry1Be1, cry1Ea7, cry1La1, cry1Ia2, cry1If1, cry1Ig1, cry1Na, cry2Aa9, cry2Ab3, cry2Ah1, vip3Aa4, vip3Ad2, cyt</i>	15
Tổng số gen			60



Hình 3.9. Thành phần và tỉ lệ các dạng protein độc tố *Bt* khai thác được từ cơ sở dữ liệu trình tự hệ gen 4 chủng *Bt* bằng công cụ BtToxin_scanner

Các trình tự protein độc tố được tiên đoán bằng công cụ BtToxin_Scanner được tiếp tục kiểm tra bằng công cụ BLAST để loại trừ các khả năng dương tính giả. BLAST (Basic Local Alignment Search Tool) là một tập hợp các chương trình tìm kiếm được thiết kế để khám phá tất cả các cơ sở dữ liệu trình tự protein và DNA có sẵn. BLASTall có rất nhiều loại tìm kiếm khác nhau phục vụ cho nhiều mục đích nghiên cứu [87]. Ở đây, chúng tôi sử dụng BLASTp để tìm kiếm tất cả các trình tự protein tương đồng với trình tự protein cần phân tích trong cơ sở dữ liệu

protein. Kết quả Blastp có thể xác định được chức năng, nguồn gốc và đặc biệt là tính mới của gen (thông qua hệ số tương đồng tối đa). Để chạy, BLASTp chúng tôi cung cấp một chuỗi amino axit đang quan tâm (chuỗi truy vấn) và so sánh với cơ sở dữ liệu của NCBI. BLASTp sẽ tìm kiếm các chuỗi con trong chuỗi truy vấn mà giống với các chuỗi con trong cơ sở dữ liệu, sau đó sẽ cho ra kết quả sau một thời gian ngắn. Trong nghiên cứu này, chúng tôi quan tâm đến mức độ tương đồng, độ bao phủ, vùng bảo tồn, vùng xúc tác của chuỗi đích so với các trình tự có sẵn trong NCBI. Kết quả phân tích sử dụng thuật toán BLASTp cho thấy tất cả các trình tự khai thác được bằng công cụ BtToxin_Scanner đều tương đồng với các protein độc tố *Bt* được công bố trên NCBI, chứng tỏ kết quả phân tích và tiên đoán của công cụ BtToxin_Scanner có độ tin cậy cao. Để xác định các vùng chức năng độc tố trên mỗi trình tự protein độc, thuật toán Blastp/Blastx được thực thi trên cơ sở dữ liệu non-redundant protein với các thông số mặc định. Trình tự ứng viên sẽ được chấp nhận dựa trên kết quả BLAST và sự có mặt của các vùng chức năng độc tố được xác thực bởi CDD (Conserve Domain Database) của NCBI. Theo đó, chúng tôi đã nhận diện được sự có mặt của 1 hoặc cả 3 vùng chức năng đặc trưng Endotoxin_N (Pfam PF03945), Endotoxin_M (Pfam PF00555) và Endotoxin_C (Pfam PF03944) ở các trình tự thuộc nhóm tinh thể độc Cry. Trong khi đó, các trình tự thuộc nhóm Vip3A đều tìm thấy vùng chức năng độc tố Vip3A_N (PF12495). Tương tự, vùng chức năng đặc trưng cho nhóm độc tố Cyt là Bac_thur_toxin (Pfam PF01338) cũng được xác định có mặt ở 4 trình tự được tiên đoán thuộc nhóm độc tố này (Hình 3.10).



Hình 3.10. Các vùng chức năng đặc trưng cho các nhóm protein độc tố *Bt* được phát hiện trên các trình tự ứng viên khai thác được từ cơ sở dữ liệu hệ gen của 4 chủng *Bt*

*** Kết quả xác định, lựa chọn các gen mới có tiềm năng bằng tin sinh học**

Về mặt lý thuyết, để xác suất phân lập thành công các gen mã hóa độc tố có tính mới từ cơ sở dữ liệu hệ gen 4 chủng *Bt* nghiên cứu là cao nhất, chúng tôi ưu tiên các trình tự mà thành phần axit amin suy diễn có sự tương đồng < 100% so với các độc tố *Bt* đã được công bố. Mặt khác, nhằm ứng dụng trong việc diệt trừ sâu đục quả đậu tương, các trình tự ứng viên cũng được đối chiếu với dữ liệu nghiên cứu liên quan đến các gen mã hóa cho protein gây độc cho côn trùng bộ cánh vảy (thiên về các loại sâu đục thân, cành, bắp, rế) hại cây trồng. Căn cứ vào các tiêu chí trên, 6 trình tự gen độc tố được dự đoán có tính mới đã được lựa chọn bao gồm 4 gen thuộc nhóm *cry1* (bao gồm *cry1Aa*, *cry1Ac*, *cry1Be*, *cry1Na*) và 2 gen thuộc phân nhóm *cry2A* là *cry2Ab* và *cry2Ah*. Trình tự axit amin suy diễn từ các gen này có độ tương đồng nằm trong khoảng 94,01% - 99,26% với các độc tố *Bt* đã công bố. Riêng *cry1Na*, mặc dù xét về độ tương đồng, protein do gen tiên đoán mã hóa giống 100% với độc tố Cry1Na3, tuy nhiên xét về độ bao phủ, trình tự axit amin của protein tiên đoán hơn protein tham chiếu 39 axit amin ở phía đầu N, do vậy tính mới của protein tiên đoán vẫn cần được xem xét (Bảng 3.7). Dựa trên kết quả các nghiên cứu trước đây, chúng tôi thấy rằng nhiều độc tố *Bt* thuộc 2 nhóm Cry1 và Cry2A đã được chứng minh có độc lực cao với côn trùng thuộc bộ Cánh vảy, đặc biệt là các loài sâu đục thân hại lúa *Scirpophaga incertulas* [88], sâu đục thân hại mía *Sesamia inferens* [89] và sâu đục đa thực *Helicoverpa armigera* [90], [91]. Như vậy, có thể thấy 6 độc tố được lựa chọn hoàn toàn phù hợp với định hướng nghiên cứu đề ra. Các nghiên cứu tiếp theo được thực hiện nhằm khai thác và đánh giá khả năng ứng dụng của 6 độc tố *Bt* này trong diệt trừ sâu hại đậu tương *E. zinckenella* Treitschke.

Bảng 3.7. Danh sách các protein độc tố tiềm năng được dự đoán có tính mới và độ tương đồng với các protein diệt côn trùng *Bt* đã công bố

STT	Protein độc tố diệt côn trùng được tiên đoán	Chủng <i>Bt</i> mang gen mã hóa	Kích thước protein tiên đoán (amino axit)	Protein độc tố tương đồng trên Genbank	Kích thước protein tham chiếu (amino axit)	Mã Genbank của protein/gen mã hóa protein tham chiếu	Độ tương đồng (%)
1	Cry1Aa	Đ6.2	1077	Cry1Aa8	1176	I26149.1	99,26
2	Cry1Ac	Đ6.2	496	Cry1Ac30	1178	ACT22519.1	95,93
3	Cry1Be	Đ6.1	894	Cry1Be3	1227	ACV96720.1	98,99
4	Cry1Na	BD8.2	640	Cry1Na3	601	AKQ08661.1	100 (hơn kém 39 amino axit)
5	Cry2Ab	SP14.2	230	Cry2Ab3	633	AAG36762.1	94,01
6	Cry2Ah	Đ6.1	573	Cry2Ah1	632	ACL13555.1	98,42

*** Kết quả đánh giá mức độ an toàn của 6 protein độc tố *Bt* tiềm năng đối với con người**

Ngoài việc xem xét về tiềm năng ứng dụng của 6 protein độc tố *Bt* quan tâm, việc phân tích và khẳng định tính an toàn của các protein này đối với con người là rất quan trọng. Nghiên cứu này được thực hiện dựa trên công cụ web AllergenOnline (<http://www.allergenonline.org/>). AllergenOnline cung cấp quyền truy cập vào danh sách các chất gây dị ứng đã được thẩm định và cơ sở dữ liệu có thể tìm kiếm theo trình tự nhằm xác định các protein tiềm ẩn nguy cơ phản ứng chéo gây dị ứng. Trang web này được thiết kế để giúp đánh giá sự an toàn của các protein có thể được đưa vào thực phẩm thông qua kỹ thuật di truyền hoặc thông qua các phương pháp chế biến thực phẩm. Mục tiêu là để xác định các protein cần phải có thêm các xét nghiệm bổ sung, chẳng hạn như liên kết IgE huyết thanh, sự giải phóng histamin của bạch cầu ưa kiềm hoặc các thử nghiệm *in vivo* để đánh giá khả năng phản ứng chéo. Cơ sở dữ liệu được cập nhật hàng năm, có thể được truy cập tự do với mục đích cung cấp một công cụ đơn giản và hữu ích trong việc đánh giá an toàn thực phẩm. Cơ sở dữ liệu này hiện tại có 2035 trình tự và 808 nhóm protein gây dị ứng và có nguy cơ gây dị ứng. Việc so sánh và tìm kiếm trình tự protein gây

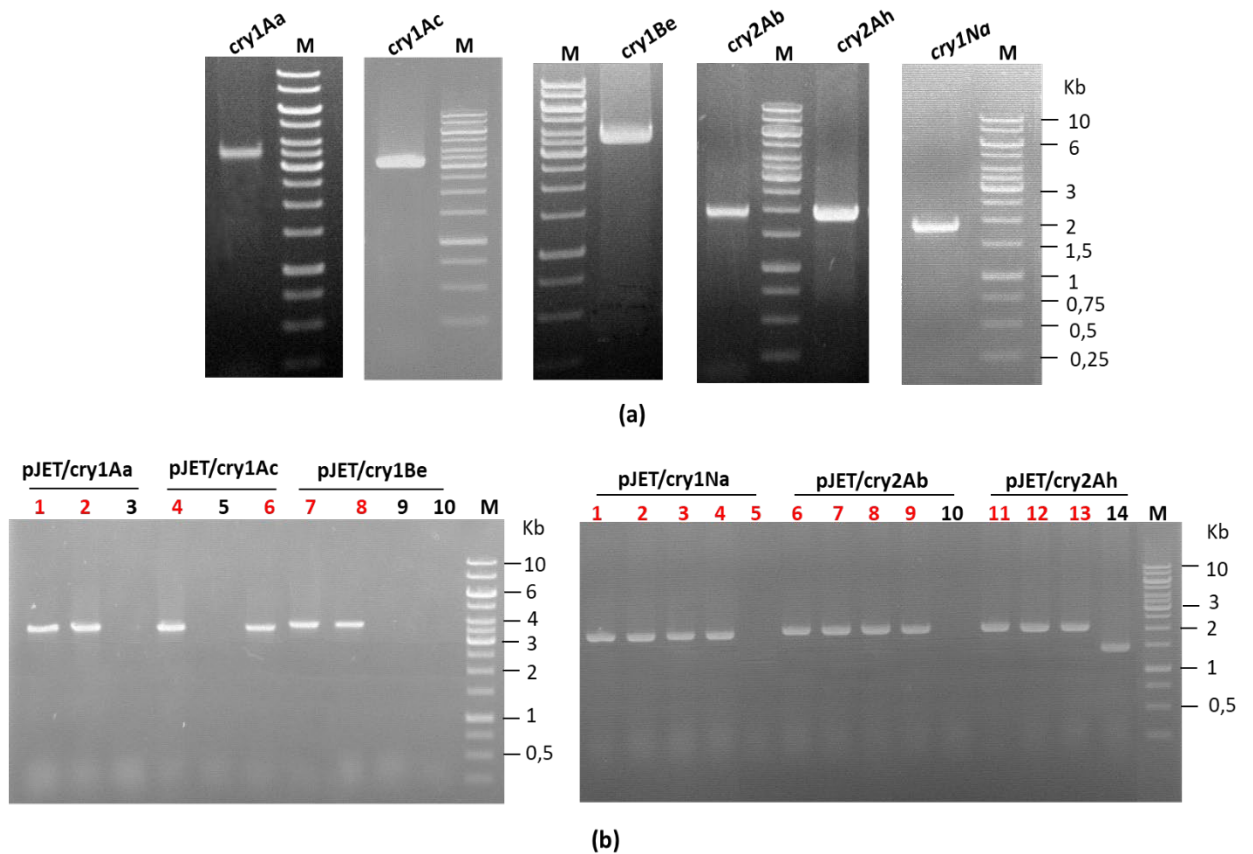
độc trên cơ sở dữ liệu này sử dụng ma trận điểm BLOSUM50 trên toàn bộ chiều dài trình tự bằng công cụ FASTA3. Các trình tự có độ tương đồng trên 70% sẽ được coi là có nguy cơ gây dị ứng.

Kết quả tìm kiếm và so sánh cho thấy 5 trình tự truy vấn gồm Cry1Aa, Cry1Ac, Cry1Be, Cry2Ab và Cry2Ah không có sự tương đồng với các trình tự tham chiếu trên ngân hàng dữ liệu các chất có khả năng gây dị ứng Allergenonline. Đối với protein Cry1Na, kết quả trích xuất cho giá trị bao phủ và tương đồng rất thấp (dưới 30%) (Hình phụ lục 2). Điều này cho thấy các protein độc tố được lựa chọn là an toàn và không có nguy cơ gây dị ứng cho con người.

3.3. Kết quả phân lập, tách dòng và xác định trình tự các gen *cry* mã hóa protein độc tố mới có tiềm năng diệt sâu đục quả đậu tương từ các chủng *Bt* lựa chọn.

3.3.1. Kết quả phân lập, tách dòng các gen độc tố *cry*

Từ bảng 3.7, có thể thấy ngoại trừ trình tự gen *cry1Na* có độ bao phủ hoàn toàn gen tham chiếu *cry1Na3*, tất cả các gen còn lại là *cry1Aa*, *cry1Ac*, *cry1Be*, *cry2Ab* và *cry2Ah* đều là các gen chưa được giải mã và lắp ráp hoàn chỉnh do không có mã mở đầu và mã kết thúc, đồng thời kích thước nhỏ hơn các gen tham chiếu là *cry1Aa8*, *cry1Ac30*, *cry1Be3*, *cry2Ab3* và *cry2Ah1*, theo thứ tự tương ứng. Trong trường hợp này, để đảm bảo phân lập được đầy đủ trình tự các gen độc tố *cry* tiên đoán, việc thiết kế mồi sử dụng cho phân lập gen được dựa trên trình tự phần đầu và cuối của các gen tham chiếu tương ứng. Riêng đối với gen *cry1Na*, cặp mồi được thiết kế cho phân lập sẽ khuếch đại hoàn toàn trình tự tiên đoán khai thác được từ dữ liệu DNA hệ gen chủng *Bt* BD8.2 (Bảng 2.1). Để đảm bảo độ chính xác của phản ứng PCR, Phusion DNA polymerase với hoạt tính đọc sửa 3'-5' exonuclease được sử dụng giúp hạn chế tối đa các sai sót trong quá trình sao chép. Kết quả điện di sản phẩm của phản ứng PCR (Hình 3.11a) cho thấy đã khuếch đại thành công 6 đoạn DNA có chiều dài phù hợp với kích thước theo dự đoán của các gen độc tố *Bt* quan tâm (hơn 3,5 kb đối với các gen *cry1Aa*, *cry1Ac*, *cry1Be* và gần 2 kb đối với các gen *cry1Na*, *cry2Ab*, *cry2Ah*). Điều này chứng minh các cặp mồi được thiết kế và chu trình nhiệt của phản ứng PCR sử dụng để phân lập gen là hiệu quả.



Hình 3.11. Kết quả điện di trên gel agarose kiểm tra các bước phân lập và nhân dòng các gen mã hóa protein độc tố *Bt*

(a): sản phẩm PCR khuếch đại các gen *cry1Aa*, *cry1Ac*, *cry1Be*, *cry1Na*, *cry2Ab* và *cry2Ah* từ DNA hệ gen 4 chủng *Bt*. (b): kết quả PCR sàng lọc các dòng *E. coli* mang plasmid tái tổ hợp pJET/*cry1Aa*, pJET/*cry1Ac*, pJET/*cry1Be*, pJET/*cry1Na*, pJET/*cry2Ab* và pJET/*cry2Ah*. M: thang DNA 1 Kb (Thermo Scientific).

Sản phẩm PCR khuếch đại các gen độc tố sau đó được tinh sạch, nối ghép vào vector tách dòng pJET1.2/blunt tạo thành các plasmid tái tổ hợp pJET/*cry1Aa*, pJET/*cry1Ac*, pJET/*cry1Be*, pJET/*cry1Na*, pJET/*cry2Ab*, pJET/*cry2Ah* và biến nạp vào tế bào *E. coli* DH5 α để dòng hóa gen. Để lựa chọn được các thể biến nạp mang các vector mong muốn, các dòng tế bào *E. coli* được nuôi cấy, tách chiết plasmid và kiểm tra sự chèn vào của gen đích bằng phản ứng PCR sử dụng cặp mồi pJET1.2_F/R được thiết kế gắn trên hai vùng biên của vector pJET1.2/blunt (Bảng 2.1). Theo kết quả phân tích sản phẩm PCR trên gel agarose 0,8% (Hình 3.11b), chúng tôi đã chọn được các dòng dương tính (màu đỏ) đối với mỗi cấu trúc plasmid tái tổ hợp pJET/*cry1Aa*, pJET/*cry1Ac*, pJET/*cry1Be*, pJET/*cry1Na*, pJET/*cry2Ab*

và pJET/cry2Ah. Các plasmid này tiếp tục được sử dụng làm nguyên liệu để giải trình tự gen bằng phương pháp Sanger sử dụng cặp mồi pJET1.2_F/R.

3.3.2. Kết quả giải và so sánh trình tự của các gen *cry* đã phân lập với các gen độc tố *Bt* đã công bố.

Kết quả giải trình tự theo phương pháp Sanger và phân tích bằng phần mềm BioEdit cũng như so sánh trên công cụ BLAST khẳng định việc phân lập thành công 6 gen mã hóa protein độc tố là *cryIAa*, *cryIAC*, *cryIBe*, *cryINa*, *cry2Ab* và *cry2Ah* từ 4 chủng *Bt* được lựa chọn từ bộ sưu tập *Bt* của Viện Công nghệ Sinh học. Tuy nhiên ngoài gen *cryINa* có trình tự phân lập được hoàn toàn khớp với trình tự gen tiên đoán, trình tự thực tế phân lập được của các gen còn lại đều có nhiều điểm sai khác so với trình tự gen tiên đoán khai thác từ cơ sở dữ liệu hệ gen *Bt* (Bảng 3.8). Điều này hoàn toàn có thể được lý giải bởi quá trình giải trình tự và lắp ráp hệ gen có tỉ lệ sai sót lớn hơn so với giải trình tự đoạn DNA ngắn theo phương pháp Sanger. Ngoài ra, các trường hợp sử dụng gen tiên đoán ở dạng gen cụt (được giải mã và lắp ráp chưa hoàn chỉnh) có độ chính xác về trình tự không cao cũng là nguyên nhân chính dẫn đến những sai khác này.

Bảng 3.8. Sự sai khác về trình tự nucleotide giữa các gen độc tố được phân lập với trình tự gen tiên đoán khai thác từ dữ liệu giải trình tự hệ gen.

STT	Gen độc tố	Vị trí nucleotide sai khác giữa gen <i>cry</i> phân lập và gen tiên đoán từ dữ liệu hệ gen <i>Bt</i>	Sự sai khác nucleotide
1	<i>cryIAa</i>	342	Thêm T
2	<i>cryIAC</i>	444	G → T
		616	T → C
		648	A → G
		684	G → A
		738	G → A
		741	C → T
		936-937	Thêm TT
3	<i>cryIBe</i>	1041	Thêm G

		1962	A → G
4	<i>cry1Na</i>	Ko	Ko
5	<i>cry2Ab</i>	248	Thêm C
		527	Thêm T
		531	Thêm T
		561	Thêm T
		695	G → A
		708	C → T
		762	A → G
		798	C → T
		813	T → C
		816	C → T
6	<i>cry2Ah</i>	214	Thêm A

Do có sự sai khác về trình tự nucleotide giữa các gen mã hóa protein độc tố *Bt* phân lập được với các trình tự gen tiên đoán ban đầu, chúng tôi đã sử dụng thuật toán BLASTp để kiểm định tính mới của các protein độc tố do các gen này mã hóa. Kết quả được trình bày trong bảng 3.9 cho thấy sự thay đổi về tính mới của một số gen thực tế phân lập được so với tiên đoán ban đầu. Cụ thể, hai gen *cry1Aa* và *cry1Ac* có nguồn gốc từ chủng *Bt* Đ6.2 ban đầu được tiên đoán có tính mới được xếp hạng 4 (Rank 4) do trình tự axit amin tương đồng 99,26% và 95,93% (nằm trong khoảng > 95%) với gen *cry1Aa8* và *cry1Ac30*, theo thứ tự tương ứng. Tuy nhiên, kết quả phân lập thực tế đã chứng minh 2 gen này chính là *cry1Aa8* và *cry1Ac30* do có trình tự nucleotide giống 100% với gen tham chiếu, vì vậy không có tính mới. Gen *cry2Ab* ban đầu được dự đoán có tính mới xếp hạng 3 do trình tự protein mã hóa tương đồng 94,01% (nằm trong khoảng > 78% và < 95%) với độc tố Cry2Ab3, tuy nhiên trình tự gen *cry2Ab* thực tế phân lập được có độ tương đồng cao hơn dự đoán (99,21%, nằm trong khoảng > 95%), nên độ mới chỉ được xếp hạng 4. Ngoài ra, so với tiên đoán ban đầu, mức độ tương đồng của các protein độc tố mã hóa bởi các gen *cry1Be* và *cry2Ah* thực tế phân lập được so với các protein tham chiếu Cry1Be3 và Cry2Ah1 cũng có sự tăng nhẹ (từ 98,99% lên 99,56% và từ

98,42% lên 99,53%), tuy nhiên tính mới của hai gen này vẫn giữ nguyên ở hạng 4. Trong số 6 gen *cry* đã phân lập, chỉ duy nhất gen *cry1Na* không có sự sai khác giữa trình tự thực tế và trình tự tiên đoán ban đầu. Mặc dù có độ tương đồng 100% với trình tự gen *cry1Na3*, nhưng đầu 5' của *cry1Na* có thêm 1 đoạn gồm 117 nucleotide phía đầu 5' mã hóa cho 39 axit amin đầu N của chuỗi polypeptide. Vì vậy, tính mới của độc tố Cry1Na cần được cân nhắc và xem xét.

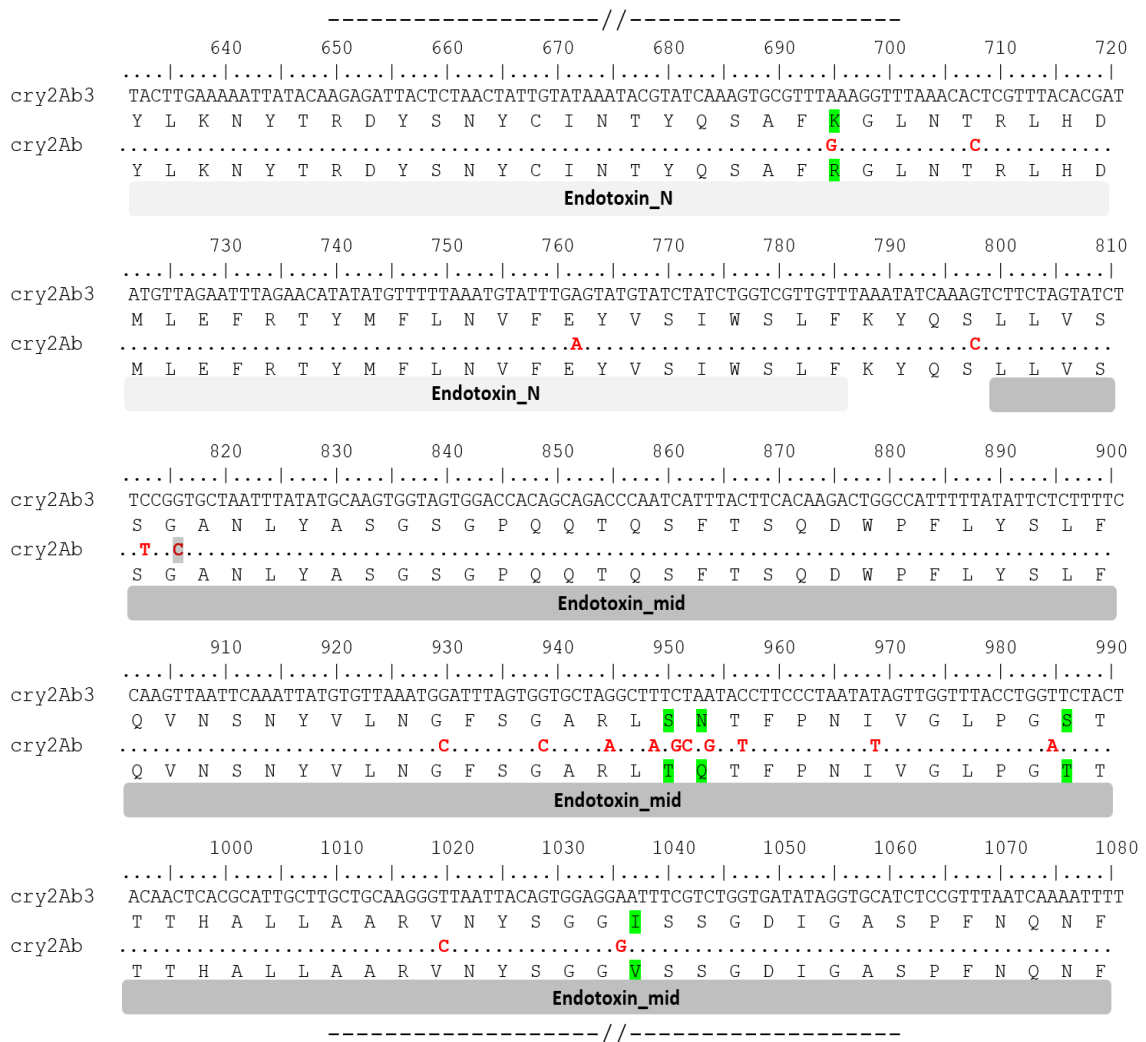
Bảng 3.9. Mức độ tương đồng của 6 protein độc tố đã phân lập được gen mã hóa với các protein đã công bố trên Genbank

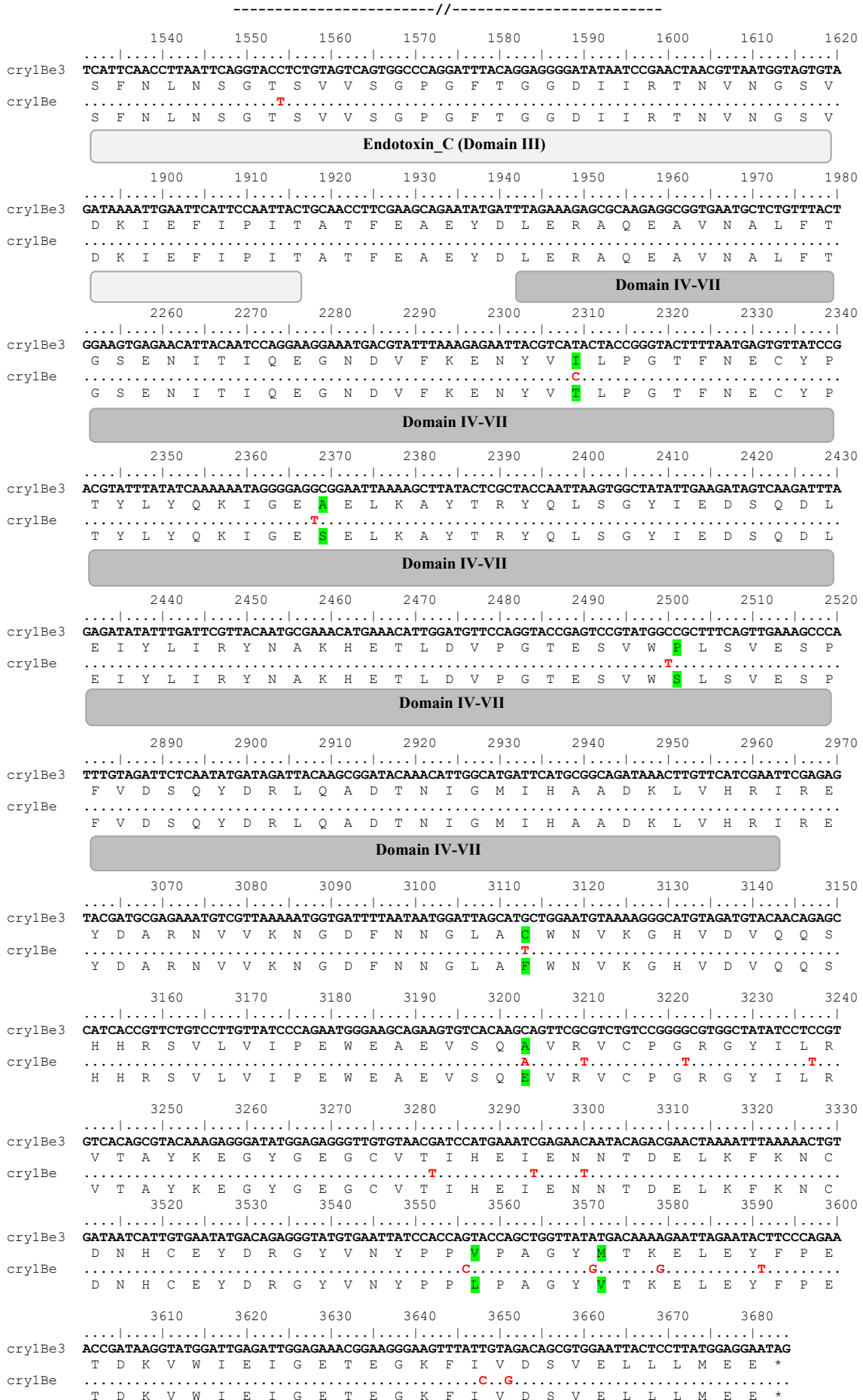
TT	Độc tố <i>Bt</i> phân lập	Kích thước gen mã hóa phân lập được (bp)	Protein tham chiếu tương đồng trên Genbank	Mức độ tương đồng dự đoán (%)	Mức độ tương đồng thực tế (%)	Xếp hạng về tính mới
1	Cry1Aa	3531	Cry1Aa8	99,26	100	Không
2	Cry1Ac	3537	Cry1Ac30	95,93	100	Không
3	Cry1Na	1890	Cry1Na3	100	100 (hơn protein tham chiếu 39 amino axit đầu N)	Cần xem xét
4	Cry1Be	3681	Cry1Be3	98,99	99,56	Hạng 4
5	Cry2Ab	1902	Cry2Ab3	94,01	99,21	Hạng 4
6	Cry2Ah	1899	Cry2Ah1	98,42	99,53	Hạng 4

Sự sai khác về trình tự nucleotide và axit amin giữa các gen mã hóa protein độc tố có tính mới là *cry1Na*, *cry1Be*, *cry2Ab* và *cry2Ah* với các trình tự gen tham chiếu được liệt kê cụ thể tại hình 3.12 và bảng 3.10.

Gen *cry2Ab* phân lập từ chủng SP14.2 gồm 1902 nucleotide, mã hóa cho chuỗi polypeptide gồm 633 axit amin với khối lượng phân tử ước tính là 70 kDa. Trình tự nucleotide và axit amin suy diễn của gen này có độ tương đồng 99,05% và 99,21%, theo thứ tự tương ứng, với gen *cry2Ab3* trên GenBank với mã truy cập AF164666.1 (Hình phụ lục 3). Giống như các thành viên khác thuộc họ độc tố Cry có cấu trúc ba vùng chức năng, cấu trúc protein Cry2Ab cũng gồm ba vùng bảo tồn điển hình là *endotoxin_N* (pfam03945), *endotoxin_mid* (pfam09131) và *endotoxin_C* (pfam03944). Kết quả giống hàng trình tự của gen *cry2Ab* phân lập

được với gen tham chiếu *cry2Ab3* cho thấy 18 vị trí sai khác về nucleotide. Tuy nhiên, chỉ có 7 vị trí trong số đó dẫn đến sự thay đổi về 5 axit amin, cụ thể là ở các codon 232 (K232R), 317 (S317T), 318 (N318Q), 329 (S329T) và 346 (I346V) (Hình 3.12, Bảng 3.10). Ngoài trừ vị trí sai khác axit amin ở codon 232 nằm trong vùng I (endotoxin_N) chịu trách nhiệm chèn qua màng và tạo lỗ trên màng tế bào ruột giữa của côn trùng đích, tất cả các vị trí sai khác còn lại đều nằm trong vùng II (endotoxin_mid), là vùng liên quan đến nhận biết và gắn kết với thụ thể, quyết định tính đặc hiệu côn trùng của độc tố *Bt*.





-----//-----

190 200 210 220 230 240 250 260 270
cry2Ah1 CTTTAAAGAAAGTGGGGAGTCTTGTGAAAAAAGAATATTAAGTGAAGTTACGGAATTAATATTCCTAGTGGCAGTACAAATCTAATG
cry2Ah L L K K V G S L V **G** K R I L S E L R N L I F P S G S T N L M
L L K K V G S L V **G** K R I L S E L R N L I F P S G S T N L M

Endotoxin N

730 740 750 760 770 780 790 800 810
cry2Ah1 ATGTTAGAATTTAGAACATATATGTTTTAAATGATTTGAATATGTATCTATCTGGTCGTTGTTTAAATATCAAGCCTTCTAGTATCT
cry2Ah M L E F R T Y M F L N V F E Y V S I W S L F K Y Q S L L V S
M L E F R T Y M F L N V F E Y V S I W S L F K Y Q S L L V S

Endotoxin_N

820 830 840 850 860 870 880 890 900
cry2Ah1 TCTGGCGTAATTTATATGCAAGTGGTAGTGGACCACAGCAGACCCCAATCATTACTTCACAAGACTGGCCATTTTATATCTCTTTTC
cry2Ah S G A N L Y A S G S G P Q Q T Q S F T S Q D W P F L Y S L F
S G A N L Y A S G S G P Q Q T Q S F T S Q D W P F L Y S L F

Endotoxin_mid

1360 1370 1380 1390 1400 1410 1420 1430 1440
cry2Ah1 AGTCCCTCAGGAACCTGGTGGATTACGAGCTTATATGGTATCTGTGCATAACAGAAAAATAATATCTATGCCGTTTCATGAAATGGT
cry2Ah S P S G T P G G L R A Y M V S V H N R K N N I Y A V H E N G
S P S G T P G G L R A Y M V S V H N R K N N I Y A V H E N G

Endotoxin_mid

1450 1460 1470 1480 1490 1500 1510 1520 1530
cry2Ah1 ACTATGATTCATTAGCGCCGGAAGATTATACAGGATTTACTATCGCCGATACATGCAACTCAAGTGAATAATCAACCGCAACATT
cry2Ah T M I H L A P E D Y T G F T I S P I H A T Q V N N N Q T R T F
T M I H L A P E D Y T G F T I S P I H A T Q V N N N Q T R T F

Endotoxin_C

1810 1820 1830 1840 1850 1860 1870 1880 1890
cry2Ah1 AATGTACCGTTAGATATAAATGTGACATAAATTCGGGTACTCAATTTGAGCTTATGAATATATGTTTGTCCCACTAATCTCCACCA
cry2Ah N V P L D I N V T L N S G T Q F E L M N I M F V P T N **A** **T** **S** **P**
N V P L D I N V T L N S G T Q F E L M N I M F V P T N **A** **T** **S** **P**

Endotoxin_C

.....
cry2Ah1 CTTTATTA
L Y *
cry2Ah
L Y *

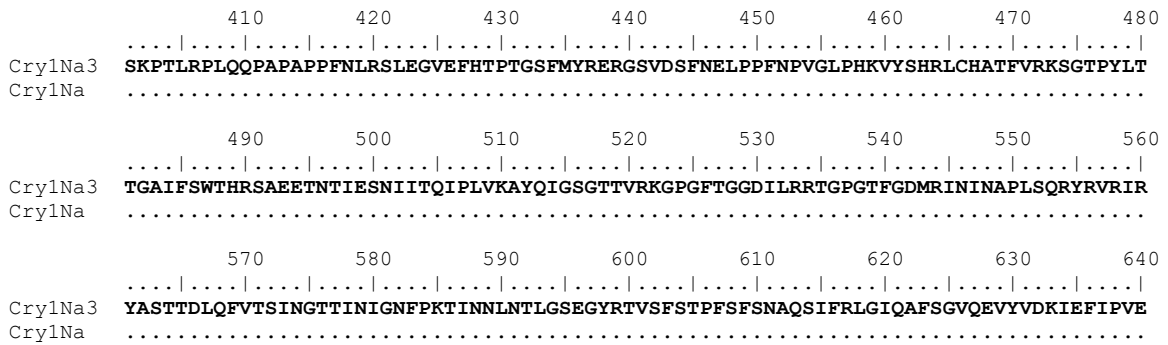
10 20 30 40 50 60 70 80
Cry1Na3 MNSKEHDYLRKVCNDLSDANINMERFDKNDALEIGMSIVSEL
Cry1Na SLILKFLFLNVMSELKGNFKKSTNRTCCLLKIINIGGRG.....

90 100 110 120 130 140 150 160
Cry1Na3 IGMIPGGTALQFVFNQLWSRLGDSGWNAFMEHVEELIDTKIEGYAKNKALSELAGIQRNLETYYIQLRNEWENDIENSKAQ
Cry1Na

170 180 190 200 210 220 230 240
Cry1Na3 GKVANYYESLEQAVERSMPQFAVGNFEVPLLTVYVQAANLHLLLRDVSVYGKRWGWSEQKIKIYYDKQIKYTHEYTNHC
Cry1Na

250 260 270 280 290 300 310 320
Cry1Na3 VNWNKGLERLKNKGSSYQDWYNNRFRREMTLTVLDIVALFPHYDVQTYPIITVAQLTREVYTDPLLNFPKLSVSQL
Cry1Na

330 340 350 360 370 380 390 400
Cry1Na3 PSFSDMENATIRTPHLMFLRMLTIYTDWYVGRNYVGGHRVTSYHVGGENIRSPLYGREANQEVPRDFYFYGPVFKTL
Cry1Na



Hình 3.12. Kết quả giống hàng thể hiện sự sai khác về trình tự nucleotide và axit amin giữa gen độc tố *cry2Ab*, *cry1Be*, *cry2Ah*, *cry1Na* và các gen tham chiếu (Chữ màu đỏ và chữ được bôi xanh thể hiện các vị trí sai khác về nucleotide và axit amin, theo thứ tự tương ứng. Phần trình tự giống nhau giữa hai gen được rút ngắn và biểu diễn bởi các đường đứt nét)

Gen *cry1Be* phân lập từ chủng Đ6.1 gồm 3681 nucleotide, mã hóa cho chuỗi polypeptide gồm 1226 axit amin với khối lượng phân tử ước tính là 136 kDa. Trình tự nucleotide và axit amin suy diễn của gen này có độ tương đồng 99,48% và 99,35%, theo thứ tự tương ứng, với gen *cry1Be1* trên GenBank với mã truy cập AF077326.1 (Hình phụ lục 4). Về cấu trúc, ngoài 3 vùng chức năng chính tạo thành vùng lõi độc tố điển hình cho các δ -endotoxin sản sinh bởi *Bt*, protein Cry1Be còn có vùng đầu C mở rộng gồm 4 domain (domain IV-VII). Kết quả giống hàng trình tự của gen *cry1Be* phân lập được với gen tham chiếu *cry1Be1* cho thấy 18 vị trí sai khác về nucleotide. Tuy nhiên, trong số này chỉ có 7 vị trí sai khác dẫn đến sự thay đổi về axit amin, cụ thể là ở các codon 770 (I770T), 790 (A790S), 834 (P834S), 1038 (C1038F), 1068 (A1068E), 1186 (V1186L) và 1191 (M1191V). Đặc biệt, toàn bộ các vị trí axit amin sai khác này đều nằm trong vùng đầu C mở rộng (Hình 3.12, Bảng 3.10). Vùng này được cho là có chức năng chính trong việc hình thành cấu trúc thể vùi dạng tinh thể của protein tiền độc tố (protoxin) do thành phần axit amin rất giàu cysteine giúp tạo các cầu nối disulphide liên phân tử.

Gen *cry2Ah* phân lập từ chủng Đ6.1 gồm 1899 nucleotide, mã hóa cho chuỗi polypeptide gồm 632 axit amin với khối lượng phân tử ước tính là 70 kDa. Trình tự nucleotide và axit amin suy diễn của gen này có độ tương đồng 99,84% và 99,53%, theo thứ tự tương ứng, với gen *cry2Ah1* trên GenBank với mã truy cập EU939453.1

(Hình phụ lục 5). Tương tự với độc tố thuộc nhóm Cry2, protein Cry2Ah cũng chỉ có cấu trúc ba vùng chức năng điển hình là endotoxin_N, endotoxin_mid và endotoxin_C. Kết quả giống hàng trình tự của gen *cry2Ah* phân lập được với gen tham chiếu *cry2Ah1* cho thấy 3 vị trí sai khác về nucleotide, đồng thời dẫn đến thay đổi về thành phần axit amin tại 3 vị trí codon 70 (E70G), 628 (L628I) và 629 (P629S). Trong đó, codon 70 nằm trong vùng I (endotoxin_N), 2 codon còn lại nằm trong vùng III (endotoxin_C) (Hình 3.12, Bảng 3.10). Cả 2 vùng này đều có chức năng liên quan đến việc chèn vào màng và tạo lỗ trên màng tế bào ruột giữa của côn trùng đích.

Gen cuối cùng là *cry1Na* phân lập từ chủng BD8.2. Gen này gồm 1890 nucleotide mã hóa cho chuỗi polypeptide gồm 629 axit amin với khối lượng phân tử ước tính là 70 kDa. Protein Cry1Na cũng chỉ có cấu trúc ba vùng chức năng điển hình là endotoxin_N, endotoxin_mid và endotoxin_C. Kết quả giống hàng trình tự của gen *cry1Na* phân lập được với gen tham chiếu *cry1Na3* (mã Genbank KT273482.1) cho thấy sự tương đồng 100%. Tuy nhiên, gen *cry1Na* có thêm 1 đoạn gồm 117 nucleotide phía đầu 5' mã hóa cho 39 axit amin đầu N của chuỗi polypeptide (Hình 3.12, Hình phụ lục 6).

Bảng 3.10. Vị trí các nucleotide và axit amin sai khác của các gen độc tố phân lập được so với gen tham chiếu

Gen độc tố phân lập	Gen độc tố tham chiếu	Vị trí nucleotide sai khác	Sự sai khác nucleotide giữa gen phân lập và gen tham chiếu	Sự sai khác amino axit
<i>cry2Ab</i>	<i>cry2Ab3</i>	695	A → G	K → R
		708	T → C	-
		762	G → A	-
		798	T → C	-
		813	C → T	-
		816	T → C	-
		930	A → C	-
		939	T → C	-
		945	G → A	-
		949	T → A	S → T
951	T → G			

		952	A → C	N → Q
		954	T → G	
		957	C → T	-
		969	A → T	-
		985	T → A	S → T
		1020	T → C	-
		1036	A → G	I → V
cry1Be	cry1Be3	1554	T → C	-
		2309	C → T	I → T
		2368	T → G	A → S
		2500	T → C	P → S
		2874	C → T	-
		3113	T → G	C → F
		3203	A → C	A → E
		3210	T → C	-
		3222	T → G	-
		3237	T → C	-
		3282	T → G	-
		3294	T → C	-
		3300	T → C	-
		3556	C → G	V → L
		3571	G → A	M → V
		3579	G → A	-
		3591	T → C	-
3648	C → T	-		
3651	G → A	-		
cry2Ah	cry2Ah1	209	A → G	E → G
		1882	C → A	L → I
		1885	C → T	P → S
Cry1Na	Cry1Na3	1-117	Gen tham chiếu không có	SLILKLFLL NVMSELK GNFKKSTN RTCCLKII NIGGRG

Như vậy, từ kết quả khai thác dữ liệu genome 4 chủng *Bt* Đ6.1, Đ6.2, SP14.2, BD8.2 và phân lập 6 gen mới tiềm năng, chúng tôi đã thu được 4 gen mã hóa cho các protein độc tố diệt côn trùng có tính mới cấp độ 4 và không có nguy cơ gây dị ứng là *cry1Be*, *cry1Na*, *cry2Ab* và *cry2Ah*. Đáng chú ý, phần lớn các thành viên thuộc nhóm gen *cry1* và *cry2* đã được nhiều nghiên cứu trước đây chứng minh là mã hóa cho các protein gây độc cho côn trùng bộ cánh vảy, trong đó có các loài

sâu đục thân, sâu đục quả... Điều này cho thấy 4 gen độc tố *Bt* đã phân lập rất có tiềm năng và phù hợp với mục tiêu nghiên cứu của đề tài. Trong bước tiếp theo, 4 gen này sẽ được đưa vào hệ biểu hiện của vi khuẩn *E. coli* nhằm thu được protein tái tổ hợp để đánh giá khả năng diệt sâu đục quả hại đậu tương *Etiella zinckenella* Treitschke.

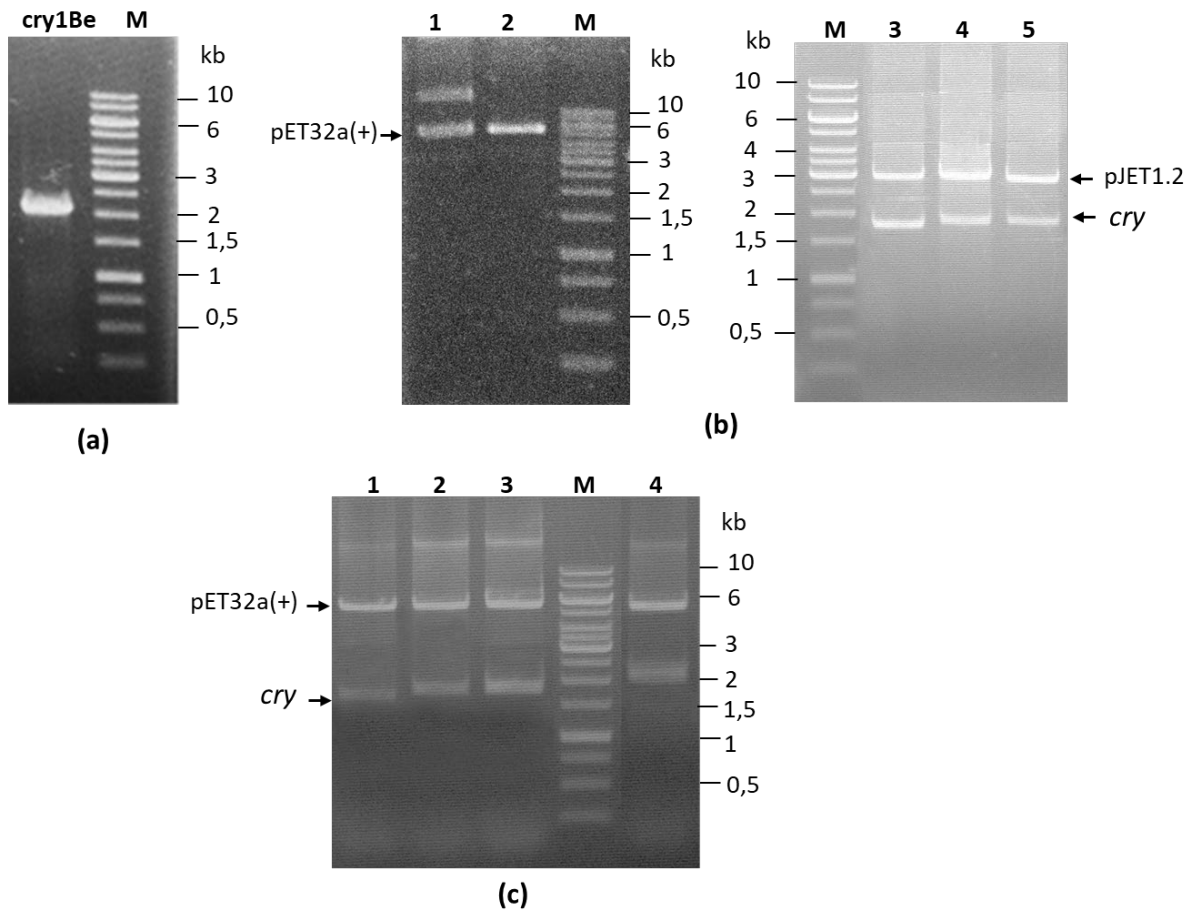
3.4. Kết quả thiết kế vector, biểu hiện các gen độc tố *cry* của các chủng *Bt* trong tế bào *E. coli* và đánh giá hoạt tính diệt sâu đục quả đậu tương của protein tái tổ hợp

3.4.1. Thiết kế các vector biểu hiện pET32a(+) mang gen *cry1Na*, *cry2Ab*, *cry2Ah* và *cry1Be*

Vector pJET1.2/blunt chỉ được sử dụng cho mục đích tách dòng, hoàn toàn không có khả năng biểu hiện gen ngoại lai. Do đó, chúng tôi tiến hành cắt và ghép nối các gen *cry1Na*, *cry2Ab*, và *cry2Ah* đã được dòng hóa trong vector pJET1.2/blunt sang vector biểu hiện pET32a(+) - một vector có nhiều ưu điểm thuận lợi cho việc biểu hiện gen ngoại lai do có mang T7lac promoter là một loại promoter mạnh được cảm ứng bởi IPTG. Đối với *cry1Be*, đoạn gen phân lập dài 3681 bp, mã hóa cho một tiền độc tố có khối lượng phân tử lớn do có chứa cả các trình tự thuộc vùng mở rộng đầu C không có vai trò trong quá trình gây độc của protein cho côn trùng đích (vùng mở rộng đầu C này sẽ bị cắt bỏ bởi các protease khi tiền độc tố vào ruột giữa của côn trùng để tạo thành độc tố dạng hoạt hóa). Mặt khác, kích thước gen lớn cũng gây khó khăn cho quá trình thiết kế vector cũng như việc biểu hiện gen. Vì vậy, trong trường hợp này, chúng tôi lựa chọn biểu hiện một phân đoạn gen *cry1Be* mã hóa protein độc tố với kích thước 2063 bp bao gồm 3 domain đầu tiên nằm trong vùng lõi độc, chịu trách nhiệm gây độc cho côn trùng đích.

Phân đoạn gen *cry1Be* được khuếch đại nhờ phản ứng PCR sử dụng khuôn là plasmid pJET/cry1Be và cặp mồi đặc hiệu có gắn thêm vị trí nhận biết của *Bam*HI và *Xho*I (Bảng 2.1) giúp xử lý cắt tạo đầu so le cho đoạn DNA thu được. (Hình 3.13a). Đối với 3 gen còn lại là *cry1Na*, *cry2Ab*, và *cry2Ah*, do trong quá trình phân lập gen, khi thiết kế mồi (Bảng 2.1), chúng tôi đã đưa trình tự nhận biết của hai enzyme hạn chế *Bam*HI và *Xho*I vào hai đầu các gen nên vấn đề trở nên đơn giản khi sử dụng chính cặp enzyme này để cắt và thu các đoạn gen đích (Hình

3.13b). Cùng với đó, để tiến hành nối ghép các gen quan tâm vào vector biểu hiện pET32a(+) thì bản thân vector cũng phải được cắt mở vòng bằng chính hai enzyme trên để tạo đầu dính bổ sung (Hình 3.13b). Sau khi các enzyme cắt hoàn toàn, tiến hành thu, tinh sạch các đoạn gen đích và vector pET32a(+) mở vòng để thực hiện phản ứng nối ghép nhờ xúc tác của T4-DNA ligase tạo thành các vector tái tổ hợp pET32/cry1Be, pET32/cry1Na, pET32/cry2Ab, pET32/cry2Ah và được biến nạp vào chủng vi khuẩn biểu hiện *E. coli* BL21 bằng phương pháp sốc nhiệt.



Hình 3.13. Kết quả điện di trên gel agarose kiểm tra các bước thiết kế các vector biểu hiện pET32a(+) mang gen *cry1Na*, *cry2Ab*, *cry2Ah* và *cry1Be*

(a) Sản phẩm PCR khuếch đại phân đoạn 2063 bp của gen *cry1Be*

(b) Kết quả xử lý các vector bằng cặp enzyme *Bam*HI và *Xho*I để tạo nguyên liệu cho phản ứng nối ghép. 1: plasmid pET32a(+); 2-5: pET32a(+) và pJET/cry1Na, pJET/cry2Ab, pJET/cry2Ah được cắt bằng *Bam*HI và *Xho*I, theo thứ tự tương ứng.

(c) Cắt kiểm tra các plasmid tái tổ hợp pET32/cry1Na, pET32/cry2Ab, pET32/cry2Ah và pET32/cry1Be bằng *Bam*HI và *Xho*I (1-4, theo thứ tự tương ứng). M: Thang DNA 1 Kb (Thermo Scientific).

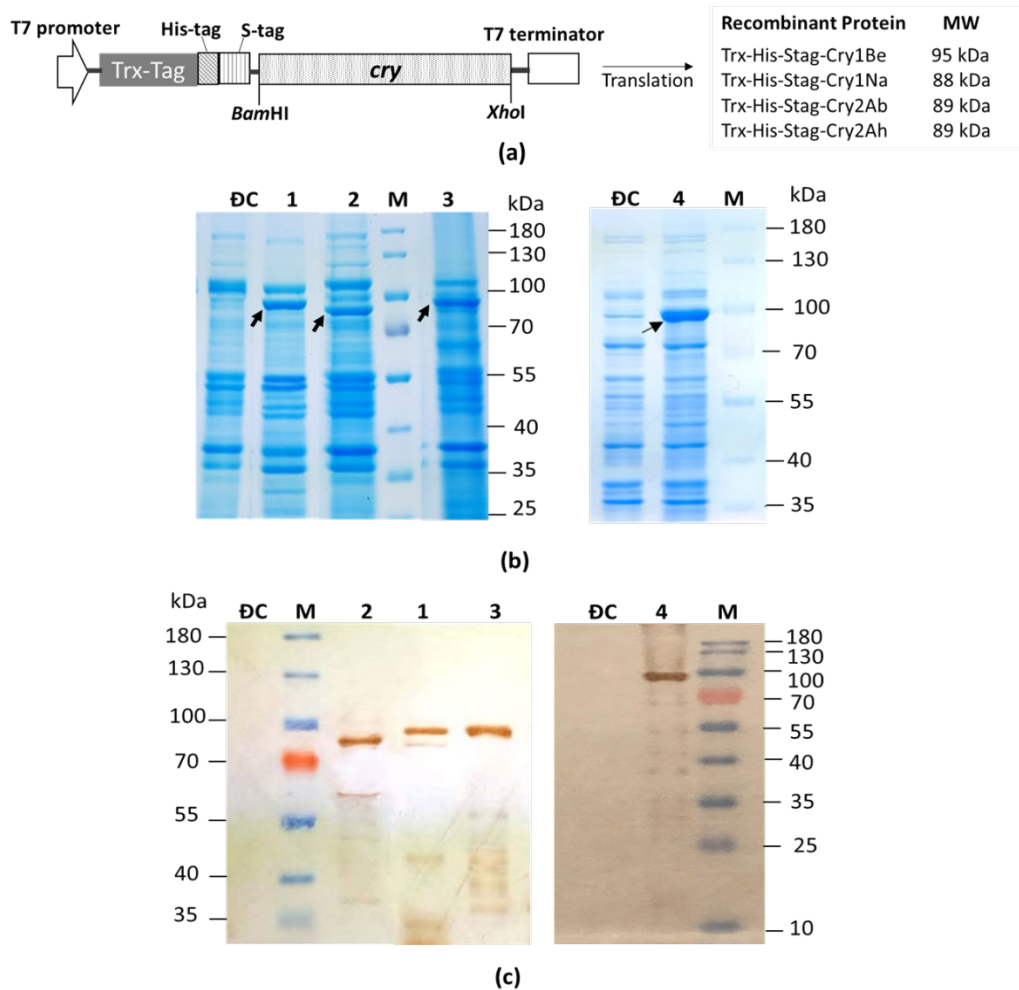
Để xác định được các dòng tế bào mang các plasmid tái tổ hợp mong muốn, chúng tôi tiến hành tách chiết plasmid và cắt kiểm tra bằng cặp enzyme hạn chế *Bam*HI và *Xho*I. Về lý thuyết, nếu quá trình nối ghép gen thành công thì các vector thành phẩm khi được xử lý bằng cặp enzyme hạn chế này sẽ tạo ra 2 đoạn ADN bao gồm: đoạn lớn dài gần 6 Kb (khung vector pET32a(+)) và đoạn nhỏ có kích thước đúng bằng các gen đích (khoảng 2 Kb). Kết quả điện di kiểm tra sản phẩm các phản ứng cắt trên gel agarose (Hình 3.13c) cho thấy đã chọn được các dòng tế bào *E. coli* BL21 mang plasmid cho các băng đúng như tính toán ứng với từng cấu trúc pET32/cry1Be, pET32/cry1Na, pET32/cry2Ab và pET32/cry2Ah. Các vector tái tổ hợp này sau đó tiếp tục được giải trình tự để khẳng định chắc chắn quá trình thiết kế vector không xảy ra các sai sót ảnh hưởng đến khung đọc của gen đích. Các dòng tế bào *E. coli* BL21 mang vector biểu hiện đã được kiểm chứng chặt chẽ được sử dụng để nghiên cứu biểu hiện các độc tố *Bt* quan tâm.

3.4.2. Biểu hiện các gen *cry1Be*, *cry1Na*, *cry2Ab*, và *cry2Ah* trong tế bào *E. coli* BL21

Dựa trên đặc điểm của vector biểu hiện pET32a(+), với cách thiết kế các plasmid tái tổ hợp nêu trên, các gen *cry* sẽ được biểu hiện ở dạng protein dung hợp. Cụ thể, đầu 5' của gen *cry* được lai với *trx* mã hóa cho thioredoxin cùng với hai đoạn trình tự mã hóa cho S.tag và hexahistidine (6xHis), kết quả biểu hiện là protein tái tổ hợp Trx-His-Stag-Cry (Hình 3.14a). Việc biểu hiện ở dạng dung hợp với thioredoxin đã được chứng minh giúp tăng khả năng hình thành cấu trúc bậc bốn chính xác cho protein ngoại lai, đồng thời hạn chế sự thủy phân bởi các enzyme protease trong tế bào chủ. Ngoài ra, nhờ dung hợp với hexahistidine, protein tái tổ hợp có thể được tinh chế bằng cột sắc ký dựa vào ái lực giữa 6xHis và Ni²⁺ trên cột.

Chủng vi khuẩn *E. coli* BL21 được sử dụng làm chủng biểu hiện gen. Đây là chủng chứa đột biến *lon* protease (protease nội bào) và *ompT* protease (protease màng ngoài tế bào). Vì vậy, chủng có ưu điểm là hạn chế được sự phân cắt của protease đối với protein ngoại lai, giúp cho protein ngoại lai ổn định hơn trong tế

bào sau khi được tổng hợp. Để biểu hiện gen, chủng tái tổ hợp phải được nuôi cấy trong môi trường có chất cảm ứng promoter. Các gen *cry* trong vector tái tổ hợp pET32/*cry* hoạt động dưới sự điều khiển của promoter T7. Promoter này được kiểm soát chặt chẽ bởi chất cảm ứng IPTG. Vì vậy, để biểu hiện các gen độc tố *cry*, tế bào *E. coli* BL21 mang plasmid tái tổ hợp được nuôi cấy trong môi trường chọn lọc đến khi OD₆₀₀ đạt 0,6 - 0,8 thì bổ sung thêm chất cảm ứng IPTG đến nồng độ cuối cùng là 0,1 mM và nuôi tiếp trong khoảng 4 giờ. Protein tổng số của dịch chiết tế bào được phân tích trên gel polyacrylamide (Hình 3.14b).



Hình 3.14. Kết quả biểu hiện các gen *cry1Be*, *cry1Na*, *cry2Ab39* và *cry2Ah* trong tế bào *E. coli* BL21

- (a) Mô hình cassette biểu hiện protein tái tổ hợp dạng dung hợp Trx-His-Stag-Cry.
 (b), (c) Phân tích SDS-PAGE và Western blot kiểm tra protein tổng số các chủng *E. coli* biểu hiện các protein dung hợp Trx-His-Stag-Cry2Ab (1), Trx-His-Stag-

Cry1Na (2), Trx-His-Stag-Cry2Ah (3) và Trx-His-Stag-Cry1Be (4). ĐC: protein tổng số của tế bào *E. coli* BL21. M: thang protein 10–180 kDa (Thermo Scientific).

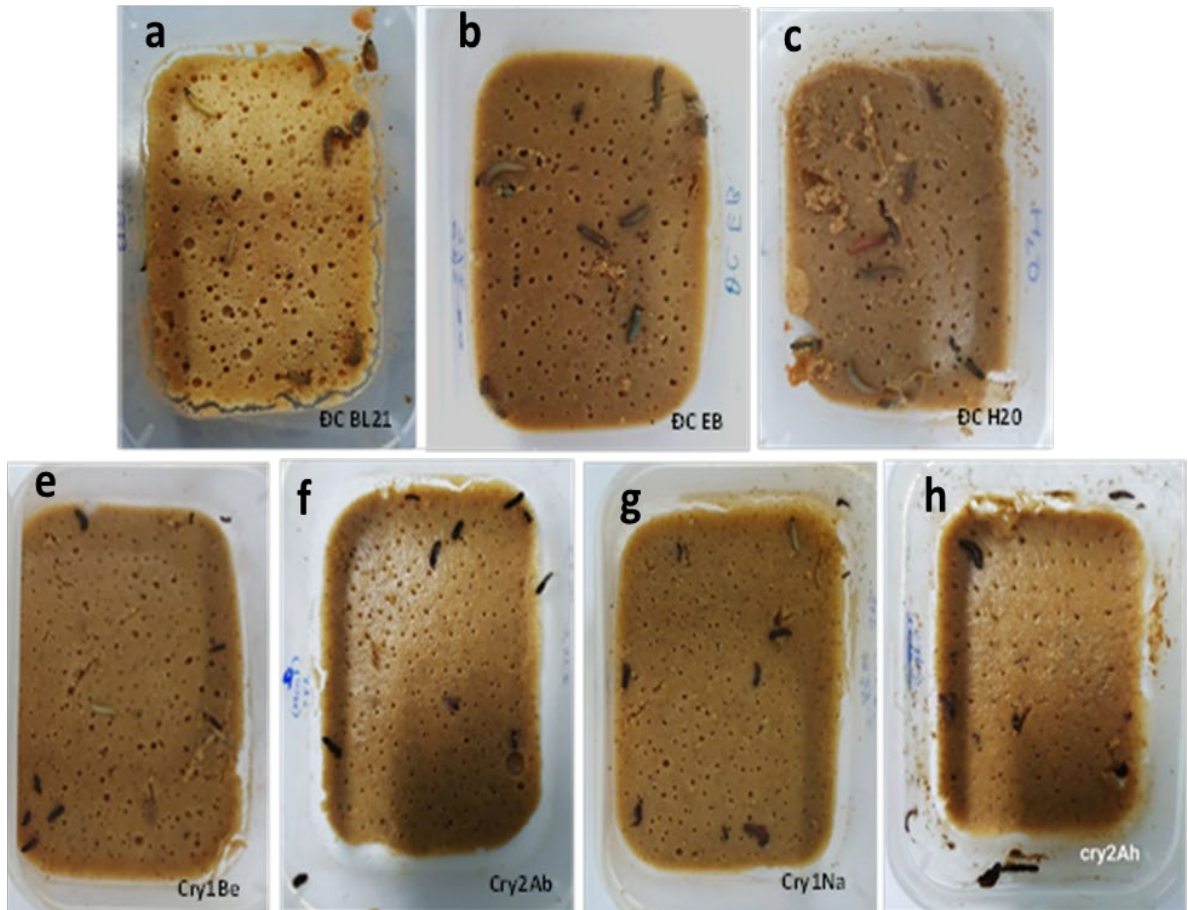
Kết quả trên điện di đồ cho thấy, trong điều kiện được cảm ứng bởi IPTG, protein tổng số thu được từ các chủng *E. coli* BL21 mang các vector pET32/cry1Be, pET32/cry1Na, pET32/cry2Ab và pET32/cry2Ah chứa các băng protein đậm nét có kích thước tương đương với các protein dung hợp Trx-His-Stag-Cry1Be (95 kDa), Trx-His-Stag-Cry1Na (88 kDa), Trx-His-Stag-Cry2Ab và Trx-His-Stag-Cry2Ah (89 kDa). Trong khi đó, ở mẫu đối chứng là tế bào *E. coli* BL21 thì không thấy sự xuất hiện các băng protein này. Việc biểu hiện thành công các protein tái tổ hợp mong muốn còn được khẳng định thông qua phép lai miễn dịch Western blot với kháng thể kháng hexahistidine (6xHis) (Hình 3.14c). Các kết quả thu được, chúng tôi đi đến kết luận đã biểu hiện thành công các gen *cry1Be*, *cry1Na*, *cry2Ab*, và *cry2Ah* trong tế bào *E. coli* BL21.

3.4.3. Thử hoạt tính diệt sâu đục quả đậu tương của các protein tái tổ hợp thu được.

Để kiểm tra khả năng diệt sâu đục quả đậu tương *E. zinckenella* Treitschke, các chủng *E. coli* BL21 biểu hiện các protein độc tố tái tổ hợp Cry1Be, Cry1Na, Cry2Ab và Cry2Ah được nuôi cấy ở điều kiện cảm ứng và thu sinh khối của 150 ml dịch khuẩn sau 4 giờ. Tế bào *E. coli* sau đó được hòa trong 10 ml dung dịch EB và siêu âm phá tế bào để thu dịch protein tổng số có chứa các protein mong muốn. Thí nghiệm thử hoạt tính diệt sâu đục quả đậu tương *E. zinckenella* Treitschke được thực hiện tương tự như đã làm để sàng lọc các chủng *Bt* trong bộ sưu tập *Bt* bản địa Việt Nam của Viện Công nghệ Sinh học. Cụ thể, bổ sung 2 ml dịch protein tổng số vào cốc chứa 20 g môi trường thức ăn nhân tạo. Mỗi cốc được thả 10 cá thể sâu non tuổi 2-3 để thử nghiệm. Song song với các công thức thí nghiệm, thực hiện tương tự với các công thức đối chứng sử dụng dịch protein tổng số của vi khuẩn *E. coli* BL21, dung dịch đệm EB và nước deion. Thí nghiệm được lặp lại 3 lần.

Căn cứ vào tỉ lệ sâu chết sau 7 ngày theo dõi (Bảng 3.11, Hình 3.15) có thể thấy protein Cry1Be có hoạt lực diệt sâu yếu nhất với tỉ lệ sâu chết đen là 66,67%. Hai protein Cry1Na và Cry2Ah có độc lực mạnh hơn với ấu trùng sâu đục quả đậu, với tỉ lệ diệt sâu là 86,67% và 80%. Đáng chú ý, tỉ lệ sâu chết ở công thức sử dụng

Cry2Ab là cao nhất (96,67%), chứng tỏ protein Cry2Ab có khả năng diệt sâu đục quả đậu tốt nhất trong số 4 loại độc tố được thử nghiệm. Protein này được lựa chọn để thực hiện các phân tích tiếp theo nhằm xác định liều gây chết trung bình LC_{50} và các nghiên cứu tiếp theo.



Hình 3.15. Kết quả thử hoạt tính diệt sâu đục quả đậu tương *E. zinckenella*

Treitschke của các protein tái tổ hợp

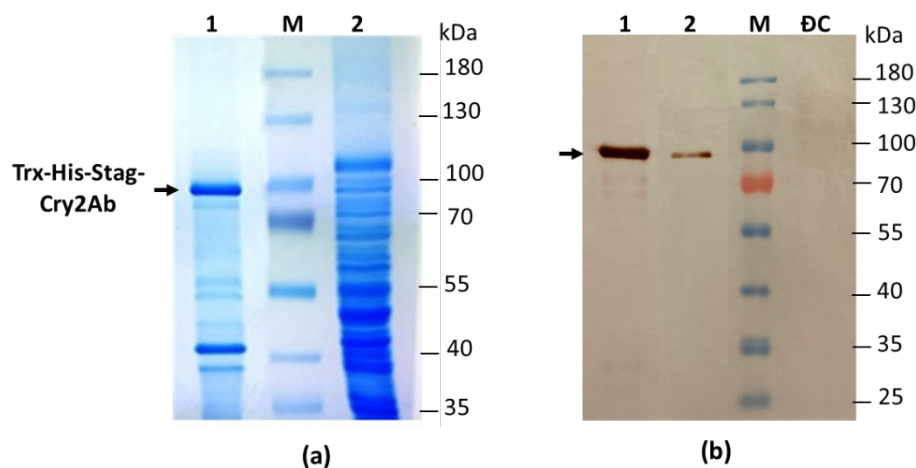
(a-h: Môi trường thức ăn nhân tạo bổ sung dịch protein tổng số chủng *E. coli* BL21, đệm EB, nước deion, dịch protein tổng số các chủng *E. coli* BL21 biểu hiện các protein độc tố tái tổ hợp Cry1Be, Cry2Ab, Cry1Na và Cry2Ah, theo thứ tự tương ứng)

Bảng 3.11. Kết quả thử hoạt tính diệt sâu đục quả *E. zinckenella* Treitschke của 4 protein độc tái tổ hợp

STT	Mẫu thử/Protein tái tổ hợp	Tổng sâu bắt vào	Tổng sâu sống + chết trắng	Tổng sâu chết đen	Hoạt lực diệt sâu (%)
1	Đối chứng (Protein tổng số của <i>E. coli</i> BL21/ đệm EB/ nước deion)	10	10	0	0
		10	10	0	
		10	10	0	
2	Cry1Be	10	4	6	66,67
		10	3	7	
		10	3	7	
3	Cry1Na	10	1	9	86,67
		10	1	9	
		10	2	8	
4	Cry2Ah	10	3	7	80
		10	2	8	
		10	1	9	
5	Cry2Ab	10	1	9	96,67
		10	0	10	
		10	0	10	

3.4.4. Tối ưu điều kiện biểu hiện gen *cry2Ab* trong tế bào *E. coli* BL21 nhằm tăng lượng protein tái tổ hợp trong pha tan.

Một trong những vấn đề thường gặp khi biểu hiện gen trong tế bào *E. coli* là protein tái tổ hợp biểu hiện ở pha cặn, gây khó khăn cho khâu tinh chế và thu hồi protein mong muốn ở dạng có hoạt tính tốt nhất. Để kiểm tra khả năng tan của protein Trx-His-Stag-Cry2Ab, chúng tôi tiến hành siêu âm phá tế bào, ly tâm tách pha protein hòa tan và tua để chạy điện di SDS PAGE và phân tích Western blot (Hình 3.16). Kết quả cho thấy lượng protein dung hợp Trx-His-Stag-Cry2Ab39 có trong pha tan là rất ít mà tồn tại chủ yếu ở dạng thể vùi không tan nằm trong tế bào chất. Đây là một trong những vấn đề thường gặp khi protein ngoại lai được biểu hiện ở mức độ cao trong tế bào *E. coli*. Các protein ở trạng thái không tan thường không đảm bảo được hoạt tính sinh học và đòi hỏi quá trình tái tạo cấu trúc (refolding) tương đối phức tạp. Do vậy, các điều kiện biểu hiện gen đã được tối ưu hóa để tăng lượng protein đích được tổng hợp ở pha tan, giúp đơn giản hóa cho bước tinh chế cũng như thử hoạt tính của protein tái tổ hợp thu được.



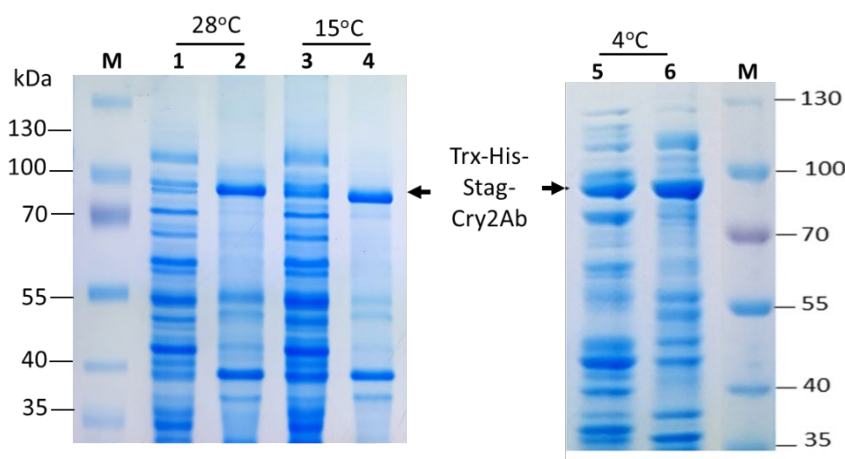
Hình 3.16. Phân tích độ hòa tan của protein tái tổ hợp Trx-His-Stag-Cry2Ab bằng SDS-PAGE (a) và Western blot (b)

1, 2: protein pha cặn và pha tan của tế bào *E. coli* BL21 mang pET32/cry2Ab được cảm ứng IPTG, theo thứ tự tương ứng. ĐC: protein tổng số của tế bào *E. coli* BL21. M: thang protein 10–180 kDa (Thermo Scientific).

3.4.4.1. Ảnh hưởng của nhiệt độ nuôi cấy cảm ứng

Trong biểu hiện gen, nhiệt độ là một trong những yếu tố quan trọng ảnh hưởng đến hiệu suất tổng hợp cũng như chất lượng của protein tái tổ hợp. Về hiệu suất, thông thường các protein ngoại lai được tổng hợp mạnh nhất ở 37°C do đây là nhiệt độ sinh trưởng tốt nhất cho tế bào chủ *E. coli*. Tuy nhiên, khi protein tái tổ hợp được tổng hợp một cách ồ ạt, trong một số trường hợp sẽ dẫn đến tình trạng phần lớn protein đích tạo thành không có khả năng hòa tan, hay nói cách khác là tồn tại ở dạng thể vùi (inclusion body) như trường hợp biểu hiện protein dung hợp Trx-His-Stag-Cry2Ab nêu trên. Nguyên nhân có thể do các protein tham gia vào công đoạn xử lý, đóng gói các chuỗi polypeptide mới được dịch mã để tạo thành phân tử protein có cấu trúc bậc 3 hoặc bậc 4 đặc trưng trong bộ máy sinh tổng hợp protein của tế bào chủ không đáp ứng đủ với lượng quá lớn protein tái tổ hợp được sinh ra. Vì vậy, việc giảm nhiệt độ nuôi cấy nhằm giảm tốc độ sinh trưởng của chủng cũng như giảm hiệu suất quá trình sinh tổng hợp protein có thể là biện pháp hữu hiệu giúp tăng lượng protein tái tổ hợp đảm bảo về hoạt tính và chức năng sinh học. Theo đó, các mức nhiệt độ nuôi cấy cảm ứng thấp hơn 37°C là 28°C, 15°C và 4°C được lựa chọn để khảo sát khả năng tan của protein ngoại lai được biểu hiện. Kết quả phân tích bằng điện di SDS-PAGE (Hình 3.17) cho thấy ở các nhiệt độ cảm ứng

là 28°C và 15°C, protein dung hợp Trx-His-Stag-Cry2Ab vẫn chủ yếu tồn tại ở pha cặn. Trong khi đó, khi hạ nhiệt độ cảm ứng xuống 4°C, kết quả điện di ghi nhận sự thay đổi đáng kể về lượng protein tái tổ hợp thu được ở pha tan, mặc dù hơn 50% lượng protein này vẫn biểu hiện ở dạng thể vùi. Do đó, nhiệt độ nuôi cấy cảm ứng chủng biểu hiện ở 4°C được lựa chọn để thu được nhiều nhất có thể dạng tan của Trx-His-Stag-Cry2Ab.



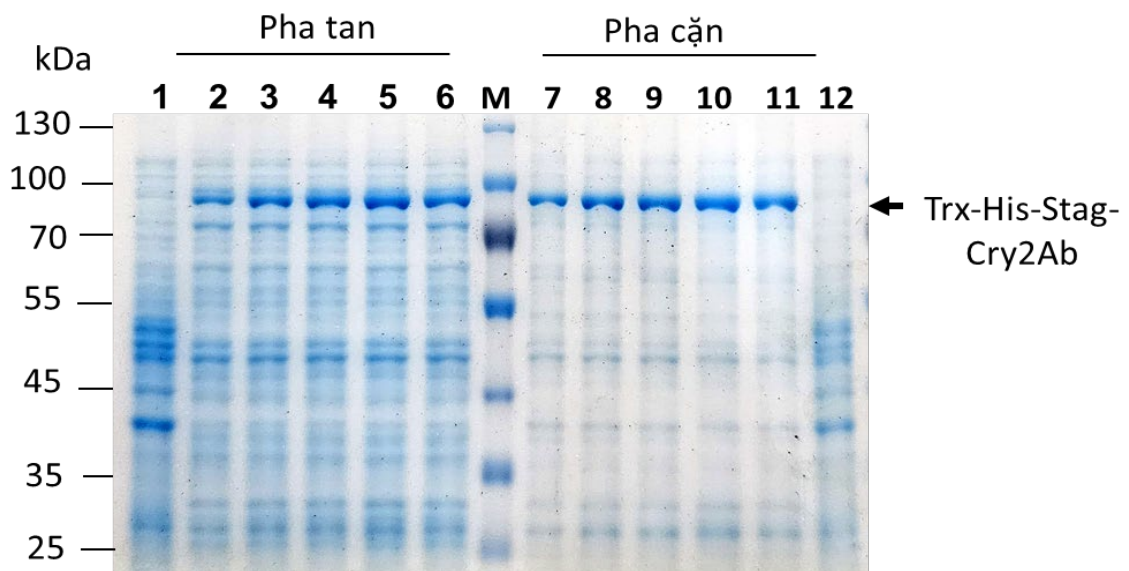
Hình 3.17. Điện di SDS-PAGE phân tích sự biểu hiện của Trx-His-Stag-Cry2Ab ở các nhiệt độ khác nhau

1, 3, 5: protein pha tan của tế bào *E. coli* BL21 mang pET32/cry2Ab được cảm ứng IPTG ở 28°C, 15°C và 4°C, theo thứ tự tương ứng. 2, 4, 6: protein pha cặn của tế bào *E. coli* BL21 mang pET32/cry2Ab được cảm ứng IPTG ở 28°C, 15°C và 4°C, theo thứ tự tương ứng. M: thang protein 10–180 kDa (Thermo Scientific).

3.4.4.2. Ảnh hưởng của nồng độ IPTG

Trên vector biểu hiện pET32/cry2Ab, gen *cry2Ab* được đặt dưới sự điều khiển của promoter T7. Hoạt động của promoter này bị ức chế bởi protein LacI, là sản phẩm của gen *lacI* cùng được thiết kế nằm trên vector pET32a(+). Khi trong môi trường không có chất cảm ứng, gen *lacI* được phiên mã và sinh tổng hợp tạo ra LacI. Chất ức chế này có ái lực với *lac* operator nằm xuôi chiều promoter T7 nên bám chặt vào vùng này, kết quả là ngăn cản sự hoạt động của promoter T7. Khi chất cảm ứng IPTG được thêm vào môi trường nuôi cấy, chất này bám vào LacI làm cho nó không còn ái lực với *lac* operator và tách khỏi phức hệ này. Nhờ đó T7 RNA polymerase có thể bám vào vị trí trên promoter T7 để khởi

động quá trình phiên mã tạo mRNA làm nguyên liệu cho quá trình sinh tổng hợp protein. Tuy nhiên, ở nồng độ cao, IPTG lại gây độc, ức chế sự sinh trưởng của vi khuẩn. Vì vậy, lượng chất cảm ứng được bổ sung vào môi trường chỉ cần vừa đủ để trung hòa chất ức chế và không gây độc cho tế bào chủ. Trong thí nghiệm này, ảnh hưởng của các mức nồng độ IPTG khác nhau (0,01 mM; 0,05 mM; 0,1 mM; 0,5 mM, 1 mM) lên khả năng sinh tổng hợp protein dung hợp Trx-His-Stag-Cry2Ab ở pha tan được khảo sát. Kết quả phân tích bằng điện di SDS-PAGE (Hình 3.18) cho thấy ở nồng độ cảm ứng IPTG 0,5 mM, lượng protein tái tổ hợp thu được ở pha tan là nhiều nhất.



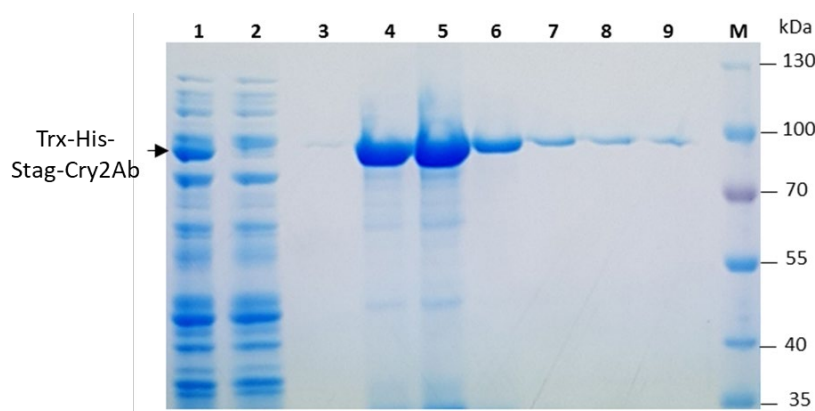
Hình 3.18. Điện di SDS-PAGE phân tích sự biểu hiện của Trx-His-Stag-Cry2Ab ở các nồng độ chất cảm ứng IPTG khác nhau

1, 12: protein pha tan và pha cặn của mẫu đối chứng không được cảm ứng IPTG; 2-6 và 7-11: protein pha tan và pha cặn của tế bào *E. coli* BL21 mang pET32/cry2Ab được cảm ứng IPTG ở các nồng độ lần lượt là 0,01 mM; 0,05 mM; 0,1 mM; 0,5 mM, 1 mM. M: thang protein 10–180 kDa (Thermo Scientific).

Như vậy, kết quả nghiên cứu đã xác định được điều kiện biểu hiện tối ưu để thu được lượng lớn protein tái tổ hợp Trx-His-Stag-Cry2Ab ở pha tan là nhiệt độ nuôi cấy cảm ứng ở 4°C với nồng độ chất cảm ứng IPTG là 0,5 mM. Với quy trình đã được tối ưu này, chúng tôi tiến hành biểu hiện lượng lớn thu sinh khối khuẩn phục vụ cho mục đích tinh sạch protein tái tổ hợp.

3.4.5. Tinh sạch Trx-His-Stag-Cry2Ab

Protein tái tổ hợp được tinh sạch bằng phương pháp sắc kí ái lực, sử dụng cột có gắn ion Ni^{2+} như đã nêu trong phần phương pháp. Nhờ đặc tính của His-tag gồm 6 axit amin histidine nằm liền kề nhau giúp tăng ái lực với ion kim loại nên protein dung hợp Trx-His-Stag-Cry2Ab sẽ tạo liên kết tĩnh điện với ion Ni^{2+} gắn trên hạt Sepharose và được giữ lại trên cột. Trong khi đó, các protein khác không có ái lực hoặc có ái lực yếu với ion kim loại trên cột sẽ bị rửa trôi. Sau khi loại bỏ được các protein tạp nhiễm, protein đích được thu hồi bằng cách bổ sung imidazole là một chất cạnh tranh Ni^{2+} trong phức hợp Ni^{2+} -histidine. Kết quả là, liên kết tĩnh điện giữa Ni^{2+} gắn trên hạt và đuôi His-tag của protein dung hợp bị phá vỡ và protein theo dịch thu ra ngoài cột. Protein thu được trong các phân đoạn được kiểm tra bằng điện di SDS-PAGE (Hình 3.19).



Hình 3.19. Điện di SDS-PAGE phân tích sản phẩm tinh sạch protein tái tổ hợp Trx-His-Stag-Cry2Ab.

1: Protein pha tan của *E. coli* BL21 mang vector pET32/cry2Ab được cảm ứng IPTG 0,5 mM ở 4°C; 2: Dịch protein sau khi gắn cột; 3-9: protein tái tổ hợp tinh sạch thu được các phân đoạn khác nhau. M: thang protein 10–180 kDa (Thermo Scientific).

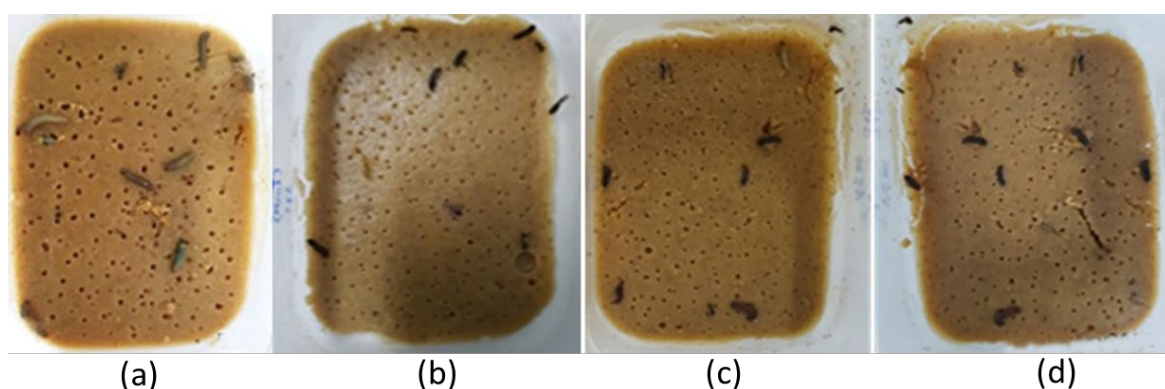
Kết quả điện di cho thấy sản phẩm tinh sạch là một protein có kích thước gần 90 kDa, tương ứng với băng protein dung hợp Trx-His-Stag-Cry2Ab trong dịch chiết protein pha tan của chủng biểu hiện. Hầu hết các protein lẫn tạp trong dịch chiết pha tan đều được loại bỏ hiệu quả. Theo tính toán từ kết quả xác định nồng độ protein bằng phương pháp Bradford, hiệu suất tinh sạch protein tái tổ hợp Trx-His-Stag-

Cry2Ab ước tính đạt 3,2 mg protein tinh chế/100 ml dịch lên men vi khuẩn. Ngoài ra, độ tinh khiết của protein tái tổ hợp trong các phân đoạn được đo bằng máy VersaDoc Imaging system cho kết quả dao động từ 87- 98%.

3.4.6. Xác định giá trị liều gây chết trung bình LC_{50} của protein tái tổ hợp Cry2Ab tinh sạch với ấu trùng sâu đục quả đậu tương *E. zinckenella* Treitschke

Các phân đoạn protein có mức độ tinh sạch trên 90% được lựa chọn để tiến hành loại muối cùng imidazole, đồng thời cô đặc bằng cách sử dụng cột 3K MWCO (hãng Thermo Scientific). Dịch protein Trx-His-Stag-Cry2Ab tinh sạch đã loại bỏ muối và imidazole được xác định nồng độ bằng phương pháp Bradford và sử dụng để xác định liều lượng gây chết 50% (LC_{50}) với ấu trùng *E. zinckenella* Treitschke tuổi 2.

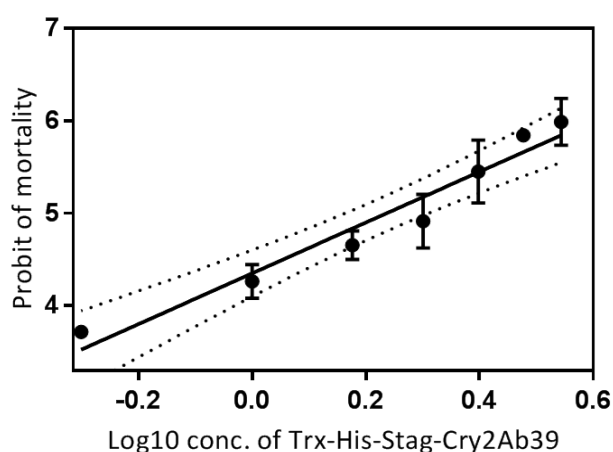
Quan sát sự phát triển của sâu non cũng như thống kê số sâu chết sau khi ăn thức ăn chứa độc tố *Bt* tái tổ hợp chúng tôi nhận thấy: ở môi trường thức ăn nhân tạo có bổ sung dịch khuẩn SP14.2 hoặc dịch tinh sạch protein Trx-His-Stag-Cry2Ab ở nồng độ cao (3,5 và 4,0 μg protein/g thức ăn), một lượng nhỏ ấu trùng (10% - 20%) chết sau 24 giờ ăn độc tố. Với môi trường thức ăn nhân tạo chứa các lượng độc tố thấp hơn (0,5; 1,0; 1,5; 2,0; 3,0 μg protein/g thức ăn) hiện tượng sâu chết xuất hiện muộn hơn (sau 2 ngày). Trong khi đó, ở lô đối chứng âm là môi trường thức ăn chỉ bổ sung đệm PBS, sâu non vẫn sống và phát triển bình thường sau 5 ngày theo dõi (Hình 3.20).



Hình 3.20. Kết quả thử hoạt tính diệt sâu đục quả đậu tương *E. zinckenella* Treitschke của protein tái tổ hợp Trx-His-Stag-Cry2Ab.

(a) Môi trường thức ăn nhân tạo bổ sung đệm PBS; (b) Môi trường thức ăn nhân tạo bổ sung dịch khuẩn *Bt* SP14.2; (c, d) Môi trường thức ăn nhân tạo bổ sung dịch protein Trx-His-Stag-Cry2Ab ở nồng độ 3,0 và 4,0 μg protein/g thức ăn.

Số liệu thể hiện mối tương quan giữa nồng độ độc tố thử nghiệm với tỉ lệ phần trăm số sâu chết sau 5 ngày được ghi nhận và phân tích bằng phương pháp hồi quy Probit của Finney (Hình 3.21). Phép nội suy dựa trên mô hình hồi quy này cho kết quả liều gây chết trung bình (LC₅₀) của Trx-His-Stag-Cry2Ab đối với ấu trùng *E. zinckenella* tuổi 2 ước tính là 1,74 µg protein/g thức ăn. Các tham số hồi quy được trình bày trong bảng 3.12 chứng minh mô hình Probit phù hợp với số liệu (kiểm định Chi square cho giá trị P = 0,88). Khoảng giới hạn tin cậy của giá trị LC₅₀ tương đối hẹp (1,33-2,16 µg protein/g thức ăn) cho thấy kết quả ước tính có độ tin cậy cao.



Hình 3.21. Đường hồi quy Probit được xây dựng dựa trên số liệu về nồng độ độc tố Trx-His-Stag-Cry2Ab sử dụng và tỉ lệ sâu chết để ước tính giá trị LC₅₀.
(Các đường chấm biểu thị khoảng tin cậy 95% của đường hồi quy)

Bảng 3.12: Giá trị liều gây chết trung bình của Trx-His-Stag-Cry2Ab đối với sâu non tuổi 2 *Etiella zinckenella* và các tham số của mô hình hồi quy Probit

LC ₅₀ (µg/g) (95% CI ^a)	Đường hồi quy (Regression line)		Mức độ phù hợp của mô hình (Goodness-of-fit test)		
	Slope ± SE ^b	Intercept ± SE ^b	χ ²	df	P ^c
1,74 (1,33-2,16)	2,74 ± 0,27	4,35 ± 0,09	2,37	6	0,88

^a Khoảng tin cậy (Confidence intervals); ^b Sai số chuẩn (Standard error); ^c Giá trị P của kiểm định Chi-square test (χ²) cho mô hình hồi quy Probit.

Từ kết quả các phân tích đã nêu, có thể khẳng định protein Cry2Ab có nguồn gốc từ chủng SP14.2 là một độc tố mới có hoạt lực tốt trong diệt trừ sâu đục quả đậu *E. zinckenella* Treitschke. Dựa trên mức độ tương đồng về trình tự axit amin (99,21%) với các protein *Bt* có hoạt tính diệt côn trùng đã được công bố, độc tố Cry2Ab được xếp hạng tính mới ở mức 4. Trình tự nucleotide và axit amin suy diễn của gen mã hóa độc tố này đã được đăng kí trên ngân hàng GenBank với mã truy cập lần lượt là MN319700.1 và QIQ19560.1. Độc tố Cry này đã được xét duyệt và công nhận là độc tố mới Cry2Ab39 có khả năng diệt sâu đục quả đậu tương bởi Ủy ban danh pháp độc tố *Bt* do Neil Crickmore đứng đầu (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/), nay là Bacterial Pesticidal Protein Resource Center (<https://www.bpprc.org/>). Đồng thời, chủng *Bt* SP14.2 sản sinh ra độc tố này đã được Cục sở hữu trí tuệ Việt Nam cấp bằng độc quyền sáng chế số 37733 ngày 30/10/2023 (Hình phụ lục 7). Ngoài ra, 2 trình tự nucleotit của 2 gen *cry1Na* của chủng BD8.2 và gen *cry2Ah* của chủng Đ6.1 cũng đã được đăng kí ngân hàng GenBank với mã truy cập lần lượt là MN725076 và MN725074 (Hình phụ lục 8).

3.5. Tối ưu mã gen *cry2Ab39* để phù hợp với hệ thống biểu hiện ở thực vật và biểu hiện tạm thời protein độc tố Cry2Ab39 trong lá thuốc lá *N. benthamiana* bằng phương pháp agroinfiltration.

Để đánh giá khả năng biểu hiện gen *cry2Ab39* trong hệ biểu hiện thực vật, làm cơ sở cho các nghiên cứu tạo cây chuyển gen kháng sâu sau này, công nghệ agroinfiltration được lựa chọn để biểu hiện tạm thời gen đích trên cây thuốc lá *N. benthamiana*. Phương pháp này có những ưu điểm như thu protein nhanh, không bị ảnh hưởng bởi vị trí gắn gen đích trong tế bào thực vật, có thể tiến hành biểu hiện trong các mô đã biệt hóa hoàn toàn như lá, protein biểu hiện ở mức độ cao và không gặp phải vấn đề an toàn sinh học hay rủi ro đến môi trường [92]

3.5.1. Tối ưu mã di truyền của gen *cry2Ab39*

Như ta đã biết, mã di truyền có tính thoái hóa, hay nói cách khác là nhiều bộ ba khác nhau có thể cùng xác định một loại axit amin. Trong quá trình dịch mã tổng hợp protein, hiện tượng ưu tiên sử dụng một số bộ mã codon phổ biến đã được phát

hiện rộng rãi ở nhiều loài sinh vật [93]. Khi sinh vật tiến hóa, mô hình sử dụng các bộ mã phổ biến luôn là một câu hỏi hấp dẫn đối với các nhà nghiên cứu. Sự ưu tiên codon là khái niệm để chỉ hiện tượng một codon mã hóa một axit amin nào đó xuất hiện với tần số cao hơn các codon khác cùng mã hóa cho axit amin đó trong phân tử mRNA. Các loài sinh vật khác nhau sử dụng các bộ mã di truyền khác nhau theo hướng ưu tiên một số bộ ba mã hóa trong số các codon đồng nghĩa cùng mã hóa cho một axit amin. Ngược lại, có nhiều bộ mã rất ít hoặc không bao giờ xuất hiện trong các gen của một số sinh vật, đó gọi là các bộ mã hiếm. Do đó có thể có codon được ưu tiên sử dụng ở loài này nhưng lại trở thành codon hiếm ở loài khác. Cũng vì lý do này, việc đưa gen ngoại lai mang mã di truyền không phù hợp vào tế bào chủ sẽ kìm hãm quá trình dịch mã của gen, giảm sinh khối protein đích.

Theo các nghiên cứu đã được công bố, trong công nghệ cây trồng *Bt*, sự biểu hiện tốt của gen *Bt* trong tế bào chủ là một yếu tố quan trọng và tiên quyết để mang lại tính kháng côn trùng cho cây chuyển gen. Một trong những chiến lược giúp tăng mức độ biểu hiện của gen chuyển là cải biến mã di truyền trong trình tự gen gốc thành các mã phù hợp hơn với hệ thống biểu hiện của vật chủ mà không làm thay đổi trình tự axit amin [94]. Vì vậy, trong nghiên cứu này, mã di truyền của gen *cry2Ab39* được tối ưu sao cho tương thích với hệ thống biểu hiện của cây thuốc lá *N. benthamiana*. Căn cứ vào bảng tần suất sử dụng các codon của *N. benthamiana* trên ngân hàng dữ liệu mã di truyền Codon Usage Database (<https://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=4100>) [83], các codon hiếm (tần suất sử dụng dưới 50%) trong trình tự của gen *cry2Ab39* kiểu dại được loại bỏ và thay thế bằng các codon được loài thực vật này sử dụng với tần suất cao hơn mà không làm thay đổi trình tự axit amin. Việc lựa chọn codon thay thế cũng theo hướng ưu tiên các codon có chứa G/C nhằm làm tăng hàm lượng GC tổng số của cả gen cũng như tạo sự phân bố GC đồng đều dọc theo chiều dài gen. Bên cạnh đó, một số motif như ATTTA (được cho là yếu tố gây bất ổn định cho phân tử mRNA), hoặc các trình tự giống với tín hiệu poly A (AATATT, AATTAA, AATAAA và AATAAT) cũng như các vị trí được dự đoán tiềm ẩn nguy cơ gây hiện tượng cắt nối phân tử mRNA cũng sẽ được loại trừ và thay thế. Ngoài ra, việc cải biến mã di truyền cũng được xem xét sao cho hạn chế tối đa khả năng hình

thành cấu trúc thứ cấp của mRNA- một trong những nguyên nhân kìm hãm quá trình dịch mã. Kết quả phân tích trình tự nucleotide sau quá trình cải biến mã cho thấy, gen *cry2Ab39* sau khi được tối ưu (kí hiệu là *cry2Ab39opt*) sai khác 24,15% so với gen nguyên bản phân lập từ vi khuẩn *Bt*, tuy nhiên các vị trí thay thế nucleotide không làm thay đổi trình tự amino axit. Ngoài ra, trong quá trình cải biến, với mục đích giữ lại protein tái tổ hợp ở lưới nội chất, chúng tôi đã thêm 69 nucleotide mã hóa cho LeB4 ER-signal peptide vào đầu 5' và đoạn mã hóa trình tự KDEL vào đầu 3' của gen đích. Đồng thời, để tạo thuận lợi cho việc phát hiện cũng như tinh sạch protein tái tổ hợp, trình tự mã hóa hexahistidine cũng được thêm vào đầu 3' của gen trước vị trí trình tự KDEL (Hình 3.22, Hình phụ lục 9).

```

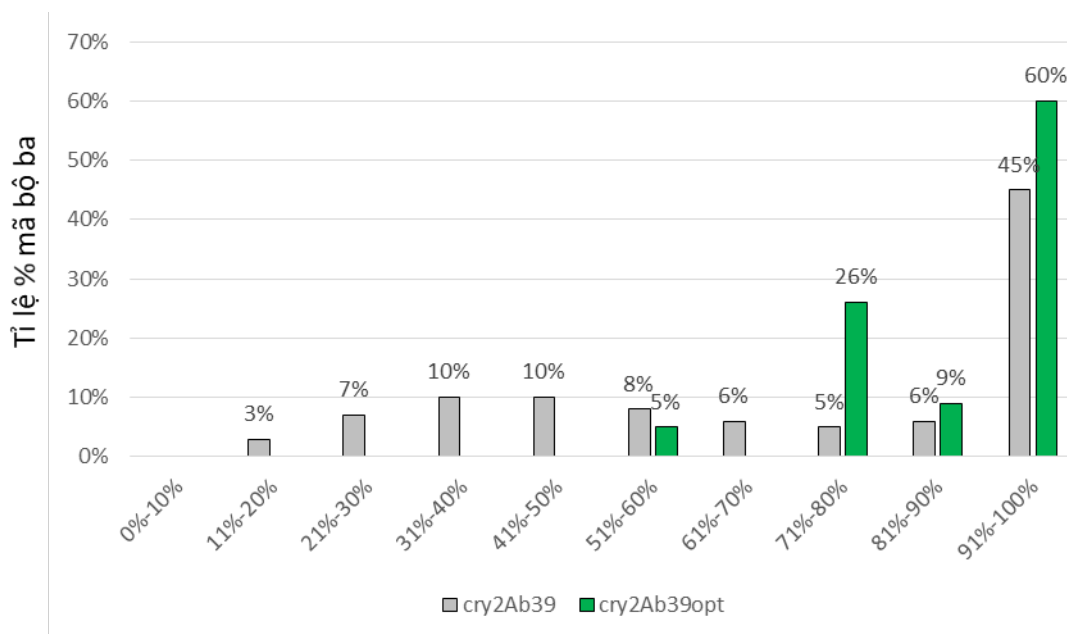
1      ATGGCTTCAAAGCCATTCTTGTCTCTTCTTTTCATTGTCTCTTTTGGCTTTTTACTTCAACTTGTGTTGGCTAAATCTGTGCT      80
81      TAACCTGGAAGGACTACTATTTGCGACGCATACAACGTTGGCAGCTCACGACCTTTTTTCATTTCAACACAAATCTTTGG      160
161     ACACCTGTGCAAAAAGAGTGGACAGAGTGGAAAAAACAATCATTCTTTTACTTGGATCCTATCGTGGGTACTGTGGCA      240
241     TCTTTCTTGTCTAAGAAGGTGGGTTCACTTGTGGTAAGAGGATCCTTTCTGAGCTTAGAAACCTTATCTTTCCATCTGG      320
321     ATCTACAAAATTTGATGCAAGATATCTTGAGGGAGACAGAAAAATTTTGAATCAGAGGTTGAACACTGACACTCTTGCTA      400
401     GGGTCAATGCTGAGTTGACTGGATTGCAAGCTAACCTGGAAGAGTTCAATAGGCAAGTTGACAACTTTTGAATCCAAAC      480
481     AGAAATGCTGTTCCCTCTTTCAATTACATCTTCTGTAAACAATGCAACAACCTTTTTCTTAATAGACTTCCACAGTTCCA      560
561     GATGCAAGGATACCAAGTTGTTGTTGCTTCCATTGTTGCTCAAGCTGCAAACTTCACTTGTTCATTATTAGGGATGTGA      640
641     TCCTTAACGCTGACGAGTGGGGTATCTCTGCTGCACATTGAGGACTTATAGAGATTACTTGAAAAACTATACTAGAGAT      720
721     TACTCTAATTTATTGTATTAACACTTATCAATCAGCTTTCAGAGGACTTAACACTAGGCTTCATGATATGCTTGAGTTAG      800
801     AACATATATGTTTTGAACGTGTTTGAATACGTTTCTATTTGGTCTCTTTTCAAGTATCAGTCACTCTTGTGTCTCTG      880
881     GAGCAAACCTGTACGCTTCTGGTCTCGGACCACAGCAGACACAGTCATTCACTTCACAAGACTGGCCATTTTTGTATTCA      960
961     CTTTTCCAAGTGAACCTAAGTATGTGTTGAAAGGATTTCTGGAGCTAGGTTGACACAGACTTTCCCTAACATCGTGGG      1040
1041    TTTGCCCGGAACACTACTACTACACACGCTTTGCTTGCAGCAAGGTTGAACACTCTGGAGGTGTGCTTCTGGAGATATCG      1120
1121    GTGCTTCTCCTTTCAACCAAAAATTTAACTGTTCTACTTTTCTTCCCTCCACTTTTGACACCTTTTGGTGGAGTCTGGTTG      1200
1201    GACTCTGGTCTGATAGAGAGGGAGTTGCAACAGTGACAACCTGGCAGACTGAATCTTTGAGACTACTCTTGGATTGAG      1280
1281    GTCTGGAGCATTACAGCAAGGGTAATTCTAACTACTTCCAGATTACTTCATTAGGAACATCTCTGGAGTGCCATTGG      1360
1361    TGGTGAGAAACGAGGACTTGAGGAGACCATGCACTACAACGAGATTAGAAATATCGCATCACCTTCTGGAACCTCCCGGA      1440
1441    GGAGCAAGGGCATACATGGTGTCTGTGCATAAATAGAAAGAAACAATATCCACGCTGTGCACGAGAAACGGATCTATGATCCA      1520
1521    CCTTGCACCTAACGACTACACTGGATTTACAATTTCAACATTCATGCAACACAAGTTAATAACCAAACAGAACATTCA      1600
1601    TCTCTGAGAAGTTCGGTAACCAAGGAGACTCATTGAGGTTGAGCAGACAACATCAACTGCTAGATATACACTTAGGGGA      1680
1681    AATGGAAACTCATATAACCTTTATTTGAGGGTTCTTCAATGGAAACTCAACTATTAGGGTTACTACTCAATGGGAGGGT      1760
1761    GTACACTGCAACTAACGTTAACACAAACAACAACAATGACGGAGTGAATGACAACGGTCTAGGTTCTCAGACATCAACA      1840
1841    TCGGAAACGTGGTGGCATCATCTAACTCAGACGTGCCTTTGGACATCAACGTGACATTGAACCTGGAACCTCAGTTGAT      1920
1921    TTGATGAACATTATGCTTGTTCCTACTAACATTTACACCTTTGTATAATGGTTCTCATCATCACCACTCACAGGACGA      2000
2001    GCTTTGA      2007
    
```

Hình 3.22. Trình tự gen đã cải biến *cry2Ab39opt*.

Các vị trí được tô màu xám đánh dấu các nucleotide được cải biến so với gen *cry2Ab39*. Trình tự nucleotide màu xanh dương mã hóa cho chuỗi peptide tín hiệu LeB4. Trình tự nucleotide màu đỏ mã hóa hexahistidine (6X His tag). Các nucleotide màu xanh lá mã hóa trình tự KDEL.

Sử dụng công cụ phân tích mã bộ ba hiếm (Rare Codon Analyzer- <https://www.biologicscorp.com/tools/RareCodonAnalyzer>) để kiểm tra sự phù hợp của trình tự gen đích với hệ biểu hiện *N. Benthamiana*, chúng tôi thấy rằng sau khi được cải biến gen *cry2Ab39opt* chứa các mã bộ ba đều có tần suất được sử dụng ở mức cao trên 50% và không còn các mã bộ ba hiếm. Tỷ lệ các mã bộ ba có tần suất

sử dụng cao từ 91-100% tăng từ 45% lên 60%, trong khi đó các bộ mã được sử dụng với tần suất 71-80% tăng từ 5% lên 26%, còn lại 5% các mã bộ ba phân bố ở mức 51%-60%, 9% phân bố ở mức 81%-90% (Hình 3.23).



Hình 3.23. Sự phân bố tỉ lệ phần trăm các mã bộ ba phù hợp của gen *cry2Ab39* với hệ biểu hiện *N. benthamiana* trước và sau khi được cải biến mã di truyền.

Bảng 3.13: So sánh các thông số về trình tự giữa gen *cry2Ab39* kiểu dại và gen được cải biến *cry2Ab39opt*

Thông số	<i>cry2Ab39</i>			<i>cry2Ab39opt</i>		
	Base	Số lượng	Tỉ lệ %	Base	Số lượng	Tỉ lệ %
Hàm lượng các base	A	661	32.45	A	573	28.13
	T	659	32.35	T	626	30.73
	G	368	18.07	G	414	20.32
	C	349	17.13	C	424	20.81
GC%		35.2%		45.1%		
Các motif gây bất ổn mRNA		76		0		
Cấu trúc mRNA thứ cấp		12		0		

Ngoài ra, việc tối ưu mã cũng giúp tăng hàm lượng GC của gen đích từ 35,2% lên 45,1%. Đáng chú ý các motif hoặc trình tự tiềm ẩn nguy cơ tạo cấu trúc mRNA thứ cấp hoặc gây bất ổn cho phân tử mRNA đều được loại trừ hoàn toàn (Bảng 3.13). Trình tự gen *cry2Ab39* sau khi được cải biến mã được thiết kế thêm

các điểm cắt của *NcoI* và *XmaI* ở 2 đầu được đặt tổng hợp nhân tạo và nhân dòng trong vector thương mại pUC57/*cry2Ab39opt*.

3.5.2. Thiết kế các vector chuyển gen pFGC/35S_ *cry2Ab39opt* và pFGC/*rbcS1A-pro_cry2Ab39*

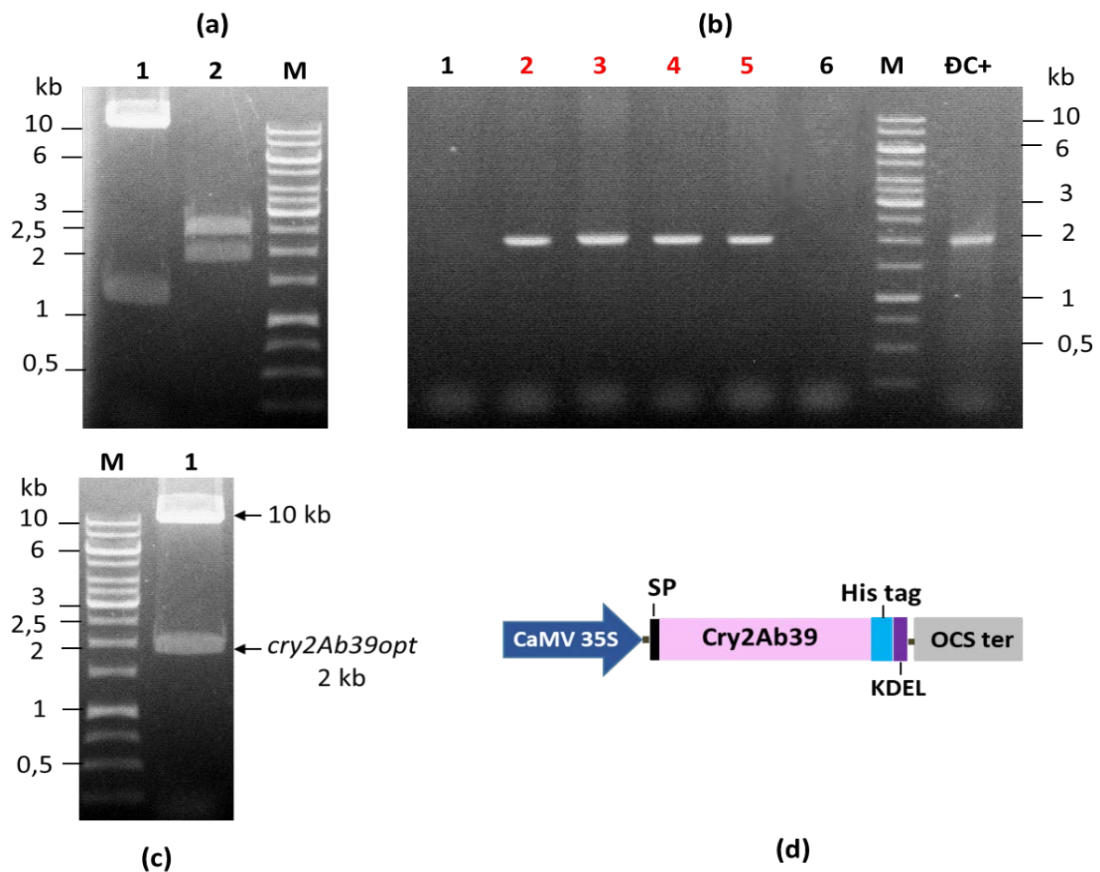
3.5.2.1. Thiết kế các vector chuyển gen pFGC/35S_ *cry2Ab39opt*

Như đã nói, gen *cry2Ab39opt* đã được tổng hợp nhân tạo và tách dòng trong vector pUC57/*cry2Ab39opt*. Trong khi đó, vector thương mại pFGC5941 đã có sẵn cassette biểu hiện gen trong thực vật kèm vị trí đa nối để chèn thêm gen đích. Cassette biểu hiện này được điều khiển bởi promoter 35S và kết thúc bởi OCS (octopine synthase) terminator. Do vậy, để tạo các đầu bắt cặp bổ sung nhằm nối ghép gen *cry2Ab39opt* vào pFGC5941 tạo thành plasmid tái tổ hợp pFGC/35S_ *cry2Ab39opt*, các vector pUC57/*cry2Ab39opt* và pFGC5941 đã được xử lý đồng thời bởi cặp enzyme hạn chế *NcoI* và *XmaI*. Sản phẩm cắt được điện di trên gel agarose 0,8%.

Theo tính toán lý thuyết, vector pFGC5941 có kích thước 11,4 kb khi được xử lý với *NcoI* và *XmaI* sẽ tạo thành 2 đoạn DNA: khung vector pFGC5941 chứa cassette biểu hiện (gần 10 kb) và phân mảnh *chsA* (chalcone synthase A) intron có kích thước khoảng 1,4 kb. Tương tự, sản phẩm phản ứng cắt vector pUC57/*cry2Ab39opt* cũng gồm 2 đoạn DNA là khung vector pUC57 (khoảng 2,7 kb) và một đoạn có kích thước hơn 2 kb ứng với gen *cry2Ab39opt*. Kết quả điện di kiểm tra sản phẩm của các phản ứng cắt enzyme hạn chế (Hình 3.24a) đều thu được các băng DNA đúng với tính toán.

Hai phân đoạn DNA là khung vector pFGC5941 chứa cassette biểu hiện và gen *cry2Ab39opt* được thổi gel và tinh sạch làm nguyên liệu cho phản ứng nối ghép nhờ xúc tác của T4-DNA ligase. Vector tái tổ hợp tạo thành được kí hiệu là pFGC/35S_ *cry2Ab39opt* và được nhân dòng trong *E. coli* DH5 α . Việc chọn dòng vi khuẩn mang vector mong muốn được thực hiện nhờ phản ứng PCR sử dụng cặp mồi *cry2Ab39opt_F/R* (Bảng 2.1) đặc hiệu cho gen *cry2Ab39opt*, trong đó các dòng dương tính sẽ cho kích thước sản phẩm là đoạn DNA hơn 2 kb (Hình 3.24b). Ngoài ra, để chắc chắn hơn nữa về sự nối ghép thành công vector pFGC/35S_ *cry2Ab39opt*, phản ứng cắt plasmid tái tổ hợp bằng enzyme hạn chế

NcoI và *XmaI* được tiến hành. Hình ảnh điện di kiểm tra sản phẩm cắt trên gel agarose 0,8% (Hình 3.24c) thể hiện 2 băng vạch có kích thước lần lượt là gần 10 kb (khung vector pFGC5941) và hơn 2 kb (gen *cry2Ab39opt*) đúng như tính toán. Các kết quả này khẳng định việc thiết kế thành công vector chuyển gen pFGC/35S_ *cry2Ab39opt* (Hình 3.24d). Vector này sau đó được biến nạp vào vi khuẩn *A. tumefaciens* AGL1 bằng phương pháp xung điện để phục vụ mục đích chuyển gen vào thực vật.



Hình 3.24. Kết quả thiết kế vector chuyển gen pFGC/35S_ *cry2Ab39opt*

(a) Điện di agarose sản phẩm xử lý các vector pFGC5941 (đường chạy 1) và pUC57/*cry2Ab39opt* (đường chạy 2) bằng cặp enzyme hạn chế *NcoI* và *XmaI*. M: Thang DNA 1 Kb (Thermo Scientific).

(b) Điện di agarose sản phẩm PCR sàng lọc các dòng *E. coli* mang plasmid tái tổ hợp pFGC/35S_ *cry2Ab39opt*. 1-6: các dòng vi khuẩn được sàng lọc, ĐC+: mẫu sử dụng khuôn cho phản ứng PCR là plasmid pUC57/*cry2Ab39opt*.

(c) Điện di agarose kiểm tra sản phẩm cắt plasmid tái tổ hợp pFGC/35S_ *cry2Ab39opt* bằng *NcoI* và *XmaI*.

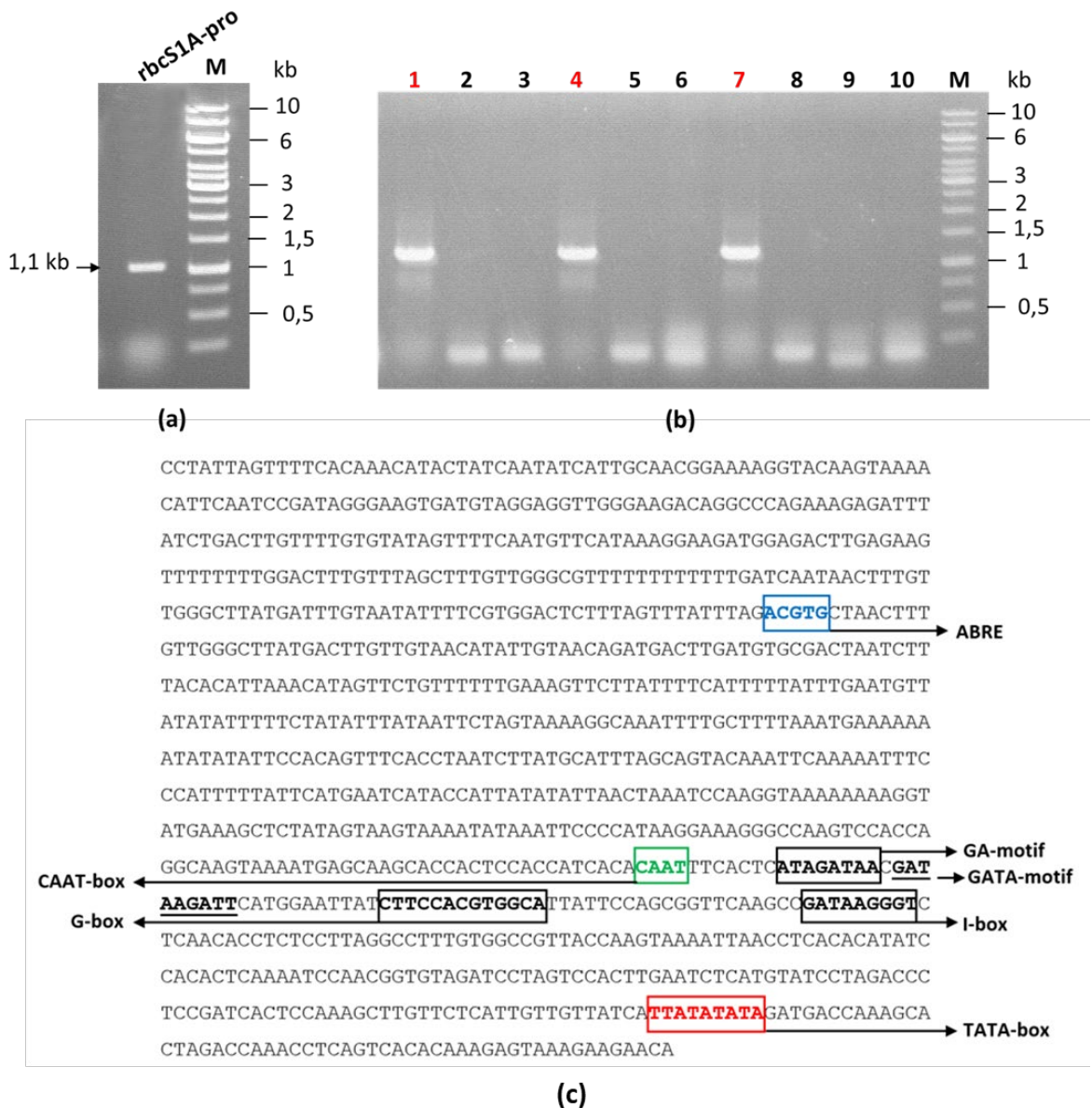
(d) Mô hình cassette biểu hiện gen *cry2Ab39opt* được điều khiển bởi promoter 35S trên vector pFGC/35S_ *cry2Ab39opt*.

3.5.2.2. Thiết kế các vector chuyển gen pFGC/*rbcS1A-pro_cry2Ab39opt*

❖ *Phân lập, tách dòng và giải trình tự promoter rbcS1A từ Arabidopsis thaliana*

Dựa trên trình tự nucleotide của gen *rbcS1A* mã hóa cho tiểu đơn vị ribulose 1,5-biphosphate carboxylase ở *Arabidopsis thaliana* trên ngân hàng Genbank với mã số X13611.1, cặp mồi đặc hiệu *rbcS1A-pro_F/R* (Bảng 2.1) đã được thiết kế nhằm khuếch đại vùng promoter của gen này từ DNA tổng số đã được tách chiết từ lá cây *Arabidopsis*. Để thuận tiện cho việc thiết kế vector trong các thí nghiệm tiếp theo, cả mồi xuôi và mồi ngược đều được gắn thêm trình tự các điểm cắt tương ứng của *EcoRI* và *XhoI*. Kết quả điện di (Hình 3.25a) cho thấy sản phẩm PCR chỉ có một băng duy nhất có kích thước khoảng 1,1 kb, tương đương với kích thước đoạn promoter *rbcS1A* chúng tôi muốn nhân bản. Đoạn DNA này sau đó được tinh sạch, nối ghép vào vector tách dòng pJET1.2/blunt và biến nạp vào chủng *E. coli* DH5 α . Kết quả sàng lọc các thể biến nạp bằng phương pháp PCR với cặp mồi pJET1.2_F/R (Hình 3.25b) cho thấy đã thu được các dòng dương tính.

Để khẳng định việc nhân dòng thành công đoạn promoter *rbcS1A*, plasmid từ dòng khuẩn *E. coli* dương tính được tách chiết và giải trình tự theo phương pháp Sanger. Kết quả phân tích cho thấy đoạn DNA phân lập được có chiều dài 1,1 kb, tương đồng 99,9% với vùng promoter của gen *rbcS1A* ở *Arabidopsis* (mã Genbank X13611.1). Việc phân tích chi tiết hơn về các yếu tố điều hòa *cis* định vị trên vùng trình tự này (Hình 3.25c) sử dụng công cụ web Plant Care (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) đã giúp xác định được các trình tự đặc trưng cho một promoter của sinh vật nhân chuẩn như TATA-box, CAAT-box. Ngoài ra đoạn promoter *rbcS1A* đã phân lập có chứa các nhóm yếu tố hoạt hóa *cis* (*cis-acting element*) liên quan tới đáp ứng với ánh sáng bao gồm: ABRE, G-box, I-box, GA motif, GATA motif.

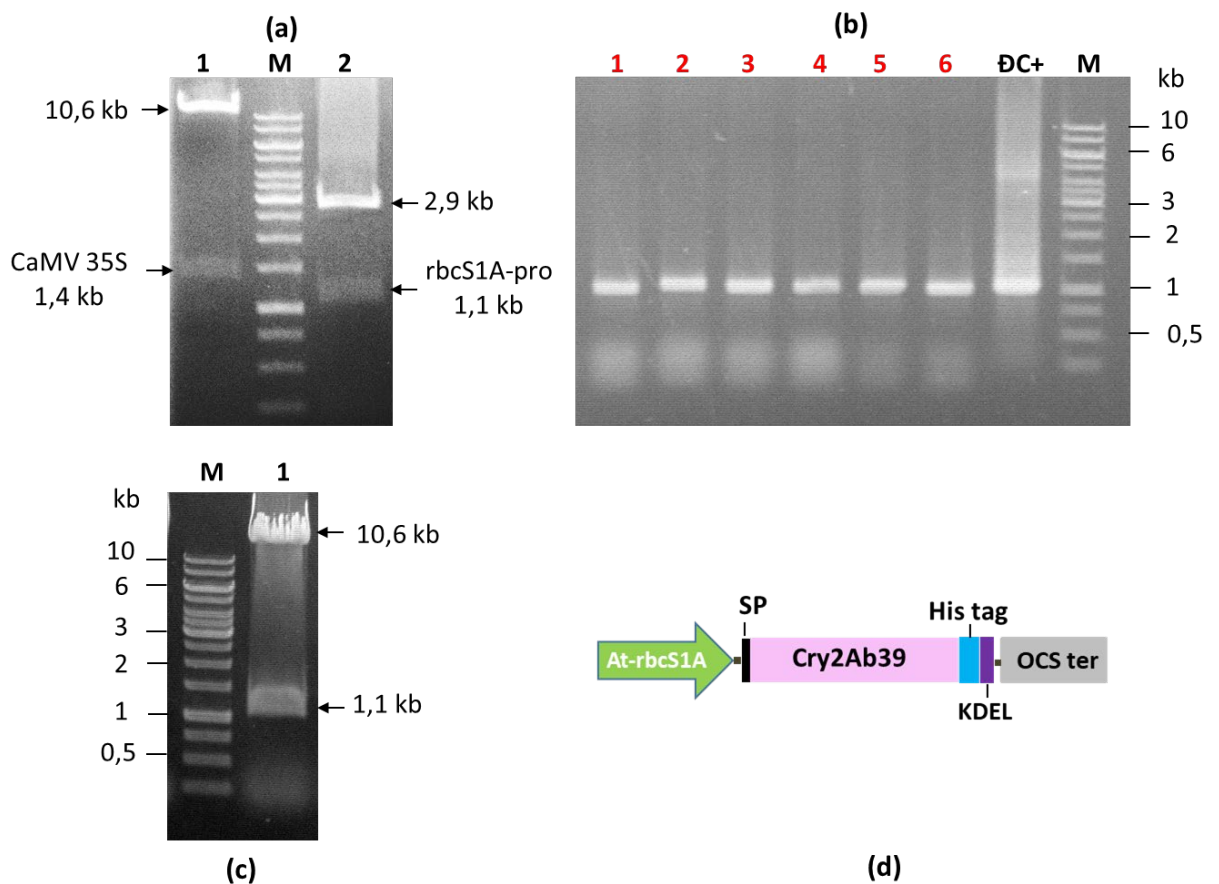


Hình 3.25. Kết quả phân lập, nhân dòng và phân tích trình tự promoter *rbcS1A* (a): sản phẩm PCR khuếch đại promoter *rbcS1A* từ DNA tổng số tách chiết từ *Arabidopsis*, M: thang DNA 1 Kb (Thermo Scientific). (b): kết quả PCR sàng lọc các dòng *E. coli* mang plasmid tái tổ hợp pJET/*rbcS1A-pro*. (c): các yếu tố điều hòa *cis* định vị trên promoter *rbcS1A*

❖ *Thiết kế vector chuyển gen pFGC/rbcS1A-pro_cry2Ab39opt*

Để loại bỏ promoter 35S trên plasmid pFGC/35S_cry2Ab39opt và thay thế bằng promoter *rbcS1A*, các vector pFGC/35S_cry2Ab39opt và pJET/*rbcS1A-pro* được xử lý đồng thời bởi cặp enzyme hạn chế *EcoRI* và *XhoI*. Sản phẩm cắt được điện di trên gel agarose 0,8%. Theo tính toán lý thuyết, vector

pFGC/35S_cry2Ab39opt khi được cắt bởi *EcoRI* và *XhoI* sẽ tạo thành 2 đoạn DNA: phân đoạn promoter 35S có kích thước khoảng 1,4 kb và phân đoạn 10,6 kb tương ứng với phần còn lại của vector. Tương tự, sản phẩm phản ứng cắt vector pJET/rbcS1A-pro cũng gồm 2 đoạn DNA là khung vector pJET1.2/blunt (khoảng 2,9 kb) và một đoạn có kích thước khoảng 1,1 kb ứng với promoter rbcS1A. Kết quả điện di kiểm tra sản phẩm của các phản ứng cắt enzyme hạn chế (Hình 3.26a) đều thu được các băng DNA đúng với tính toán.



Hình 3.26. Kết quả thiết kế vector chuyển gen pFGC/rbcS1A-pro_cry2Ab39opt
 (a) Kết quả xử lý các vector pFGC/35S_cry2Ab39 (đường chạy 1) và pJET/rbcS1A-pro (đường chạy 2) bằng cặp enzyme hạn chế *EcoRI* và *XhoI*. M: Thang DNA 1 Kb (Thermo Scientific).
 (b) Kết quả PCR sàng lọc các dòng *E. coli* mang plasmid tái tổ hợp pFGC/rbcS1A-pro_cry2Ab39opt. 1-6: các dòng vi khuẩn được sàng lọc, ĐC+: mẫu sử dụng khuôn cho phản ứng PCR là plasmid pJET/rbcS1A-pro
 (c) Cắt kiểm tra plasmid tái tổ hợp pFGC/rbcS1A-pro_cry2Ab39opt bằng *EcoRI* và *XhoI*.

(d) Mô hình cassette biểu hiện gen *cry2Ab39opt* được điều khiển bởi promoter *rbcS1A* trên vector pFGC/*rbcS1A-pro_cry2Ab39opt*.

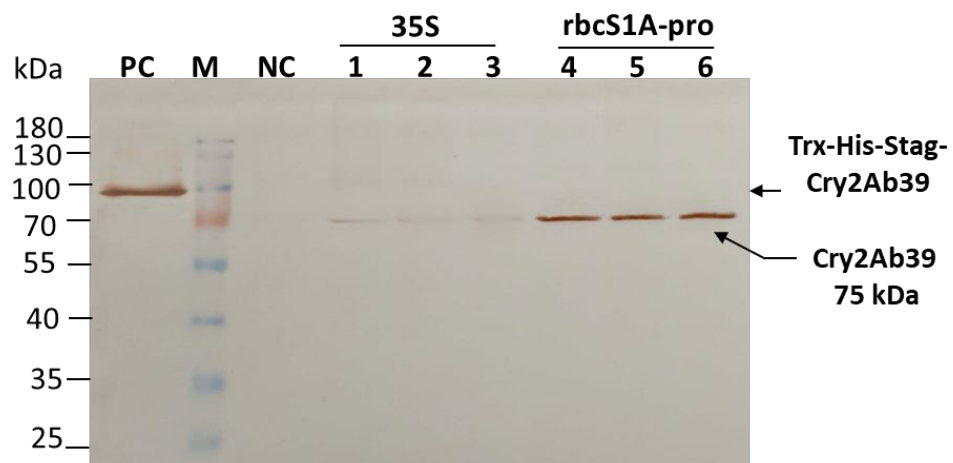
Hai phân đoạn DNA là promoter *rbcS1A* và khung vector pFGC/35S_*cry2Ab39opt* đã loại bỏ promoter 35S được thổi gel và tinh sạch làm nguyên liệu cho phản ứng nối ghép nhờ xúc tác của T4-DNA ligase. Vector tái tổ hợp tạo thành được kí hiệu là pFGC/*rbcS1A-pro_cry2Ab39opt* và được nhân dòng trong *E. coli* DH5 α . Việc chọn dòng vi khuẩn mang vector mong muốn được thực hiện nhờ phản ứng PCR sử dụng cặp mồi *rbcS1A_F/R* (Bảng 2.1) đặc hiệu cho gen promoter *rbcS1A*, trong đó các dòng dương tính sẽ cho kích thước sản phẩm là đoạn DNA khoảng 1,1 kb (Hình 3.26b). Ngoài ra, để chắc chắn hơn nữa về sự nối ghép thành công vector pFGC/*rbcS1A-pro_cry2Ab39opt*, plasmid tái tổ hợp được cắt kiểm tra bằng chính cặp enzyme hạn chế *EcoRI* và *XhoI*. Hình ảnh điện di kiểm tra sản phẩm cắt trên gel agarose 0,8% (Hình 3.26c) thể hiện 2 băng vạch có kích thước lần lượt là 10,6 kb (khung vector pFGC5941 chứa gen *cry2Ab39opt*) và 1,1 kb (promoter *rbcS1A*) đúng như tính toán. Các kết quả này khẳng định việc thiết kế thành công vector chuyển gen pFGC/*rbcS1A-pro_cry2Ab39opt* (Hình 3.26d). Vector này sau đó được biến nạp vào vi khuẩn *A. tumefaciens* AGL1 bằng phương pháp xung điện để phục vụ mục đích chuyển gen vào thực vật.

3.5.3. Biểu hiện tạm thời protein tái tổ hợp Cry2Ab39 trong lá thuốc lá *N. benthamiana*

Agroinfiltration là một phương pháp biểu hiện tạm thời protein ở thực vật sử dụng vi khuẩn *Agrobacterium tumefaciens*. Để đạt được mức độ biểu hiện protein tái tổ hợp tốt nhất, nhiều yếu tố liên quan cần phải được tối ưu. Trong nghiên cứu này, chúng tôi khảo sát ba yếu tố tác động đến hiệu quả tổng hợp protein đích trong tế bào lá thuốc lá *N. benthamiana* là promoter điều khiển cassette biểu hiện gen ngoại lai, mật độ vi khuẩn sử dụng cho lây nhiễm và thời gian thu hoạch lá sau khi lây nhiễm.

3.5.3.1. Ảnh hưởng của promoter đến sự biểu hiện tạm thời của gen *cry2Ab39opt*

Promoter là một thành phần quan trọng trong cấu trúc của tất cả các gen và đóng vai trò then chốt trong việc biểu hiện gen. Trong lĩnh vực công nghệ gen thực vật, nhiều loại promoter được sử dụng rất rộng rãi để điều khiển hoạt động của gen ngoại lai trong cây trồng biến đổi gen. Theo khả năng biểu hiện gen, promoter được chia làm 2 loại chính là promoter cơ định và promoter đặc hiệu. Tùy thuộc vào đặc tính của gen chuyển hoặc mục đích biểu hiện gen mà ta lựa chọn các loại promoter cho phù hợp.



Hình 3.27. Phân tích sự biểu hiện của gen *cry2Ab39opt* dưới sự điều khiển của promoter 35S và *rbcS1A-pro* bằng Western blot sử dụng kháng thể anti-6X His. 1, 2, 3: 5 μ g dịch chiết protein tổng số từ lá cây thuốc lá được xâm nhiễm với *Agrobacterium* mang cấu trúc chuyển gen pFGC/35S_*cry2Ab39*; 4, 5, 6: 5 μ g dịch chiết protein tổng số từ lá cây thuốc lá được xâm nhiễm với *Agrobacterium* mang cấu trúc chuyển gen pFGC/*rbcS1A-pro*_*cry2Ab39*; NC (đối chứng âm): dịch chiết protein tổng số từ lá cây không lây nhiễm; PC (đối chứng dương): protein dung hợp Trx-His-Stag-Cry2Ab39 (khoảng 90 kDa) được biểu hiện trong *E. coli* BL21; M: thang protein 10 -180 kDa (Thermo Scientific).

Trong thí nghiệm này, tính hiệu quả của 2 promoter là CaMV 35S và *rbcS1A-pro* trong việc điều khiển biểu hiện gen *cry2Ab39* trong lá thuốc lá *N. benthamiana* được phân tích và so sánh. Sau 5 ngày biến nạp, lượng protein tái tổ hợp có trong dịch chiết thô từ lá thuốc lá được lây nhiễm vi khuẩn mang 2 cấu trúc

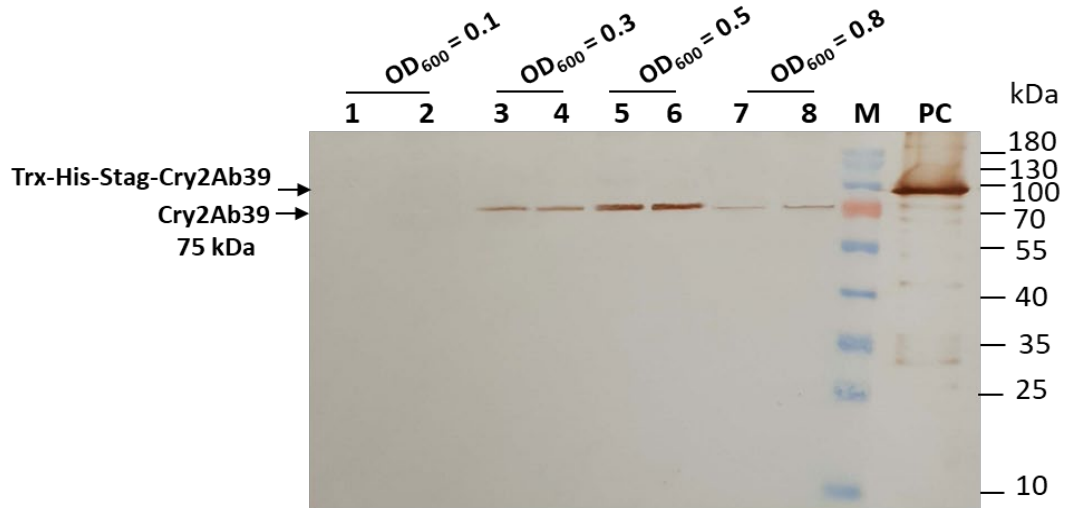
chuyển gen được phân tích bằng kỹ thuật Western blot sử dụng kháng thể anti-6X His. Về lý thuyết, protein Cry2Ab39 tái tổ hợp gắn đuôi His tag sau khi được tổng hợp sẽ có kích thước gần 75 kDa. Kết quả lai Western blot (Hình 3.27) phát hiện được băng protein có kích thước đúng như tính toán ở cả 2 loại mẫu lá được xâm nhiễm khuẩn. Trong khi đó, ở mẫu đối chứng âm là cây thuốc lá không lây nhiễm thì không có sự xuất hiện của băng protein này, chứng tỏ gen *cry2Ab39opt* đã hoạt động dưới sự điều khiển của cả 2 promoter. Tuy nhiên, mức độ biểu hiện của gen này ở lá thuốc lá dưới sự kiểm soát của 2 promoter có sự khác biệt rõ ràng. Cụ thể, băng protein ở các mẫu chuyển gen với cấu trúc pFGC/rbcS1A-pro_cry2Ab39 có độ dày và đậm hơn hẳn so với băng protein ở các mẫu chuyển gen với cấu trúc pFGC/35S_cry2Ab39. Kết quả này chỉ ra rằng promoter rbcS1A hiệu quả hơn so với promoter 35S trong việc biểu hiện tạm thời gen *cry2Ab39* trong lá thuốc lá *N. benthamiana*.

3.5.3.2. Ảnh hưởng của mật độ vi khuẩn xâm nhiễm đến sự biểu hiện tạm thời của gen *cry2Ab39opt*

Mật độ tế bào vi khuẩn được xác định bởi giá trị quang phổ OD ở bước sóng 600 nm của dịch lỏng vi khuẩn, có liên quan tới sinh khối tế bào của chúng hay số lượng tế bào trong một thể tích nhất định. Trong nghiên cứu biểu hiện tạm thời, việc xác định mật độ vi khuẩn phù hợp, giai đoạn vi khuẩn sinh trưởng mạnh mẽ có ý nghĩa quan trọng để hiệu quả chuyển gen được cao nhất. Thông thường, hàm lượng protein biểu hiện sẽ tỉ lệ thuận với nồng độ dịch khuẩn biến nạp. Tuy nhiên, nồng độ dịch khuẩn khi vượt quá ngưỡng thích hợp thì không những không làm tăng thêm mà thậm chí có thể ức chế sự biểu hiện của protein đích.

Để xác định nồng độ khuẩn phù hợp nhất cho biểu hiện tạm thời gen *cry2Ab39opt* trong *N. benthamiana*, thí nghiệm agroinfiltration được tiến hành với dịch khuẩn mang cấu trúc chuyển gen pFGC/rbcS1A-pro_cry2Ab39 được pha loãng tới các giá trị OD₆₀₀ = 0,1; 0,3; 0,5 và 0,8. Các mẫu lá được thu thập sau 5 ngày nhiễm khuẩn, và mức độ biểu hiện tạm thời của *cry2Ab39opt* được đánh giá bằng Western blot. Từ kết quả lai miễn dịch (Hình 3.28) và phân tích hình ảnh bằng phần mềm Image J để so sánh độ đậm nhạt của băng protein, chúng tôi nhận thấy sự tổng hợp protein Cry2Ab39 đạt hiệu quả tốt nhất ở các mẫu được lây nhiễm với

dịch khuẩn có $OD_{600} = 0,5$. Trong khi đó, với các mẫu được xâm nhiễm với dịch khuẩn có $OD_{600} = 0,1$, phương pháp Western blot chưa đủ nhạy để phát hiện được sự biểu hiện của protein đích. Kết quả này chỉ ra rằng dịch huyền phù vi khuẩn *Agrobacterium* với mật độ $OD_{600} = 0,5$ là tối ưu cho sự biểu hiện tạm thời của gen *cry2Ab39opt* trong lá thuốc lá *N. benthamiana* bằng phương pháp agroinfiltration.



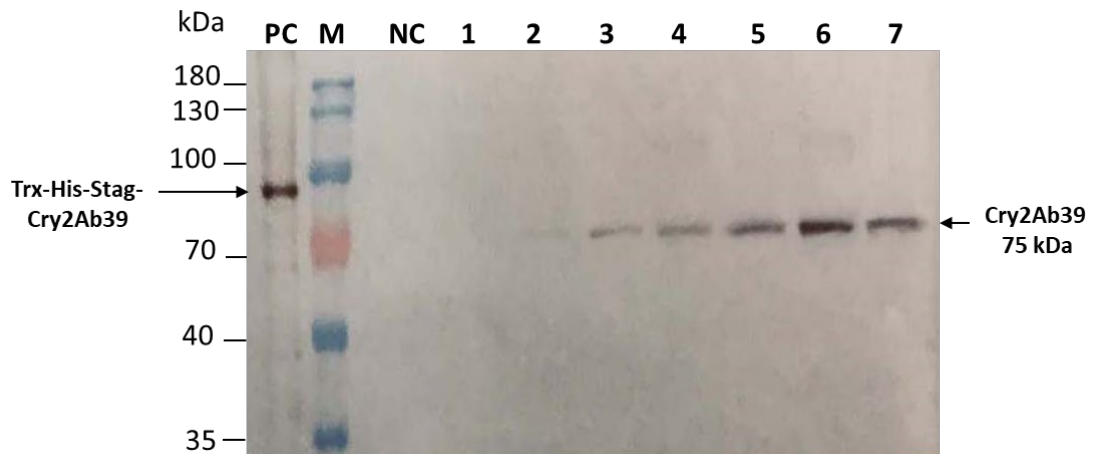
Hình 3.28. Phân tích ảnh hưởng của mật độ khuẩn *Agrobacterium* lây nhiễm đến biểu hiện tạm thời của gen *cry2Ab39opt* trong *N. benthamiana* bằng Western blot sử dụng kháng thể anti-6X His

1 và 2; 3 và 4; 5 và 6; 7 và 8: 5 μ g dịch chiết protein tổng số từ lá cây thuốc lá được xâm nhiễm với dịch khuẩn có OD_{600} lần lượt là 0,1; 0,3; 0,5 và 0,8; PC (Đối chứng dương): protein dung hợp Trx-His-Stag-Cry2Ab39 (khoảng 90 kDa) được biểu hiện trong *E. coli* BL21; M: thang protein 10 -180 kDa (Thermo Scientific).

3.5.3.3. Tối ưu thời gian thu hoạch lá sau lây nhiễm

Thời gian thu hoạch lá sau khi lây nhiễm khuẩn bằng phương pháp agroinfiltration cũng được chứng minh là ảnh hưởng đến hàm lượng protein tái tổ hợp sinh tổng hợp trong lá. Nếu thu lá quá sớm thì sự tích lũy không đạt được tối đa. Ngược lại, nếu thu lá quá muộn, khi đó lá đi vào giai đoạn già hóa và chết do tổn thương của quá trình agroinfiltration dẫn đến sự phân hủy protein đặc biệt là các protein ngoại lai. Trong nghiên cứu này, dịch huyền phù tế bào *A. tumefaciens* AGL1 mang cấu trúc pFGC/rbcS1A-pro_{cry2Ab39} ở mật độ $OD_{600} = 0,5$ được xâm nhiễm vào lá thuốc lá và khảo sát thu lá ở 7 thời điểm khác nhau là 1, 2, 3, 4, 5, 6 và

7 ngày sau khi lây nhiễm. Protein tổng số được tách chiết từ các mẫu lá, điện di trên SDS-PAGE và lai Western blot với kháng thể anti-6X His (Hình 3.29). Kết quả cho thấy sự biểu hiện của Cry2Ab39 tái tổ hợp có thể phát hiện được từ ngày thứ 2 sau biến nạp. Mức độ biểu hiện của protein này tăng dần qua các ngày sau đó và đạt cao nhất vào ngày thứ 6, tuy nhiên đến ngày thứ 7 sau biến nạp, lượng protein tái tổ hợp có sự giảm nhẹ, căn cứ vào độ đậm nhạt của các băng protein trên màng lai. Do đó, 6 ngày sau biến nạp là mốc thời gian tốt nhất cho thu hoạch lá để thu được lượng protein Cry2Ab39 ở mức cao nhất.

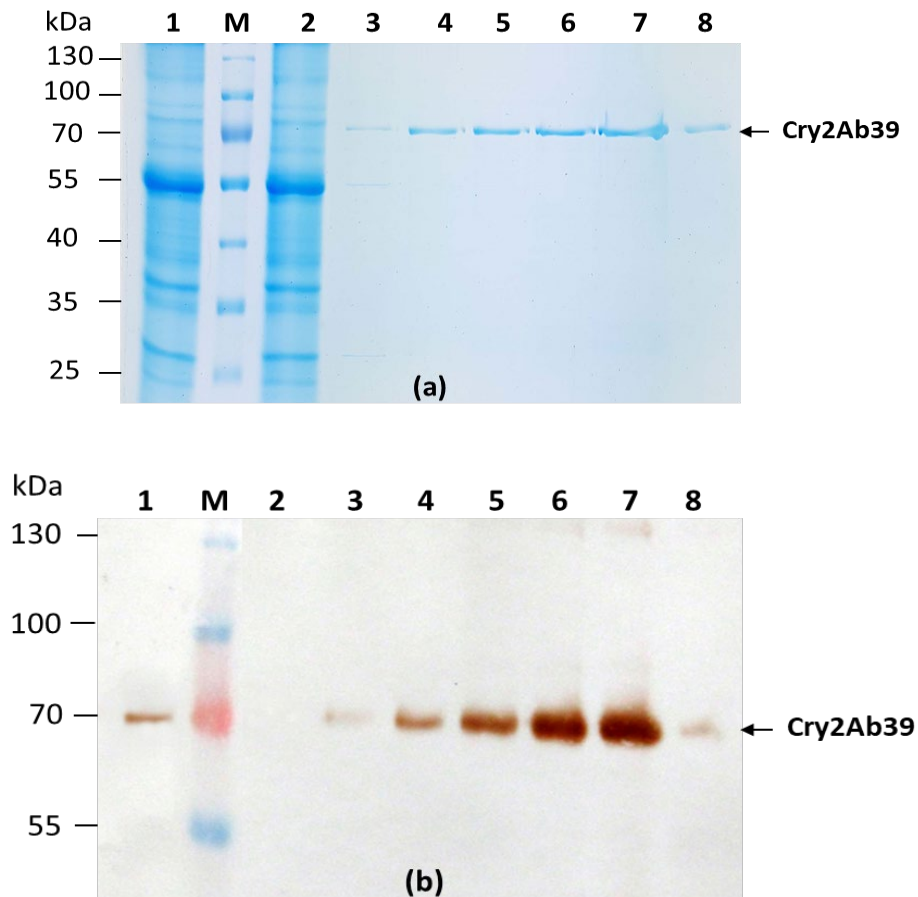


Hình 3.29. Phân tích sự biểu hiện của protein Cry2Ab39 trong lá *N. benthamiana* sau các ngày biến nạp bằng Western blot sử dụng kháng thể anti-6X His 1-7: 10 μ g dịch chiết protein tổng số từ lá thuốc lá thu hoạch ở ngày thứ 1, 2, 3, 4, 5, 6, 7 sau khi biến nạp. NC (mẫu đối chứng âm): dịch chiết protein tổng số từ lá cây không lây nhiễm; PC (Đối chứng dương): protein dung hợp Trx-His-Stag-Cry2Ab39 (khoảng 90 kDa) được biểu hiện trong *E. coli* BL21; M: thang protein 10-180 kDa (Thermo Scientific).

Từ kết quả nghiên cứu nói trên, các điều kiện tối ưu được lựa chọn để biểu hiện tạm thời gen *cry2Ab39opt* trong lá thuốc lá *N. benthamiana* bằng phương pháp agroinfiltration giúp thu được lượng protein tái tổ hợp ở mức cao nhất là: sử dụng cấu trúc chuyên gen được điều khiển bởi promoter *rbcS1A*, mật độ dịch khuẩn cho xâm nhiễm là $OD_{600} = 0,5$, và thu hoạch lá sau 6 ngày biến nạp.

3.5.4. Tinh sạch protein tái tổ hợp Cry2Ab39 từ lá thuốc lá

Trên cơ sở quy trình biểu hiện tạm thời với các điều kiện đã tối ưu, chúng tôi tiến hành biểu hiện tạm thời protein Cry2Ab39 và tách chiết protein chuẩn bị cho bước tinh chế. Do gen *cry2Ab39opt* được thiết kế cho phép biểu hiện protein tái tổ hợp ở dạng dung hợp với hexahistidine ở đầu C, việc tinh sạch Cry2Ab39 có thể thực hiện nhờ phương pháp sắc kí ái lực sử dụng cột có gắn ion Ni²⁺.



Hình 3.30. Phân tích kết quả tinh sạch protein Cry2Ab39 bằng SDS-PAGE (a) và lai miễn dịch Western blot (b).

1: Dịch chiết protein tổng số từ lá thuốc lá được nhiễm khuẩn để biểu hiện tạm thời protein Cry2Ab39; 2: dịch qua cột; 3-8: các phân đoạn thu hồi protein; M: thang protein 10 -180 kDa (Thermo Scientific).

Protein Cry2Ab39 có chứa đuôi hexahistidine có trong dịch chiết protein tổng số từ lá sẽ được giữ lại một cách hiệu quả trên cột. Để loại trừ khả năng bám dính của các protein tạp nhiễm, một lượng nhỏ imidazole được bổ sung vào đệm

gắn và đệm rửa. Cuối cùng, protein đích sẽ được đẩy khỏi cột trong môi trường đệm có chứa nồng độ lớn imidazole đóng vai trò chất cạnh tranh liên kết Ni^{2+} . Protein thu được từ các phân đoạn được phân tích trên SDS-PAGE và lai Western blot sử dụng kháng thể anti-6X His (Hình 3.30). Kết quả cho thấy protein Cry2Ab39 với kích thước khoảng 75 kDa đã được tinh sạch thành công. Protein tái tổ hợp thu được có độ tinh khiết cao, hầu như đã loại bỏ được hoàn toàn các protein tạp có trong dịch chiết thô từ lá. Từ kết quả xác định nồng độ protein sau khi tinh chế, hiệu suất của quá trình thu hồi Cry2Ab39 từ thuốc lá *N. benthamiana* ước tính đạt khoảng 1,5 g protein/ kg lá tươi.

Chương 4. THẢO LUẬN

4.1. Ứng dụng phương pháp giải trình tự thế hệ mới trong tìm kiếm và khai thác các gen độc tố diệt sâu *Bt* mới.

Do khả năng diệt côn trùng phổ rộng của các độc tố *Bt*, thuốc trừ sâu có nguồn gốc *Bt* cũng như cây trồng chuyển gen *Bt* đã được phát triển và ứng dụng rộng rãi để kiểm soát sâu hại trong nông nghiệp, đồng thời giúp giảm việc sử dụng thuốc trừ sâu hóa học trên quy mô toàn cầu [95]. Một ví dụ cụ thể là việc canh tác bông chuyển gen *Bt* trên các cánh đồng ở Ấn Độ đã giúp giảm 70% lượng thuốc trừ sâu cần sử dụng, tiết kiệm tới 30 đô la Mỹ/1 ha và tăng sản lượng bông thu hoạch thêm 80%-87% [96]. Tuy nhiên, mối đe dọa chính đối với việc sử dụng hiệu quả và lâu dài thuốc trừ sâu *Bt* cũng như cây trồng chuyển gen *Bt* là sự tiến triển khả năng kháng thuốc của sâu bệnh. Một thực tế đã được chứng minh rằng côn trùng có thể phát triển khả năng kháng độc tố *Bt* trong phòng thí nghiệm và thuốc trừ sâu *Bt* trên đồng ruộng [97], [98], [99], [100], [101], [102], [103]. Do đó, việc nghiên cứu tìm ra các loại độc tố *Bt* diệt sâu mới, có tính đặc hiệu cao đối với côn trùng đích được coi là một yêu cầu cấp thiết [97], [104], [105]. Trước đây, phương pháp tiếp cận thường được sử dụng trong các nghiên cứu tìm kiếm và phân lập các gen mã hóa protein độc tố từ các chủng *Bt* chưa được mô tả là kỹ thuật PCR [106], [107], [108], [109]. Tuy nhiên, phương pháp này có nhiều hạn chế bao gồm không hiệu quả trong việc phát hiện các gen *Bt* mới, đồng thời khó thu được các gen có trình tự đầy đủ [108], [110]. Kể từ năm 2008 tới nay, các phương pháp giải trình tự gen thế hệ mới (next generation sequencing - NGS) ra đời cho phép giải mã toàn bộ hệ gen (whole-genome sequencing - WGS) một cách nhanh chóng với lượng thông tin lớn và chi phí tiết kiệm đã trở thành công cụ mạnh mẽ và hữu ích giúp các nhà khoa học khám phá thêm nhiều gen độc tố *Bt* mới [104], [110], [111], [112], [113], [114]. Trong nghiên cứu này, chúng tôi đã ứng dụng thành công phương pháp giải trình tự gen thế hệ mới trong việc tìm kiếm và phân lập được gen độc tố mới *cry2Ab39* từ chủng vi khuẩn *Bt* SP14.2 bản địa của Việt Nam. Toàn bộ hệ gen của chủng *Bt serovar canadensis* SP142 đã được giải mã bằng công nghệ giải trình tự gen tức thời đơn

phân tử (single molecular real time sequencing - SMRT) do Pacific Biosciatics (Pac-Bio) phát triển. Dữ liệu bộ gen hoàn chỉnh của chủng này thể hiện các đặc điểm điển hình tương tự như các chủng *Bt* đã được giải mã và công bố, bao gồm kích thước hệ gen, hàm lượng GC và số lượng CDS dự đoán. Việc tiên đoán các gen mã hóa độc tố diệt côn trùng được thực hiện bằng cách sử dụng BtToxin_scanner, một công cụ tin sinh học online thông lượng cao giúp khai thác các gen độc tố có nguồn gốc từ *Bt*, đã giúp chúng tôi xác định được một gen mới tiềm năng thuộc nhóm *cry2Ab*. Kết quả phân lập, nhân dòng và giải trình tự gen quan tâm bằng phương pháp Sanger đã chứng minh tính chính xác của phương pháp giải trình tự hệ gen cũng như tiên đoán gen độc tố chúng tôi đã sử dụng. Gen *cry2Ab* phân lập được từ chủng SP14.2 có độ tương đồng 99,05% về trình tự nucleotide với gen *cry2Ab3* trên Genbank (AF164666.1), trong khi trình tự axit amin suy diễn của chúng có mức tương đồng 99,21% với 5 vị trí sai khác. Với mức độ sai khác đã nêu, độc tố Cry2Ab phát hiện được từ chủng SP14.2 có tính mới được xếp hạng 4. Đáng chú ý là, hầu hết các vị trí axit amin sai khác đều nằm trong vùng chức năng Endotoxin_mid chịu trách nhiệm nhận biết và liên kết với thụ thể đặc hiệu trên màng tế bào ruột giữa của côn trùng đích, cho thấy khả năng cao có sự sai khác về phổ tác động cũng như tính đặc hiệu côn trùng đích của độc tố mới được phát hiện so với độc tố Cry2Ab3 đã được công bố. Trình tự nucleotide của gen mã hóa cũng như trình tự axit amin của protein độc tố này đã được chúng tôi đăng kí trên ngân hàng Genbank với mã truy cập lần lượt là MN319700.1 và QIQ19560.1. Đồng thời, độc tố mới này đã được đặt danh pháp chính thức là Cry2Ab39 bởi Ủy ban danh pháp độc tố *Bt* (nay là Bacterial Pesticidal Protein Resource Center (BPPRC)- <https://www.bpprc.org/>). Chủng SP14.2 mang gen này cũng đã được cấp bằng độc quyền sáng chế bởi Cục sở hữu trí tuệ Việt Nam.

4.2. Ảnh hưởng của nhiệt độ nuôi cấy cảm ứng đến sự biểu hiện protein tái tổ hợp Cry2Ab39 dạng tan trong tế bào *E. coli*.

Như đã đề cập ở trên, nhiều nỗ lực nghiên cứu đã được thực hiện để tìm kiếm các gen độc tố *Bt* mới từ các chủng *B. thuringiensis* tiềm năng. Tuy nhiên, việc xác định được đặc tính phân tử cũng như độc tính của các độc tố mới phát hiện để đưa vào ứng dụng cũng đối mặt với không ít thách thức. Phương pháp tiếp cận

được sử dụng phổ biến nhất là biểu hiện gen *Bt* trong một hệ biểu hiện hiệu suất cao để thu được đủ lượng độc tố tái tổ hợp cho các nghiên cứu tiếp theo. *Escherichia coli* được coi là một hệ thống sản xuất hoàn hảo đối với các protein chức năng không đòi hỏi quá trình glycosyl hóa vì khả năng biểu hiện nhanh, sản xuất quy mô lớn và tiết kiệm chi phí [115]. Tuy nhiên, hiện tượng protein tái tổ hợp sau khi được tổng hợp không được chế biến, cuộn gập tạo cấu trúc không gian chính xác mà có xu hướng tích tụ lại tạo thành dạng thể vùi không tan là một trở ngại lớn cho việc tái cấu trúc nhằm thu được các protein có hoạt tính [116], [117], [118]. Nhiều nghiên cứu đã cho thấy, sự cuộn gập không đúng là nguyên nhân khiến các protein biểu hiện ở dạng thể vùi thường không có hoạt tính sinh học [119]. Theo báo cáo của Hire và cộng sự, protein Cry2Aa14 dạng tan sau khi được tinh chế có độc tính cao với *S. litura* cũng như *C. quinquefasciatus*, trong khi đó, dạng thể vùi của protein này không gây độc cho cả 2 loài côn trùng nói trên ngay cả ở nồng độ cao gấp 5 lần [120]. Cho đến nay, chưa có phương pháp phổ biến nào hiệu quả trong việc ngăn chặn sự hình thành dạng thể vùi khi biểu hiện protein tái tổ hợp trong *E. coli*. Tuy nhiên, các nghiên cứu đã tìm ra một số chiến lược được chứng minh là hữu ích giúp cải thiện khả năng tan của protein đích được biểu hiện trong *E. coli* bao gồm: tối ưu các điều kiện nuôi cấy cũng như kiểm soát tốc độ sinh trưởng của tế bào vi khuẩn [121], biểu hiện ở dạng dung hợp với một số protein phân tử nhỏ [122], [123], hoặc đồng biểu hiện với các phân tử protein chaperone và foldase giúp hỗ trợ quá trình lắp ráp, cuộn gập protein [124], [125]... Trong nghiên cứu này, chúng tôi đã thử nghiệm hai yếu tố nhằm tăng khả năng biểu hiện ở pha tan của protein Cry2Ab39 tái tổ hợp là biểu hiện dung hợp với thioredoxin (Trx) và giảm tốc độ tổng hợp protein. Các hệ thống biểu hiện dung hợp gen dựa trên Trx đã được chứng minh là một trong những công cụ lý tưởng để tránh sự hình thành thể vùi trong quá trình sản xuất protein ở *E. coli coli* [126]. Hơn nữa, phân tử protein Trx đã được chứng minh là không làm ảnh hưởng đến hoạt tính sinh học của các protein tái tổ hợp [120]. Do đó, với mục đích thu được protein mục tiêu ở dạng hòa tan, cấu trúc biểu hiện gen đích dung hợp với *trx* đã được thiết kế bằng cách ghép nối gen *cry2Ab39* vào vector pET-32a(+). Cấu trúc này cho phép biểu hiện protein tái tổ hợp Trx-His-Stag-Cry2Ab39, trong đó việc gắn thêm đuôi His tag để phục vụ nghiên cứu tinh chế,

còn hai thành phần Trx và S tag đóng vai trò tăng cường độ hòa tan và tính ổn định cho protein tái tổ hợp. Mặc dù Trx-His-Stag-Cry2Ab39 được biểu hiện ở mức cao trong tế bào *E. coli* BL21 như mong đợi, nhưng chỉ một lượng rất nhỏ protein dung hợp này tồn tại ở pha tan, còn lại phần lớn tích tụ ở dạng thể vùi trong điều kiện cảm ứng ở 37°C. Ở đây, nguyên nhân dẫn đến sự hình thành ưu thế của thể vùi được nhận định là do sự sản xuất quá mức protein tái tổ hợp dẫn đến sự mất cân bằng giữa tốc độ tổng hợp và tốc độ gấp cuộn protein của tế bào. Để giải quyết vấn đề này, trạng thái cân bằng giữa hai quá trình trên đã được thiết lập lại bằng cách giảm tốc độ sinh trưởng của tế bào vật chủ thông qua việc giảm nhiệt độ nuôi cấy cảm ứng. Kết quả tối ưu cho thấy, khi hạ nhiệt độ cảm ứng xuống 4°C, lượng Trx-His-Stag-Cry2Ab39 ở dạng tan thu được tăng rõ rệt (chiếm gần 50% tổng sinh khối protein dung hợp được tạo ra) (Hình 3.17). Điều này cho thấy, việc biểu hiện ở mức nhiệt thấp đã giúp cải thiện khả năng tan của protein Cry2Ab39 tái tổ hợp. Kết quả này phù hợp với công bố của Hire và cộng sự khi nghiên cứu biểu hiện độc tố Cry2Aa14 ở dạng đơn hoặc dạng dung hợp với Trx. Theo đó, lượng protein tái tổ hợp Cry2Ab14 cũng như Trx-Cry2Ab14 dạng tan và có hoạt tính sinh học thu được nhiều hơn trong điều kiện nhiệt độ cảm ứng biểu hiện thấp (25°C) so với khi biểu hiện ở mức nhiệt 37°C [120]. Tuy nhiên, theo quan sát hình ảnh điện di SDS-PAGE, mức nhiệt cảm ứng 25°C mới chỉ giúp cải thiện một phần nhỏ lượng protein tái tổ hợp dạng tan mong muốn, còn lại phần lớn vẫn nằm ở pha không tan. Trong nghiên cứu của chúng tôi, nhiệt độ cảm ứng khi được hạ xuống mức cực kỳ thấp (4°C) đã giúp tế bào vi khuẩn tổng hợp được một lượng lớn protein Trx-His-Stag-Cry2Ab39 dạng tan (chiếm gần 50% sinh khối protein tái tổ hợp tổng số). Như vậy, việc biểu hiện gen *cry2Ab39* trong *E. coli* BL21 ở mức nhiệt độ cảm ứng cực kỳ thấp đã làm giảm sự hình thành và tích lũy protein tái tổ hợp dạng thể vùi. Có thể thấy đây là một chiến lược tiềm năng cho các nghiên cứu biểu hiện gen độc tố *Bt* nói chung trong hệ biểu hiện *E. coli*. Việc sử dụng điều kiện cảm ứng ở nhiệt độ thấp khi biểu hiện các loại protein khác như enzyme cũng đã được báo cáo bởi San-Miguel và cộng sự. Cụ thể, khi biểu hiện *P5βR2* mã hóa cho progesterone 5β-reductase phân lập từ *Digitalis* trong tế bào vi khuẩn *E. coli*, nhóm tác giả thấy rằng, việc protein enzyme tái tổ hợp thu được có chức năng sinh học hay không hoàn toàn phụ thuộc

vào nhiệt độ cảm ứng. Chỉ những tế bào được cảm ứng ở các mức nhiệt độ thấp (15°C hoặc 4°C) mới tạo ra được enzyme P5βR2 dạng hoạt động. Trong khi đó, khi biểu hiện ở 37°C hoặc 25°C, protein tái tổ hợp thu được hoàn toàn không có hoạt tính [127].

4.3. Cải biến mã di truyền gen *cry2Ab39* có nguồn gốc từ vi khuẩn để phù hợp với hệ biểu hiện của thực vật

Theo kết quả công bố của nhiều nghiên cứu trước đây, các gen *cry* của vi khuẩn *Bt* khi được chuyển vào cây trồng đều có mức độ biểu hiện rất thấp và thường không thể giúp cây chủ có khả năng kháng lại các loài côn trùng gây hại [128]. Lý do chính dẫn đến sự biểu hiện ở mức độ cực kỳ thấp của các gen *cry* dạng đại trong tế bào thực vật là do các bản phiên mã mRNA của chúng không bền và bị phân hủy một cách nhanh chóng. Nguyên nhân mất ổn định của các phân tử mRNA này là do sự khác biệt rõ rệt trong thành phần DNA giữa gen *cry* của vi khuẩn và các gen của thực vật. Cụ thể, tỷ lệ A-T trong các gen *cry* dạng đại cao hơn nhiều so với gen thực vật điển hình. Các vùng giàu A-T này nhiều khả năng được tế bào thực vật nhận dạng như các vùng intron hoặc các tín hiệu poly A, dẫn đến việc phân cắt sợi mRNA mục tiêu hoặc gây ra sự kết thúc sớm quá trình dịch mã [129], [130], [131], [132]. Ngoài ra, sự khác biệt trong việc ưu tiên sử dụng các bộ ba mã hóa giữa hệ thống biểu hiện của thực vật và vi khuẩn cũng góp phần dẫn đến hiệu quả dịch mã thấp [131], [132]. Chính vì vậy, trong nghiên cứu này, nhằm tăng cường khả năng biểu hiện của gen *cry2Ab39* trong thuốc lá *N. benthamiana*, năm chiến lược tối ưu hóa đã được sử dụng bao gồm: (1) loại bỏ các bộ ba mã hiếm đối với thuốc lá bằng cách thay thế bằng các bộ mã vật chủ ưu tiên sử dụng; (2) tăng tần số các bộ mã giàu GC nhằm tăng hàm lượng GC tổng số của toàn bộ gen; (3) loại trừ các vị trí tiềm ẩn nguy cơ gây ra hiện tượng cắt nối mRNA hoặc các trình tự giống tín hiệu poly A ở thực vật; (4) tránh sự hình thành cấu trúc thứ cấp của phân tử mRNA hoặc các motif mRNA kém ổn định; (5) vận chuyển và giữ protein tái tổ hợp ở lưới nội chất ER bằng cách thêm trình tự peptide tín hiệu LeB4 vào đầu N và trình tự KDEL ở đầu C. Kết quả biểu hiện và tinh chế cho thấy mức độ biểu hiện tạm thời của protein Cry2Ab39 tái tổ hợp trong lá thuốc lá *N. benthamiana* là hơn 1,5 μg protein/g lá, tương đương hoặc cao hơn so với các báo cáo về mức độ biểu

hiện của các gen *cry* khác đã được cải biến mã khi được chuyển vào thuốc lá hay các đối tượng cây trồng khác [133], [134], [135], cho thấy chiến lược tối ưu hóa được sử dụng trong nghiên cứu đã mang lại hiệu quả.

4.4. Lựa chọn promoter phù hợp để biểu hiện gen *cry2Ab39opt* trong lá thuốc lá

Việc điều khiển sự phiên mã của gen chuyển trong tế bào vật chủ là một trong những bước quan trọng, góp phần quyết định đến hiệu quả tạo ra protein tái tổ hợp. Promoter là yếu tố điều hòa chính chịu trách nhiệm kiểm soát trực tiếp quá trình phiên mã của các gen nói chung. Người ta đã chứng minh việc sử dụng các promoter khác nhau có thể dẫn đến sự sai khác về mức độ biểu hiện protein đích trong tế bào chủ lên đến 80% [136]. Vì vậy, lựa chọn được một promoter phù hợp là hết sức cần thiết để việc biểu hiện gen đạt được mức tối ưu [137]. Trong nghiên cứu này, ảnh hưởng của hai promoter 35S (có nguồn gốc từ Cauliflower Mosaic Virus-CaMV) và *rbcS1A* (phân lập từ *Arabidopsis*) đến sự biểu hiện tạm thời của gen *cry2Ab39opt* trong mô lá thuốc lá đã được phân tích và so sánh. CaMV 35S được xem là một promoter mạnh, thuộc loại cơ định, thường được sử dụng khá phổ biến để biểu hiện protein tái tổ hợp ở mức độ cao trong tế bào thực vật [138]. Trong khi đó, promoter *rbcS1A* thuộc nhóm promoter đặc hiệu mô, được nhiều nghiên cứu sử dụng trong việc biểu hiện một cách đặc hiệu gen chuyển trong lá và quả giai đoạn còn xanh [139], [140]. Kết quả bán định lượng protein tái tổ hợp bằng Western blot cho thấy hàm lượng Cry2Ab39 trong lá thuốc lá chuyển gen với cấu trúc được kiểm soát bởi promoter *rbcS1A* cao hơn đáng kể so với mẫu lá chuyển gen được điều khiển bởi promoter CaMV 35S (Hình 3.27), chứng tỏ hiệu quả vượt trội của promoter *rbcS1A* đối với biểu hiện tạm thời protein đích trong lá *N. benthamiana*. Cùng với kết quả của những công bố khác, nghiên cứu của chúng tôi chứng minh một thực tế là các promoter mạnh không phải lúc nào cũng là sự lựa chọn tốt nhất cho sự biểu hiện ở mức cao các protein tái tổ hợp. Trong một số trường hợp, việc sử dụng promoter mạnh để điều khiển sự phiên mã của gen chuyển không chỉ dẫn đến sự tích lũy một lượng lớn mRNA của gen mục tiêu trong tế bào chất mà còn làm tăng số lượng RNA bất thường và kích hoạt quá trình làm câm gen của tế bào chủ, dẫn đến hậu quả không mong muốn là giảm hiệu quả biểu hiện gen [141], [142].

Hơn thế nữa, kết quả của nghiên cứu cũng góp phần cho thấy việc sử dụng promoter *rbcS1A* khi cần biểu hiện đặc hiệu gen đích ở các mô xanh của thực vật là một hướng tiếp cận hiệu quả giúp cải thiện mức độ biểu hiện của gen chuyển [143], [144], [145].

4.5. Điều kiện tối ưu để biểu hiện tạm thời Cry2Ab39 trong lá thuốc lá

Agroinfiltration là một kỹ thuật đơn giản nhưng hiệu quả giúp đưa gen ngoại lai vào tế bào thực vật và nhanh chóng tạo ra protein tái tổ hợp [146]. Mặc dù phương pháp này đã được sử dụng rộng rãi trên cây thuốc lá *N. benthamiana* [147], [148], [149], [150], tuy nhiên tùy thuộc vào từng loại gen chuyển mà một số điều kiện cần được xem xét và điều chỉnh để tăng khả năng lây nhiễm *Agrobacterium* vào mô lá cũng như tăng lượng protein tái tổ hợp thu được. Trong nghiên cứu này, hai yếu tố bao gồm mật độ của tế bào vi khuẩn dùng để lây nhiễm và thời gian thu hoạch lá sau khi nhiễm khuẩn đã được chứng minh là có ảnh hưởng đáng kể đến sự biểu hiện của gen *cry2Ab39opt*. Theo nhiều công bố trước đây, mật độ vi khuẩn sử dụng trong thí nghiệm agroinfiltration là một trong những yếu tố ảnh hưởng lớn đến hiệu quả chuyển gen cũng như sự biểu hiện tạm thời của gen chuyển [146], [151], [152]. Nếu dịch khuẩn bị pha loãng quá mức có thể dẫn đến tỷ lệ tế bào vi khuẩn/tế bào thực vật thấp, làm giảm hiệu suất chuyển gen. Ngược lại, nếu mật độ vi khuẩn quá đậm đặc có thể gây ra sự phát triển quá mức của vi khuẩn trong mô lá dẫn đến sự tổn thương nghiêm trọng cho cây chủ [146], [153]. Trong các điều kiện OD₆₀₀ của dịch huyền phù *Agrobacterium* được thử nghiệm để lây nhiễm, mức biểu hiện cao nhất của protein Cry2Ab39 được quan sát thấy ở OD₆₀₀ = 0,5 (Hình 3.28). Mật độ tế bào *Agrobacterium* ở mức này cũng được nhiều nhóm nghiên cứu lựa chọn trong thí nghiệm agroinfiltration trên đối tượng *N. benthamiana* để thu được sự biểu hiện ở mức cao của protein tái tổ hợp [154], [155], [156]. Ngoài mật độ vi khuẩn, thời gian thu hoạch lá sau khi lây nhiễm cũng có ảnh hưởng nhất định đến mức độ biểu hiện của protein mong muốn [157]. Theo đó, nên thu mẫu lá đã lây nhiễm *Agrobacterium* ở thời điểm mô lá chưa xuất hiện các dấu hiệu tổn thương nặng gây ra do sự phát triển của vi khuẩn như lá vàng úa, mất lục lạp, lá héo quá mức hoặc bị hoại tử. Các tổn thương này có thể ảnh hưởng đến khả năng tạo năng lượng và do đó làm giảm hiệu suất tổng hợp protein của tế bào chủ. Kết quả ghi nhận được trong

thí nghiệm của chúng tôi cho thấy thời điểm 6 ngày sau khi lây nhiễm là tốt nhất để thu mẫu lá và chiết protein, giúp nhận được nhiều nhất lượng protein tái tổ hợp Cry2Ab39 (Hình 3.29). Thời gian thu hoạch lá trong thí nghiệm biểu hiện tạm thời theo các nghiên cứu đã công bố có sự khác nhau phụ thuộc vào loại cây kí chủ cho biểu hiện, chủng *Agrobacterium* được sử dụng để chuyển gen cũng như gen ngoại lai mong muốn được đưa vào hệ biểu hiện thực vật [146], [158].

4.6. Tiềm năng ứng dụng của protein độc tố Cry2Ab39

Việc nghiên cứu và tìm ra hàng trăm gen độc tố từ vi khuẩn *Bt* có khả năng kháng côn trùng gây hại kết hợp với những tiến bộ vượt nhanh chóng trong công nghệ gen đã thúc đẩy việc tạo ra cây trồng chuyển gen *Bt* trên phạm vi toàn cầu [60], [159]. Trong thời kỳ đầu, các gen độc tố *cry1* là nguồn gen phổ biến cho các nghiên cứu chuyển gen tạo ra thế hệ cây trồng *Bt* đầu tiên. Tuy nhiên, do thời gian lưu hành lâu dài cũng như quy mô canh tác lớn, một số quần thể côn trùng đích hiện nay đã bắt đầu xuất hiện hiện tượng tiến hóa khả năng kháng độc để trốn thoát và tồn tại [160], [161], [162]. Vì vậy, hiện nay các độc tố *Bt* thuộc nhóm Cry2 đang được xem là lựa chọn thay thế cho protein Cry1 để phát triển cây trồng chuyển gen kháng côn trùng thế hệ thứ hai [120]. Độc tố thuộc nhóm này được xác định có khả năng diệt côn trùng thuộc bộ Cánh vảy và Hai cánh [163]. Đặc biệt, các protein Cry2A đã được chứng minh có thể gây độc cho một số sâu hại thuộc bộ Cánh vảy mà các protein Cry1 không diệt được [164]. Hơn nữa, với việc nhận diện các thụ thể trên màng tế bào ruột giữa của côn trùng khác với thụ thể của Cry1, các độc tố nhóm Cry2 cho thấy khả năng kháng chéo của các côn trùng kháng Cry1 là rất thấp [165], [166]. Ngoài ra, do kích thước phân tử nhỏ hơn so với gen *cry1*, các gen mã hóa Cry2 được cho là có lợi thế lớn hơn trong việc chuyển gen cũng như biểu hiện ở thực vật [120]. Vì những lý do trên, việc nghiên cứu tìm ra các protein Cry2 mới có độc lực cao và phổ hoạt tính rộng là cần thiết, không chỉ để tiếp tục phát triển cây trồng *Bt* giúp kiểm soát sâu hại mà còn quản lý và giải quyết được tình trạng kháng độc của côn trùng.

Trong nghiên cứu này, một độc tố mới thuộc nhóm Cry2A là Cry2Ab39 có nguồn gốc từ chủng *Bt Serovar canadensis* SP14.2 bản địa của Việt Nam đã được phát hiện và mô tả các đặc trưng về mặt phân tử cũng như đánh giá hoạt tính sinh

học. Kết quả thử nghiệm đã chứng minh protein Cry2Ab39 tái tổ hợp gây độc cho sâu non *Etiella zinckenella* Treitschke – một loài sâu đục quả gây hại cho các cây họ đậu, với liều gây chết trung bình $LC_{50} = 1,74 \mu\text{g protein/g}$ thức ăn. Đây là nghiên cứu thứ hai trên thế giới được công bố liên quan đến độc tố *Bt* có khả năng diệt *E. zinckenella* Treitschke [24], [167]. Tuy nhiên, nghiên cứu này là công bố đầu tiên phân lập được protein độc tố mới có độc lực cao với loài côn trùng này. *E. zinckenella* Treitschke thuộc bộ Lepidoptera, họ Pyralidae và phân bố rộng rãi trên khắp thế giới, từ miền nam và miền đông châu Âu, Bắc Mỹ và Úc đến vùng nhiệt đới và cận nhiệt đới của Châu Phi và Châu Á. Các biện pháp phòng trừ bằng thuốc trừ sâu hóa học không hiệu quả vì ấu trùng *E. zinckenella* ăn bên trong vỏ quả đậu khiến thuốc khó tiếp cận. Do đó, gen *cry2Ab39* có tiềm năng ứng dụng trong nghiên cứu tạo các cây họ đậu chuyển gen giúp kiểm soát loại sâu hại này. Ngoài ra, cần có các nghiên cứu tiếp theo nhằm đánh giá đầy đủ phổ diệt côn trùng cũng như độc lực của độc tố mới này với các loài côn trùng khác để có hướng ứng dụng hiệu quả.

KẾT LUẬN VÀ KIẾN NGHỊ

KẾT LUẬN

1. 221 chủng *Bt* được hoạt hóa, nhân nuôi, thu sinh khối và thử nghiệm hoạt tính diệt sâu đục quả đậu tương *E. zinckenella* Treitschke. Lựa chọn được 4 chủng là SP14.2, Đ6.1, Đ6.2, BD8.2 có hoạt lực diệt sâu cao (trên 85%).
2. Hệ gen 4 chủng SP14.2, Đ6.1, Đ6.2 và BD8.2 đã được giải trình tự, lắp ráp và chú giải với kích thước 6,8 – 7,1 Mb. Khai thác và lựa chọn được 6 gen mới tiềm năng trong diệt trừ sâu hại đậu tương *E. zinckenella*, gồm có *cry1Aa*, *cry1Ac*, *cry1Be*, *cry1Na*, *cry2Ab* và *cry2Ah*.
3. 4 gen mã hóa protein độc tố có tính mới là *cry1Be*, *cry1Na*, *cry2Ab* và *cry2Ah* đã được phân lập và nhân dòng. Các protein Cry tái tổ hợp được biểu hiện trong tế bào *E. coli* có khả năng diệt sâu đục quả đậu tương *E. zinckenella* từ 66,67% - 96,67%, trong đó độc tố Cry2Ab có hoạt lực mạnh nhất (96,67%).
4. Protein tái tổ hợp Cry2Ab đã được biểu hiện ở dạng tan trong tế bào *E. coli*. Protein Cry2Ab tinh sạch gây độc cho sâu non tuổi 2 với liều gây chết trung bình LC50 = 1,74 µg protein/g thức ăn. Trình tự nucleotide và axit amin của gen *cry2Ab* đã được đăng kí trên ngân hàng GenBank với mã truy cập lần lượt là MN319700.1 và QIQ19560.1. Độc tố mới phát hiện được xét duyệt và đặt tên là Cry2Ab39 bởi Ủy ban danh pháp nội độc tố *Bt*. Ngoài ra, chủng *Bt* SP14.2 mang gen mã hóa độc tố này đã được cấp bằng độc quyền sáng chế bởi Cục sở hữu trí tuệ Việt Nam.
5. Gen *cry2Ab39* đã được cải biến mã để phù hợp với hệ biểu hiện của thuốc lá. Xây dựng thành công quy trình biểu hiện tạm thời protein Cry2Ab39 trong lá thuốc lá *N. benthamiana* bằng phương pháp agroinfiltration với các điều kiện tối ưu. Protein tái tổ hợp Cry2Ab39 được biểu hiện và tinh sạch từ lá thuốc lá với hiệu suất thu hồi là 1,5 g protein/ kg lá tươi.

KIẾN NGHỊ

Trên cơ sở các kết quả đạt được, để có thể tiếp tục phát triển các nghiên cứu tiếp theo, chúng tôi đề xuất một số kiến nghị sau:

1. Tiếp tục nghiên cứu khả năng biểu hiện gen *cry2Ab39* trên cây đậu tương làm cơ sở cho nghiên cứu chuyển gen tạo giống đậu tương có khả năng kháng sâu đục quả.
2. Thực hiện các nghiên cứu tiếp theo nhằm đánh giá đầy đủ phổ diệt côn trùng cũng như độc lực của độc tố mới Cry2Ab39 với các loài côn trùng đích khác để có hướng ứng dụng hiệu quả.

DANH MỤC CÔNG TRÌNH CÔNG BỐ LIÊN QUAN ĐẾN LUẬN ÁN

1. **Le Thu Ngọc**, Le Thi Minh Thanh, Pham Bich Ngoc, Trinh Thi Thu Ha, Dong Van Quyen, Ngo Dinh Binh, Chu Hoang Ha, Hoang Ha, Nguyen Van Dong (2022) Detection of a novel Cry2Ab toxin against *Etiella zinckenella* Treitschke from the *Bacillus thuringiensis* serovar *canadensis* SP142 strain. *Plant Prot Sci* 58:158–169. <https://doi.org/10.17221/59/2021-PPS>
2. **Le Thu Ngọc**, Tran Thi Huyen, Trinh Thai Vy, Nguyen Thi Tra, Chu Hoang Ha, Do Tien Phat, Le Thi Minh Thanh, Pham Bich Ngoc (2022) Transient Expression of plant-codon-optimized *cry2ab39* gene by agroinfiltration in *N. benthamiana*. *J Mod Agric Biotechnol* 1(2): 11. DOI: 10.53964/jmab.2022011.
3. Lê Thị Minh Thành, Trịnh Thị Thu Hà, Ngô Đình Bình, Chu Hoàng Hà, Phạm Bích Ngọc, **Lê Thu Ngọc** (2020). Chứng vi khuẩn *Bacillus thuringiensis* serovar *canadensis* SP14.2 thuần khiết về mặt sinh học mang gen mã hóa protein độc tố Cry2Ab39 diệt sâu đục quả đậu tương. Bằng độc quyền sáng chế số 37733 được Cục sở hữu trí tuệ Việt Nam cấp ngày 30/10/2023.

DANH MỤC TÀI LIỆU THAM KHẢO

1. Zhang B., Chen P., Florez-Palacios S.L., Shi A., Hou A. & Ishibashi T., 2010, Seed quality attributes of food-grade soybeans from the U.S. and Asia, *Euphytica*, 173, pp. 387–396.
2. De Maria M., Robinson E., Kangile J., Kadigi R., Dreoni I., Couto M., 2020, Global soybean trade - the geopolitics of a bean, *UKRI GCRF TRADE Hub*. doi: 10.34892/7yn1-k494
3. Ngô Thế Dân, Trần Đình Long, Trần Văn Lại, 1999, *Cây đậu tương*, NXB Nông nghiệp Hà Nội, tr. 128-130.
4. Montoya F., García C., Pintos F. & Otero A., 2017, Effects of irrigation regime on the growth and yield of irrigated soybean in temperate humid climatic conditions, *Agricultural Water Management*, 193, pp. 30–45.
5. USDA, 2020, World agricultural production [Online]. Available: <https://apps.fas.usda.gov/PSDOnline/Circulars/2020/01/production.pdf>
6. USDA, 2022, World agricultural production [Online]. Available: <https://apps.fas.usda.gov/PSDOnline/Circulars/2022/01/production.pdf>”
7. USDA, 2024, World agricultural production [Online]. Available: <https://apps.fas.usda.gov/PSDOnline/Circulars/2024/01/production.pdf>”
8. Tổng cục thống kê Việt Nam, Niên giám thống kê tóm tắt -2023 [Online]. <http://www.gso.gov.vn/du-lieu-va-so-lieu-thong-ke/2023/07/niem-giam-thong-ke-tom-tat-2023/>
9. Tổng cục thống kê Việt Nam, Niên giám thống kê tóm tắt -2020 [Online]. <http://www.gso.gov.vn/du-lieu-va-so-lieu-thong-ke/2020/07/niem-giam-thong-ke-tom-tat-2020/>
10. Metcalf R.L., 1996, Applied entomology in the twenty-first century: needs and prospects, *American entomologist*, 42(4), pp. 216-227.
11. Dhaliwal G.S., Jindal V., Mohindru B., 2015, Crop Losses due to insect pests: Global and Indian Scenario, *Indian Journal of Entomology*, 77, pp. 165. <https://doi.org/10.5958/0974-8172.2015.00033.4>
12. Taghizadeh R., Talebi A., Fathipour Y., Khalghani J., 2012, Effect of ten soybean cultivars on development and reproduction of lima bean pod borer, *Etiella zinckenella* (Lepidoptera: Pyralidae) under laboratory conditions, *Applied Entomology and Phytopathology*, 79, pp. 15–28.
13. Abd El-Rahman S., Abdel-Wahab E., 2020, Efficiency of certain bio-insecticides for reducing the yield losses due to the bean pod borer, *Etiella*

- zinckenella* (Treitschke) in soybean fields, *Journal of Plant Protection and Pathology*, 11(1), pp. 29-36.
14. Ginting S., Pujiwati H., Djoko U., Murcitra B., Susilo E., 2022, The attack of *Etiella zinckenella* Treitschke on soybean varieties, *Jurnal Hama dan Penyakit Tumbuhan Tropika*, 22, pp. 83–89.
 15. Tohamy H.T., El-Hafez G.A., 2005, Integrated crop management system for controlling cowpea pod worm, *Etiella zinckenella* (Treit.) in relation to soybean yield at Minia and New valley regions. *Egyptian Journal of Agricultural Research*, 83, pp. 1079–1098.
 16. Schlickmann Tank J., Stefani L., Burgos-Cantoni C., Mongelós-Franco J., Enciso-Maldonado G., 2023, Record of *Etiella zinckenella*(Treitschke) affecting soybeans in Itapúa, Paraguay, *Revista de la Sociedad Científica del Paraguay*, 28, pp. 269–279.
 17. Tabata J., Yasuda K., 2011, The sex pheromone of the lima bean pod borer *Etiella zinckenella* (Treitschke), *Japan Agricultural Research Quarterly*, 45, pp. 47–50.
 18. Nguyễn Phước Đăng, Thái Kim Tuyền, 2017, Đánh giá phản ứng của chín dòng đậu nành đối với sâu đục trái (*Etiella zinckenella*) tại Cần Thơ ở vụ Xuân Hè 2015, *Tạp chí Khoa học Đại học Cần Thơ*, tr. 46–53.
 19. Ahmad M., Sayyed A.H., Saleem M.A., Ahmad M., 2008, Evidence for field evolved resistance to newer insecticides in *Spodoptera litura* (Lepidoptera: Noctuidae) from Pakistan. *Crop Protection*, 27, pp. 1367–1372.
 20. Huang S., Han Z., 2007, Mechanisms for multiple resistances in field populations of common cutworm, *Spodoptera litura* (Fabricius) in China. *Pesticide Biochemistry and Physiology*, 87, pp. 14–22.
 21. Abdel-Wahab E., 2019, Potential of some soybean varieties for resistance to lima bean pod borer (*Etiella zinckenella*) under field conditions, *Research on Crops*, 20, pp. 389–398.
 22. Kuswantoro H., Bayu M.S.Y.I., Baliadi Y., Tengkan W., 2017, Resistance of advanced soybean lines to pod borer (*Etiella zinckenella*), *Biosaintifika: Journal of Biology & Biology Education*, 9, pp. 317–324.
 23. Martins-Salles S., Machado V., Massochin-Pinto L., Fiuza L.M., 2017, Genetically modified soybean expressing insecticidal protein (Cry1Ac): Management risk and perspectives. *FACETS*, 2(1), pp. 496-512.
 24. Hang P.L.B., Linh N.N., Ha N.H., Dong N.V., Hien L.T.T., 2021, Genome sequence of a Vietnamese *Bacillus thuringiensis* strain TH19 reveals two

- potential insecticidal crystal proteins against *Etiella zinckenella* larvae. *Biological Control* 152, pp. 104473.
25. Jensen G.B., Hansen B.M., Eilenberg J., Mahillon J., 2003, The hidden lifestyles of *Bacillus cereus* and relatives, *Environmental Microbiology*, 5, pp. 631–640.
 26. Syed T., Askari M., Meng Z., Li Y., Abid M.A., Wei Y., 2020, Current insights on vegetative insecticidal proteins (Vip) as next generation pest killers, *Toxins*, 12. <https://doi.org/10.3390/toxins12080522>
 27. Palma L., Muñoz D., Berry C., Murillo J., Caballero P., 2014, *Bacillus thuringiensis* toxins: an overview of their biocidal activity, *Toxins*, 6, pp. 3296–3325.
 28. Höfte H., Whiteley H.R., 1989, Insecticidal crystal proteins of *Bacillus thuringiensis*, *Microbiological Reviews*, 53, pp. 242–255.
 29. Loutfi H., Fayad N., Pellen F., Le Jeune B., Chakroun M., Benfarhat D., 2021, Morphological study of *Bacillus thuringiensis* crystals and spores, *Applied Sciences*, 11, pp. 155. <https://doi.org/10.3390/app11010155>
 30. Crickmore N., Zeigler D.R., Feitelson J., Schnepf E., Van Rie J., Lereclus D., 1998, Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins, *Microbiology and molecular biology reviews: MMBR*, 62, pp. 807–813.
 31. Crickmore N., Berry C., Panneerselvam S., Mishra R., Connor T.R., Bonning B.C., 2020, A structure-based nomenclature for *Bacillus thuringiensis* and other bacteria-derived pesticidal proteins, *Journal of Invertebrate Pathology*, 186, pp. 107438. <https://doi.org/10.1016/j.jip.2020.107438>
 32. Chougule N.P., Bonning B.C., 2012), Toxins for transgenic resistance to hemipteran pests, *Toxins*, 4, pp. 405–429.
 33. Xie Q.D., Huang T., Wang X.M., Salem H.H., Ali B., Zhang X.Y., 2010, Effect of *Bacillus thuringiensis* var. *israelensis* endotoxin on the intermediate snail host of *Schistosoma japonicum*, *Current Research in Bacteriology*, 3(1), pp. 37-41.
 34. van Frankenhuyzen K., 2013, Cross-order and cross-phylum activity of *Bacillus thuringiensis* pesticidal proteins, *Journal of Invertebrate Pathology*, 114, pp.76–85.
 35. van Frankenhuyzen K., 2009, Insecticidal activity of *Bacillus thuringiensis* crystal proteins, *Journal of Invertebrate Pathology*, 101, pp. 1–16.

36. Ohba M., Mizuki E., Uemori A., 2009, Parasporin, a new anticancer protein group from *Bacillus thuringiensis*, *Anticancer Research*, 29, pp. 427–433.
37. Xu C., Wang B.-C., Yu Z., Sun M., 2014, Structural Insights into *Bacillus thuringiensis* Cry, Cyt and parasporin toxins, *Toxins*, 6, pp. 2732–2770.
38. Ben-Dov E., 2014, *Bacillus thuringiensis subsp. israelensis* and its dipteran-specific toxins, *Toxins*, 6, pp. 1222–1243.
39. Bravo A., Gill S.S., Soberón M., 2007, Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control, *Toxicon: Official Journal of the International Society on Toxinology*, 49, pp. 423–435.
40. Cohen S., Albeck S., Ben-Dov E., Cahan R., Firer M., Zaritsky A., 2011, Cyt1Aa toxin: crystal structure reveals implications for its membrane-perforating function, *Journal of Molecular Biology*, 413, pp. 804–814.
41. Guerchicoff A., Delécluse A., Rubinstein C.P., 2001, The *Bacillus thuringiensis* cyt genes for hemolytic endotoxins constitute a gene family, *Applied and Environmental Microbiology*, 67, pp. 1090–1096.
42. Soberón M., López-Díaz J.A., Bravo A., 2013, Cyt toxins produced by *Bacillus thuringiensis*: A protein fold conserved in several pathogenic microorganisms, *Peptides*, 41, pp. 87–93.
43. Yu X., Liu T., Sun Z., Guan P., Zhu J., Wang S., 2012, Co-expression and synergism analysis of Vip3Aa29 and Cyt2Aa3 insecticidal proteins from *Bacillus thuringiensis*, *Current Microbiology*, 64, pp. 326–331.
44. Wirth M.C., Delécluse A., Walton W.E., 2001, Cyt1Ab1 and Cyt2Ba1 from *Bacillus thuringiensis subsp. medellin* and *B. thuringiensis subsp. israelensis* Synergize *Bacillus sphaericus* against *Aedes aegypti* and resistant *Culex quinquefasciatus* (Diptera: Culicidae), *Applied and Environmental Microbiology*, 67, pp. 3280–3284.
45. De Maagd R., Bravo A., Berry C., Crickmore N., Schnepf H., 2003, Structure, diversity, and evolution of protein toxins from spore-forming entomopathogenic bacteria, *Annual review of genetics*, 37, pp. 409–33.
46. Pardo-López L., Soberón M., Bravo A., 2013, *Bacillus thuringiensis* insecticidal three-domain Cry toxins: mode of action, insect resistance and consequences for crop protection, *FEMS microbiology reviews*, 37, pp. 3–22.
47. Ben-Dov E., 2014, *Bacillus thuringiensis subsp. israelensis* and its dipteran-specific toxins, *Toxins*, 6, pp. 1222–1243.

48. Jenkins J.L., Dean D.H., 2000, Exploring the mechanism of action of insecticidal proteins by genetic engineering methods, *Genetic Engineering*, 22, pp. 33–54.
49. Schnepf E., Crickmore N., Van Rie J., Lereclus D., Baum J., Feitelson J., 1998, *Bacillus thuringiensis* and its pesticidal crystal proteins, *Microbiology and Molecular Biology Reviews*, 62, pp.775–806.
50. Mendoza-Almanza G., Esparza-Ibarra E.L., Ayala-Luján J.L., Mercado-Reyes M., Godina-González S., Hernández-Barrales M., 2020, The cytotoxic spectrum of *Bacillus thuringiensis* toxins: from insects to human cancer cells, *Toxins*, 12(5), pp. 301. <https://doi.org/10.3390/toxins12050301>
51. Liu L., Boyd S.D., Bulla L.A., Winkler D.D., 2018, The defined toxin-binding region of the cadherin G-protein coupled receptor, BT-R1, for the active Cry1Ab toxin of *Bacillus thuringiensis*, *Journal of Proteomics & Bioinformatics*, 11, pp. 201–210.
52. Arenas I., Bravo A., Soberón M., Gómez I., 2010, Role of alkaline phosphatase from *Manduca sexta* in the mechanism of action of *Bacillus thuringiensis* Cry1Ab toxin, *The Journal of Biological Chemistry*, 285, pp. 12497–12503.
53. Gómez I., Dean D.H., Bravo A., Soberón M., 2003, Molecular basis for *Bacillus thuringiensis* Cry1Ab toxin specificity: two structural determinants in the *Manduca sexta* Bt-R1 receptor interact with loops alpha-8 and 2 in domain II of Cyl1Ab toxin, *Biochemistry*, 42, pp. 10482–10489.
54. Gómez I., Sánchez J., Miranda R., Bravo A., Soberón M., 2002, Cadherin-like receptor binding facilitates proteolytic cleavage of helix alpha-1 in domain I and oligomer pre-pore formation of *Bacillus thuringiensis* Cry1Ab toxin, *FEBS letters*, 513, pp. 242–246.
55. Bravo A., Gómez I., Conde J., Muñoz-Garay C., Sánchez J., Miranda R., 2004, Oligomerization triggers binding of a *Bacillus thuringiensis* Cry1Ab pore-forming toxin to aminopeptidase N receptor leading to insertion into membrane microdomains, *Biochimica Et Biophysica Acta*, 1667, pp. 38–46.
56. Pacheco S., Gómez I., Arenas I., Saab-Rincon G., Rodríguez-Almazán C., Gill S.S., 2009, Domain II loop 3 of *Bacillus thuringiensis* Cry1AB toxin is involved in a “ping pong” binding mechanism with *Manduca sexta* aminopeptidase-N and cadherin receptors, *The Journal of Biological Chemistry*, 284, pp. 32750–32757.

57. Ocelotl J., Sánchez J., Gómez I., Tabashnik B.E., Bravo A., Soberón M., 2017, ABCC2 is associated with *Bacillus thuringiensis* Cry1Ac toxin oligomerization and membrane insertion in diamondback moth, *Scientific Reports*, 7, pp. 2386.
58. ISAAA, 2017, Global status of commercialized biotech/GM crops in 2017: Biotech crop adoption surges as economic benefits accumulate in 22 years.
59. Lê Thị Thu Hiền, Phạm Lê Bích Hằng, Nguyễn Tường Vân, Lê Thị Minh Thành, Đào Thị Hằng, 2020, Nghiên cứu và phát triển các giống đậu tươn biến đổi gen sử dụng các gen kháng sâu có nguồn gốc từ vi khuẩn *Bacillus thuringiensis*, *Tạp chí Công nghệ Sinh học*, 18, tr. 1–21.
60. Xiao Y., Wu K., 2019, Recent progress on the interaction between insects and *Bacillus thuringiensis* crops, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374, pp. 20180316. <https://doi.org/10.1098/rstb.2018.0316>
61. Bernardi O., Malvestiti G.S., Dourado P.M., Oliveira W.S., Martinelli S., Berger G.U., 2012, Assessment of the high-dose concept and level of control provided by MON 87701 × MON 89788 soybean against *Anticarsia gemmatalis* and *Pseudoplusia includens* (Lepidoptera: Noctuidae) in Brazil, *Pest Management Science*, 68, pp. 1083–1091
62. MacRae T.C., Baur M.E., Boethel D.J., Fitzpatrick B.J., Gao A.-G., Gamundi J.C., 2005, Laboratory and field evaluations of transgenic soybean exhibiting high-dose expression of a synthetic *Bacillus thuringiensis cryIA* gene for control of Lepidoptera, *Journal of Economic Entomology*, 98, pp. 577–587.
63. ISAAA, 2021, Biotech crops drive socio-economic development and sustainable environment in the New Frontier.
64. Bengyella L., Yekwa E.L., Iftikhar S., Nawaz K., Jose R.C., Fonmboh D.J., 2018, Global challenges faced by engineered *Bacillus thuringiensis cry* genes in soybean (*Glycine max* L.) in the twenty-first century, *3 Biotech*, 8, pp. 464. <https://doi.org/10.1007/s13205-018-1484-8>
65. Marques L., Santos A., Castro B., Moscardini V., Rosseto J., Silva O., 2017, Field evaluation of soybean transgenic event DAS-81419-2 expressing Cry1F and Cry1Ac proteins for the control of secondary lepidopteran pests in Brazil, *Crop Protection*, 96, pp. 109–115.
66. Qin D., Liu X.-Y., Miceli C., Zhang Q., Wang P.-W., 2019, Soybean plants expressing the *Bacillus thuringiensis cry8*-like gene show resistance to

- Holotrichia parallela*, *BMC biotechnology*, 19, pp. 66.
<https://doi.org/10.1186/s12896-019-0563-1>
67. Ngoc P., Lan V., Trang T., Thuong N., Ngoc L., Ha C., 2015, Agrobacterium-mediated transformation of *cry8Db* gene in Vietnam sweet potato cultivar, *Journal of Life Sciences*, 10(6).
<https://doi.org/10.17265/1934-7391/2015.06.002>
 68. Trịnh Minh Hợp, Nguyễn Thị Nhã, Trịnh Thị Vân Anh, Võ T.X.T., Nguyễn N.U.T., 2016, Nghiên cứu tạo giống bông kháng sâu và chịu thuốc trừ cỏ bằng kỹ thuật chuyển gen, *Kỷ yếu các công trình khoa học và công nghệ nổi bật trong ngành Công Thương giai đoạn 2011-2015*.
 69. Nguyễn Thị Nhã, Trịnh Thị Vân Anh, Nguyễn N.U.T., Trịnh Minh Hợp, 2016, Kết quả chuyển gen VIP3a vào cây bông thông qua *Agrobacterium tumefaciens*, *Tạp chí KH&CN - Bộ Công Thương*.
 70. Nguyễn Thị Nhã, Trịnh Thị Vân Anh, Võ T.X.T., Nguyễn N.U.T., Trịnh Minh Hợp, Phạm Bích Ngọc, 2016, Hiệu quả chuyển gen kháng thuốc diệt cỏ ở cây bông (*Gossypium hirsutum* L.) thông qua *A. tumefaciens*, *Tạp chí Nông nghiệp và Phát triển Nông thôn*.
 71. Nguyễn Văn Đồng, Nguyễn Anh Vũ, 2013, Phân lập và thiết kế vector biểu hiện gen liên quan đến khả năng chịu hạn GmMyb ở đậu tương, *Tạp chí Khoa học và Công nghệ nông nghiệp Việt Nam*, 2, tr. 49–54.
 72. Nguyễn Văn Đồng, Nguyễn Anh Vũ, Nguyễn H.K., Dương T.B., 2013, Nghiên cứu biến nạp gen liên quan đến khả năng kháng hạn và thuốc trừ cỏ vào giống đậu tương ĐT22, *Tạp chí Nông nghiệp và Phát triển nông thôn*, 11, tr. 3–9.
 73. Trần Thị Cúc Hòa, 2010, Tạo chọn giống đậu tương biến đổi gen kháng sâu ở Việt Nam, *Tạp chí Viện Khoa học Nông nghiệp Việt Nam*, tr. 214–220.
 74. Trần Thị Cúc Hòa, Nguyễn T.H.B., Trần Thanh Hải, Nguyễn Q.V., Nguyễn Đ.C., 2013, Nghiên cứu chọn tạo các giống đậu tương chuyển gen kháng ruồi đục thân và sâu đục quả, *Hội thảo quốc gia về khoa học cây trồng lần thứ nhất*, tr. 441–449
 75. Mayjonade B., Gouzy J., Donnadiou C., Pouilly N., Marande W., Callot C., 2016, Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules, *BioTechniques*, 61, pp. 203–205.
 76. Gurevich A., Saveliev V., Vyahhi N., Tesler G., 2013, QUAST: quality assessment tool for genome assemblies, *Bioinformatics*, 29, pp. 1072–1075.

77. 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, pp. 3210–3212.
78. Waterhouse R.M., Seppey M., Simão F.A., Manni M., Ioannidis P., Klioutchnikov G., 2017, BUSCO applications from quality assessments to gene prediction and phylogenomics, *Molecular Biology and Evolution*, 35(3), pp. 543-548.
79. Wattam A.R., Davis J.J., Assaf R., Boisvert S., Brettin T., Bun C., 2017, Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center, *Nucleic Acids Research*, 45, pp. 535–542.
80. Brettin T., Davis J.J., Disz T., Edwards R.A., Gerdes S., Olsen G.J., 2015, RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes, *Scientific Reports*, 5, pp. 8365. <https://doi.org/10.1038/srep08365>
81. Bradford M.M., 1976, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical Biochemistry*, 72, pp. 248–254.
82. Finney D.J., 1972, *Probit analysis*, 3rd ed, Cambridge University Press, Cambridge, 333 p.
83. Nakamura Y., Gojobori T., Ikemura T., 2000, Codon usage tabulated from international DNA sequence databases: status for the year 2000, *Nucleic Acids Research*, 28, pp. 292. <https://doi.org/10.1093/nar/28.1.292>
84. Gutiérrez R.A., MacIntosh G.C., Green P.G., 1999, Current perspectives on mRNA stability in plants: multiple levels and mechanisms of control, *Trends in Plant Science*, 4, pp. 429–438.
85. Sambrook J., Fritsch, E.F., Maniatis T., 1989, *Molecular Cloning, A laboratory manual 2nd edn*, New York, Cold Spring Harbor Laboratory Press.
86. Zhang X., Du P., Lu L., Xiao Q., Wang W., Cao X., 2008, Contrasting effects of HC-Pro and 2b viral suppressors from Sugarcane mosaic virus and Tomato aspermy cucumovirus on the accumulation of siRNAs, *Virology*, 374, pp. 351–360.
87. Madden T., 2002, *The BLAST sequence analysis tool*, The NCBI Handbook.

88. Maqbool S., Riazuddin S., Loc N., Gatehouse A., Gatehouse J., Christou P., 2001, Expression of multiple insecticidal genes confers broad resistance against a range of different rice pests, *Molecular Breeding*, 7, pp. 85–93.
89. Gao S., Yang Y., Xu L., Guo J., Su Y., Wu Q., 2018, Particle Bombardment of the *cry2A* gene cassette induces stem borer resistance in Sugarcane, *International Journal of Molecular Sciences*, 19, pp. 1692. <https://doi.org/10.3390/ijms19061692>
90. Lenin K., Mariam M., Udayasuriyan V., 2001, Expression of a *cry2Aa* gene in an acrySTALLIFEROUS *Bacillus thuringiensis* strain and toxicity of Cry2Aa against *Helicoverpa armigera*, *World Journal of Microbiology and Biotechnology*, 17, pp. 273–278.
91. Singh S., Kumar N.R., Maniraj R., Lakshmikanth R., Rao K.Y.S., Muralimohan N., 2018, Expression of Cry2Aa, a *Bacillus thuringiensis* insecticidal protein in transgenic pigeon pea confers resistance to gram pod borer, *Helicoverpa armigera*, *Scientific Reports*, 8, pp. 8820. <https://doi.org/10.1038/s41598-018-26358-9>
92. Le Thị Thủy, Lê Thu Ngọc, Hồ Thị Thương, Nguyễn Thu Giang, Phạm Bích Ngọc, Chu Hoàng Hà, 2017, Biểu hiện Protein HA/H7N9 Polymer dung hợp IgMFC trong cây thuốc lá (*Nicotiana benthamiana*) bằng phương pháp agroinfiltration, *Tạp chí Khoa học ĐHQGHN: Khoa học Tự nhiên và Công nghệ*, 33(1), tr. 105-115.
93. Murray E., Lotzer J., Eberle M., 1989, Codon usage in plants, *Nucleic acids research*, 17, pp. 477–98.
94. Gustafsson C., Govindarajan S., Minshull J., 2004, Codon bias and heterologous protein expression, *Trends in Biotechnology*, 22, pp. 346–353.
95. Clive J., 2010, Global status of commercialized biotech/GM crops, ISAAA Briefs 42: (ISAAA, Ithaca, NY, 2011).
96. Qaim M., Zilberman D., 2003, Yield effects of genetically modified crops in developing countries, *Science (New York, N.Y.)*, 299, pp. 900–902.
97. Ferré J., Van Rie J., 2002, Biochemistry and genetics of insect resistance to *Bacillus thuringiensis*, *Annual Review of Entomology*, 47, pp. 501–533.
98. Janmaat A.F., Myers J., 2003, Rapid evolution and the cost of resistance to *Bacillus thuringiensis* in greenhouse populations of cabbage loopers, *Trichoplusia ni*, *Proceedings. Biological Sciences*, 270, pp. 2263–2270.

99. Tabashnik B.E., Van Rensburg J.B.J. & Carrière Y., 2009, Field-evolved insect resistance to *Bt* crops: definition, theory, and data, *Journal of Economic Entomology*, 102, pp. 2011–2025.
100. Kruger M., Van Rensburg J.B.J., Van den Berg J., 2009, Perspective on the development of stem borer resistance to *Bt* maize and refuge compliance at the Vaalharts irrigation scheme in South Africa, *Crop Protection*, 28, pp. 684–689.
101. Dhurua S., Gujar G., 2011, Field-evolved resistance to *Bt* toxin Cry1Ac in the pink bollworm, *Pectinophora gossypiella* (Saunders) (Lepidoptera: Gelechiidae), from India, *Pest management science*, 67, pp. 898–903.
102. Gassmann A.J., Petzold-Maxwell J.L., Keweshan R.S., Dunbar M.W., 2011, Field-evolved resistance to *Bt* maize by Western corn rootworm, *PLoS ONE*, 6, pp. e22629. <https://doi.org/10.1371/journal.pone.0022629>
103. Zhang H., Yin W., Zhao J., Jin L., Yang Y., Wu S., 2011, Early warning of cotton bollworm resistance associated with intensive planting of *Bt* cotton in China, *PLoS ONE*, 6, pp. e22874. <https://doi.org/10.1371/journal.pone.0022874>
104. Palma L., Muñoz D., Berry C., Murillo J., Caballero P., 2014, Draft genome sequences of two *Bacillus thuringiensis* strains and characterization of a putative 41.9-kDa insecticidal toxin, *Toxins*, 6, pp. 1490–1504.
105. Sayyed A.H., Haward R., Herrero S., Ferré J., Wright D.J., 2000, Genetic and biochemical approach for characterization of resistance to *Bacillus thuringiensis* toxin Cry1Ac in a field population of the diamondback moth, *Plutella xylostella*, *Applied and Environmental Microbiology*, 66, pp. 1509–1516.
106. Berón C.M., Curatti L., Salerno G.L., 2005, New strategy for identification of novel *cry*-type genes from *Bacillus thuringiensis* strains, *Applied and Environmental Microbiology*, 71, pp. 761–765.
107. Juárez-Pérez V.M., Ferrandis M.D., Frutos R., 1997, PCR-based approach for detection of novel *Bacillus thuringiensis cry* genes. *Applied and Environmental Microbiology*, 63, pp. 2997–3002.
108. Noguera P.A., Ibarra J.E., 2010, Detection of new *cry* genes of *Bacillus thuringiensis* by use of a novel PCR primer system, *Applied and Environmental Microbiology*, 76, pp. 6150–6155.
109. Tan F., Zhu J., Tang J., Tang X., Wang S., Zheng A., 2009, Cloning and characterization of two novel crystal protein genes, *cry54Aa1* and *cry30Fa1*,

- from *Bacillus thuringiensis* strain BtMC28, *Current Microbiology*, 58, pp. 654–659.
110. Ye W., Zhu L., Liu Y., Crickmore N., Peng D., Ruan L., 2012, Mining new crystal protein genes from *Bacillus thuringiensis* on the basis of mixed plasmid-enriched genome sequencing and a computational pipeline, *Applied and Environmental Microbiology*, 78, pp. 4795–4801.
 111. Lazarte N., Valacco M., Moreno S., Salerno G., Berón C., 2021, Molecular characterization of a *Bacillus thuringiensis* strain from Argentina, toxic against Lepidoptera and Coleoptera, based on its whole-genome and Cry protein analysis, *Journal of Invertebrate Pathology*, 183, pp. 107563.
 112. Naveenarani M., Suesha G.S., Srikanth J., Hari K., Sankaranarayanan C., Mahesh P., Nirmala R., Swathik C.P., Crickmore N., Bakshi R., Appunu C., Singaravelu B., 2022, Whole genome analysis and functional characterization of a novel *Bacillus thuringiensis* (Bt 62) isolate against sugarcane white grub *Holotrichia serrata* (F), *Genomics*, 114, pp. 185–195.
 113. Pacheco S., Gómez I., Chiñas M., Sánchez J., Soberón M., Bravo A., 2021, Whole genome sequencing analysis of *Bacillus thuringiensis* GR007 reveals multiple pesticidal protein genes, *Frontiers in Microbiology*, 12, pp. 3268.
 114. Sajid M., Geng C., Li M., Wang Y., Liu H., Zheng J., 2018, Whole-genome analysis of *Bacillus thuringiensis* revealing partial genes as a source of novel Cry toxins, *Applied and Environmental Microbiology*, 84, pp. e00277-18. <https://doi.org/10.1128/AEM.00277-18>
 115. Jia B., Jeon C.O., 2016, High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives, *Open Biology*, 6(8), pp. 160196. <https://doi.org/10.1098/rsob.160196>
 116. Singh A., Upadhyay V., Upadhyay A.K., Singh S.M., Panda A.K., 2015, Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process, *Microbial Cell Factories*, 14, pp. 41. <https://doi.org/10.1186/s12934-015-0222-8>
 117. Baig F., Fernando L.P., Salazar M.A., Powell R.R., Bruce T.F., Harcum S.W., 2014, Dynamic transcriptional response of *Escherichia coli* to inclusion body formation, *Biotechnology and Bioengineering*, 111, pp. 980–999.
 118. Gutiérrez-González M., Farías C., Tello S., Pérez-Etcheverry D., Romero A., Zúñiga R., 2019, Optimization of culture conditions for the expression of

- three different insoluble proteins in *Escherichia coli*, *Scientific Reports*, 9, pp. 16850. <https://doi.org/10.1038/s41598-019-53200-7>
119. Park A.-R., Jang S.-W., Kim J.-S., Park Y.-G., Koo B.-S., Lee H.-C., 2018, Efficient recovery of recombinant CRM197 expressed as inclusion bodies in *E.coli*, *PLOS One*, 13, pp. e0201060. <https://doi.org/10.1371/journal.pone.0201060>
 120. Hire R.S., Makde R.D., Dongre T.K., D'souza S.F., 2009, Expression, purification and characterization of the Cry2Aa14 toxin from *Bacillus thuringiensis subsp. Kenyae*, *Toxicon*, 54, pp. 519–524.
 121. Rosano G.L., Morales E.S., Ceccarelli E.A., 2019, New tools for recombinant protein production in *Escherichia coli*: A 5-year update, *Protein Science*, 28, pp. 1412–1422.
 122. Costa S., Almeida A., Castro A. & Domingues L., 2014, Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system, *Frontiers in Microbiology*, 5, pp. 1-20.
 123. Han X., Ning W., Ma X., Wang X., Zhou K., 2020, Improving protein solubility and activity by introducing small peptide tags designed with machine learning models, *Metabolic Engineering Communications*, 11, pp. e00138. <https://doi.org/10.1016/j.mec.2020.e00138>
 124. de Marco A., Deuerling E., Mogk A., Tomoyasu T., Bukau B., 2007, Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*, *BMC Biotechnology*, 7, pp. 32. <https://doi.org/10.1186/1472-6750-7-32>
 125. Baneyx F., Palumbo J.L., 2003, Improving heterologous protein folding via molecular chaperone and foldase co-expression, *Methods in Molecular Biology (Clifton, N.J.)*, 205, pp. 171–197.
 126. Berndt C., Lillig C.H., Holmgren A., 2008, Thioredoxins and glutaredoxins as facilitators of protein folding, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1783, pp. 641–650.
 127. San-Miguel T., Pérez-Bermúdez P., Gavidia I., 2013, Production of soluble eukaryotic recombinant proteins in *E. coli* is favoured in early log-phase cultures induced at low temperature, *SpringerPlus*, 2(1), pp. 89. <https://doi.org/10.1186/2193-1801-2-89>
 128. Estruch J.J., Carozzi N.B., Desai N., Duck N.B., Warren G.W., Koziel M.G., 1997. Transgenic plants: an emerging approach to pest control, *Nature Biotechnology*, 15, pp. 137–141.

129. Adang M.J., Brody M.S., Cardineau G., Eagan N., Roush R.T., Shewmaker C.K., 1993, The reconstruction and expression of a *Bacillus thuringiensis cryIIIA* gene in protoplasts and potato plants, *Plant Molecular Biology*, 21, pp. 1131–1145.
130. LuShui W., LiHua D., FengXiang L., GuoYing X., 2014, Optimization of the Cry2Aa gene and development of insect-resistant and herbicide-tolerant photoperiod-sensitive genic male sterile rice, *Czech Journal of Genetics and Plant Breeding*, 50, pp. 19–25.
131. Mazier M., Pannetier C., Tourneur J., Jouanin L., Giband M., 1997, The expression of *Bacillus thuringiensis* toxin genes in plant cells, In: *Biotechnology Annual Review*. (Ed. M.R. El-Gewely), Elsevier, pp. 313–347.
132. Xu C., Wei H., Wang L., Yin T., Zhuge Q., 2019, Optimization of the *cryIAh1* sequence enhances the hyper-resistance of transgenic poplars to *Hyphantria cunea*, *Frontiers in Plant Science*, 10, pp. 335. <https://doi.org/10.3389/fpls.2019.00335>
133. Cheema H., Khan A., Khan M., Aslam U., Ahmad I., Khan I., 2015, Assessment of *Bt* cotton genotypes for the Cry1Ac transgene and its expression, *The Journal of Agricultural Science*, 153, pp. 1-9.
134. Li X., Li S., Lang Z., Zhang J., Zhu L., Huang D., 2013, Chloroplast-targeted expression of the codon-optimized truncated *cryIAh* gene in transgenic tobacco confers a high level of protection against insects, *Plant cell reports*, 32(8), pp. 1299-1308.
135. Niu L., Mannakkara A., Qiu L., Wang X., Hua H., Lei C., 2017, Transgenic *Bt* rice lines producing Cry1Ac, Cry2Aa or Cry1Ca have no detrimental effects on Brown Planthopper and Pond Wolf Spider, *Scientific Reports*, 7, pp. 1940. <https://doi.org/10.1038/s41598-017-02207-z>
136. Sun M., Gao X., Zhao Z., Li A., Wang Y., Yang Y., 2020, Enhanced production of recombinant proteins in *Corynebacterium glutamicum* by constructing a bicistronic gene expression system, *Microbial Cell Factories*, 19, pp. 113. <https://doi.org/10.1186/s12934-020-01370-9>
137. Düzenli Ö., Okay S., 2020, Promoter engineering for the recombinant protein production in prokaryotic systems, *AIMS Bioengineering*, 7, pp. 62–81.
138. Seternes T., Tonheim T.C., Myhr A.I., Dalmo R.A., 2016, A plant 35S CaMV promoter induces long-term expression of luciferase in Atlantic salmon, *Scientific Reports*, 6, pp. 25096. <https://doi.org/10.1038/srep25096>

139. Abdeen A., Virgós A., Olivella E., Villanueva J., Avilés X., Gabarra R., 2005, Multiple insect resistance in transgenic tomato plants over-expressing two families of plant proteinase inhibitors, *Plant Molecular Biology*, 57, pp. 189–202.
140. Atkinson N., Leitão N., Orr D.J., Meyer M.T., Carmo-Silva E., Griffiths H., 2017, Rubisco small subunits from the unicellular green alga *Chlamydomonas* complement Rubisco-deficient mutants of *Arabidopsis*, *New Phytologist*, 214, pp. 655–667.
141. Rajeev Kumar S., Anunanthini P., Ramalingam S., 2015, Epigenetic silencing in transgenic plants, *Frontiers in Plant Science*, 6, pp. 693. <https://doi.org/10.3389/fpls.2015.00693>
142. Vaucheret H., Béclin C., Elmayan T., Feuerbach F., Godon C., Morel J.-B., 1998, Transgene-induced gene silencing in plants, *The Plant Journal*, 16, pp. 651–659.
143. Christov N.K., Imaishi H., Ohkawa H., 1999, Green-tissue-specific expression of a reconstructed *cryIC* gene encoding the active fragment of *Bacillus thuringiensis* delta-endotoxin in haploid tobacco plants conferring resistance to *Spodoptera litura*, *Bioscience, Biotechnology, and Biochemistry*, 63, pp. 1433–1444.
144. Tanabe N., Tamoi M., Shigeoka S., 2015, The sweet potato RbcS gene (IbRbcS1) promoter confers high-level and green tissue-specific expression of the GUS reporter gene in transgenic *Arabidopsis*, *Gene*, 567, pp. 244–250.
145. Wang Y., 2014, Green tissue-specific analysis of a cloned *rbcS* promoter from *Lemna gibba*, *Czech Journal of Genetics and Plant Breeding*, 50, pp. 235–240.
146. Norkunas K., Harding R., Dale J., Dugdale B., 2018, Improving agroinfiltration-based transient gene expression in *Nicotiana benthamiana*, *Plant Methods*, 14, pp. 71. <https://doi.org/10.1186/s13007-018-0343-2>
147. Davis K., Gkotsi D.S., Smith D.R.M., Goss R.J.M., Caputi L., O'Connor S.E., 2020, *Nicotiana benthamiana* as a transient expression host to produce auxin analogs, *Frontiers in Plant Science*, 11, pp. 1791. <https://doi.org/10.3389/fpls.2020.581675>
148. Goulet M.-C., Gaudreau L., Gagné M., Maltais A.-M., Laliberté A.-C., Éthier G., 2019, Production of Biopharmaceuticals in *Nicotiana benthamiana*—axillary stem growth as a key determinant of total protein

- Yield, *Frontiers in Plant Science*, 10, pp. 735.
<https://doi.org/10.3389/fpls.2019.00735>
149. Ma L., Lukasik E., Gawehns F., Takken F.L.W., 2012, The use of agroinfiltration for transient expression of plant resistance and fungal effector proteins in *Nicotiana benthamiana* leaves, *Methods in Molecular Biology (Clifton, N.J.)*, 835, pp. 61–74.
 150. Zhang Y., Chen M., Siemiatkowska B., Toleco M.R., Jing Y., Strotmann V., 2020, A highly efficient *Agrobacterium*-mediated method for transient gene expression and functional studies in multiple plant species, *Plant Communications*, 1, pp. 100028. <https://doi.org/10.1016/j.xplc.2020.100028>
 151. Cao V., Pamplona R., Kim J., Oh Y., Cho S.-K., Ahn J., 2017, Optimization of *Agrobacterium*-mediated transient expression of heterologous genes in spinach, *Plant Biotechnology Reports*, 11. <https://doi.org/10.1007/s11816-017-0457-4>
 152. Hanittinan O., Oo Y., Chaotham C., Rattanapisit K., Shanmugaraj B., Phoolcharoen W., 2020, Expression optimization, purification and *in vitro* characterization of human epidermal growth factor produced in *Nicotiana benthamiana*, *Biotechnology Reports*, 28, pp. e00524. <https://doi.org/10.1016/j.btre.2020.e00524>
 153. Mondal T., Bhattacharya A., Ahuja P., Chand P., 2001, Transgenic tea [*Camellia sinensis* (L.) O. Kuntze cv. Kangra Jat] plants obtained by *Agrobacterium*-mediated transformation of somatic embryos, *Plant Cell Reports*, 8, pp. 712–720.
 154. Alkanaimsh S., Karuppanan K., Guerrero A., Tu A.M., Hashimoto B., Hwang M.S., 2016, Transient expression of tetrameric recombinant human butyrylcholinesterase in *Nicotiana benthamiana*, *Frontiers in Plant Science*, 7, pp. 1-13.
 155. Mbewana S., Meyers A., Weber B., Mareledwane V., Ferreira M., Majiwa P., 2018, Expression of Rift Valley fever virus N-protein in *Nicotiana benthamiana* for use as a diagnostic antigen, *BMC Biotechnology*, 18(1), pp. 77.
 156. Souza Prado G., Bamogo P., Abreu J., Gillet F.-X., Santos V., Jean-Paul B., 2019, *Nicotiana benthamiana* is a suitable transient system for high-level expression of an active inhibitor of cotton boll weevil α -amylase, *BMC Biotechnology*, 19(1), pp. 15. <https://doi.org/10.1186/s12896-019-0507-9>

157. Kim K., Kang Y., Park S.R., Kim D.-S., Lee S.-W., Ko K., 2021, Effect of leaf position and days post-infiltration on transient expression of colorectal cancer vaccine candidate proteins GA733-Fc and GA733-FcK in *Nicotiana benthamiana* plant, *PeerJ*, 9, pp. e10851. <https://doi.org/10.7717/peerj.10851>
158. Mo Q., Mai R., Yang Z., Chen M., Yang T., Lai H., 2012, A hydroponic cultivation system for rapid high-yield transient protein expression in *Nicotiana* plants under laboratory conditions, *Nan Fang Yi Ke Da Xue Xue Bao- Journal of Southern Medical University*, 32, pp. 772–777.
159. Abbas M., 2018, Genetically engineered (modified) crops (*Bacillus thuringiensis* crops) and the world controversy on their safety, *Egyptian Journal of Biological Pest Control*, 28. <https://doi.org/10.1186/s41938-018-0051-2>
160. Akhurst R.J., James W., Bird L.J., Beard C., 2003, Resistance to the Cry1Ac delta-endotoxin of *Bacillus thuringiensis* in the cotton bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae), *Journal of Economic Entomology*, 96, pp. 1290–1299.
161. Walsh T., James B., Chakroun M., Ferré J., Downes S., 2018, Isolating, characterising and identifying a Cry1Ac resistance mutation in field populations of *Helicoverpa punctigera*, *Scientific Reports*, 8, pp. 2626. <https://doi.org/10.1038/s41598-018-21012-w>
162. Wang P., Zhao J.-Z., Rodrigo-Simón A., Kain W., Janmaat A.F., Shelton A.M., 2007, Mechanism of resistance to *Bacillus thuringiensis* toxin Cry1Ac in a greenhouse population of the cabbage looper, *Trichoplusia ni*, *Applied and Environmental Microbiology*, 73, pp. 1199–1207.
163. Sevim E., Eryüzlü E., Demirbağ Z., Demir I., 2012, A novel *cry2Ab* gene from the indigenous isolate *Bacillus thuringiensis* subsp. *Kurstaki*, *Journal of microbiology and biotechnology*, 22(1), pp. 133-140.
164. Yamamoto T., Powell G.K., 1993, *Bacillus thuringiensis* crystal proteins : Recent advances in understanding its insecticidal activity, *Advanced Engineered Pesticides*, New York, N.Y: Marcel Dekker, Inc, pp. 3–42
165. Elleuch J., Jaoua S., Ginibre C., Chandre F., Tounsi S., Zribi Zghal R., 2016, Toxin stability improvement and toxicity increase against Dipteran and Lepidopteran larvae of *Bacillus thuringiensis* crystal protein Cry2Aa, *Pest Management Science*, 72(12), pp. 2240-2246.
166. Gouffon C., Van Vliet A., Van Rie J., Jansens S., Jurat-Fuentes J.L., 2011, Binding sites for *Bacillus thuringiensis* Cry2Ae toxin on heliothine brush

border membrane vesicles are not shared with Cry1A, Cry1F, or Vip3A toxin, *Applied and Environmental Microbiology*, 77, pp. 3182–3188.

167. Le T.N., Le T.M.T., Pham B. N., Trinh T.T.H., Dong V.Q., Ngo D.B., Chu H.H., Nguyen V.D., 2022, Detection of a novel Cry2Ab toxin against *Etiella zinckenella* Treitschke from the *Bacillus thuringiensis* serovar canadensis SP142 strain, *Plant Protection Science*, 58, pp. 158–169.

PHỤ LỤC

Bảng phụ lục 1: Dữ liệu thử nghiệm hoạt tính kháng sâu của các chủng *Bt*

STT	Ký hiệu	Dưới loài	Địa điểm/Nguồn phân lập	OD _{600nm} (72 giờ)	Hình dạng tinh thể	Hoạt tính diệt sâu đục quả đậu tương (%)
1	JP12.7		Sa Pa	2.571		0
2	SP1.1		Sa Pa	3.048		0
3	SP1.24		Sa Pa	2.977		56.67
4	SP1.39		Sa Pa	2.841		23.33
5	SP10.17		Sa Pa	2.671		16.67
6	SP10.2		Sa Pa	3.495		0
7	SP10.5		Sa Pa	3.869		0
8	SP10.6	<i>B. t. konkukian</i>	Sa Pa	3.985		56.67
9	SP10.7	<i>B. t. morrisoni</i>	Sa Pa	3.584	Lưỡng tháp	23.33
10	SP10.9		Sa Pa	3.677		16.67
11	SP12.1		Sa Pa	2.770		33.33
12	SP12.2		Sa Pa	2.471		16.67
13	SP12.7		Sa Pa	2.480		13.33
14	SP13.1		Sa Pa	4.02		0
15	SP13.2		Sa Pa	3.409		0
16	SP13.6		Sa Pa	3.812		0
17	SP13.9	<i>B. t. tenebrionis</i>	Sa Pa	4.000	Lưỡng tháp, hình cầu	13.33
18	SP14.2	<i>B. t. canadensis</i>	Sa Pa	3.846	Lưỡng tháp	93.33
19	SP14.3	<i>B. t. kurstaki</i>	Sa Pa	3.575	Hình cầu, không xác định	13.33
20	SP15.2		Sa Pa	2.935	Hình cầu	3.33
21	SP15.3		Sa Pa	2.435	Lưỡng tháp	60
22	SP15.7		Sa Pa	3.976	Lưỡng tháp	26.67
23	SP2		Sa Pa	3.987	Lưỡng tháp	30
24	SP2.1		Sa Pa	3.446	Lưỡng tháp	80
25	SP2.11		Sa Pa	4.099	Lưỡng tháp	23.33
26	SP2.12	<i>B. t. kurstaki</i>	Sa Pa	2.968	Lưỡng tháp	13.33
27	SP2.2		Sa Pa	4.035	Lưỡng tháp	30
28	SP2.3	<i>B. t. kurstaki</i>	Sa Pa	2.536		0
29	SP2.6		Sa Pa	3.635		0
30	SP2.7		Sa Pa	2.378		26.67
31	SP2.8		Sa Pa	4.081		10
32	SP2.9		Sa Pa	3.946		3.33
33	SP20.1		Sa Pa	4.047		13.33
34	SP20.2		Sa Pa	3.236	Hình cầu	33.33

STT	Ký hiệu	Dưới loài	Địa điểm/Nguồn phân lập	OD _{600nm} (72 giờ)	Hình dạng tinh thể	Hoạt tính diệt sâu đục quả đậu tương (%)
35	SP20.3	<i>B. t. kurstaki</i>	Sa Pa	3.437	Không xác định	16.67
36	SP20.4		Sa Pa	2.446	Lưỡng tháp, không xác định	26.67
37	SP20.8		Sa Pa	2.46	Hình cầu	56.67
38	SP20.9	<i>B.t. konkukian</i>	Sa Pa	2.686		0
39	SP21.2		Sa Pa	3.579		3.33
40	SP21.4		Sa Pa	2.473		16.67
41	SP22.1		Sa Pa	2.301		63.33
42	SP22.10	<i>B.t. konkukian</i>	Sa Pa	2.753	Lưỡng tháp	0
43	SP22.16		Sa Pa	2.750		0
44	SP22.2		Sa Pa	3.973		0
45	SP22.3		Sa Pa	3.354		0
46	SP22.4		Sa Pa	3.354		13.33
47	SP22.9		Sa Pa	3.750		0
48	SP23.10	<i>B. t. morrisoni</i>	Sa Pa	3.076	Hình cầu	66.67
49	SP23.11		Sa Pa	2.494	Lưỡng tháp	0
50	SP23.13	<i>B. t. kurstaki</i>	Sa Pa	2.665	Lưỡng tháp	0
51	SP26.1	<i>B. t. kurstaki</i>	Sa Pa	2.951	Lưỡng tháp	33.33
52	SP3.1		Sa Pa	2.944	Lưỡng tháp, hình cầu	16.67
53	SP3.11		Sa Pa	3.036	Lưỡng tháp	13.33
54	SP3.15		Sa Pa	3.060	Lưỡng tháp	0
55	SP3.16	<i>B.t. aizawai</i>	Sa Pa	3.468	Lưỡng tháp	0
56	SP3.17		Sa Pa	2.968	Lưỡng tháp	0
57	SP3.20	<i>B. t. sumiyoshiensis</i>	Sa Pa	2.056	Hình cầu	0
58	SP3.21		Sa Pa	2.836		30
59	SP3.23		Sa Pa	3.956		56.67
60	SP3.25		Sa Pa	2.046	Lưỡng tháp, hình cầu, không xác định	0
61	SP3.4		Sa Pa	2.147		3.33
62	SP3.9		Sa Pa	3.068		0
63	SP31.20		Sa Pa	2.170		36.67
64	SP4.1		Sa Pa	3.089		16.67
65	SP4.13		Sa Pa	4.240		16.67
66	SP4.2		Sa Pa	3.870		50
67	SP4.5	<i>B. t. morrisoni</i>	Sa Pa	2.335	Lưỡng tháp	10
68	SP4.6	<i>B. t. kurstaki</i>	Sa Pa	3.480	Lưỡng tháp	43.33
69	SP4.7	<i>B. t. colmeri</i>	Sa Pa	2.930	Không xác định	0

STT	Ký hiệu	Dưới loài	Địa điểm/Nguồn phân lập	OD _{600nm} (72 giờ)	Hình dạng tinh thể	Hoạt tính diệt sâu đục quả đậu tương (%)
70	SP4.8	<i>B. t. kurstaki</i>	Sa Pa	3.050	Hình cầu	26.67
71	SP5.1	<i>B. t. pondicheriensis</i>	Sa Pa	3.823	Hình cầu	40
72	SP5.7	B.t. <i>pondicheriensis</i>	Sa Pa	3.509	Lưỡng tháp, không xác định	36.67
73	SP6.1	<i>B. t. fukuokaensis</i>	Sa Pa	3.401	Lưỡng tháp	30
74	SP6.2		Sa Pa	3.700		0
75	SP6.4		Sa Pa	3.401		73.33
76	SP6.5		Sa Pa	3.806		33.33
77	SP7.1	B.t. <i>aizawai</i>	Sa Pa	3.306	Lưỡng tháp, Không xác định	0
78	SP7.10		Sa Pa	2.08		16.67
79	SP7.11		Sa Pa	2.406		33.33
80	SP7.12		Sa Pa	2.870		16.67
81	SP7.15		Sa Pa	3.771		13.33
82	SP7.2		Sa Pa	3.793		0
83	SP7.3	<i>B. t. fukuokaensis</i>	Sa Pa	2.350	Lưỡng tháp	0
84	SP7.4		Sa Pa	3.461		0
85	SP7.6		Sa Pa	2.247		0
86	SP7.8		Sa Pa	2.863		0
87	SP7.9	<i>B. t. kurstaki</i>	Sa Pa	3.862	Hình cầu	33.33
88	SP8.1		Sa Pa	3.864		0
89	SP8.3		Sa Pa	3.469		0
90	SP9.1	B.t. <i>kurstaki</i>	Sa Pa	2.861	Lưỡng tháp	3.33
91	SP6.3	<i>B. t. fukuokaensis</i>	Sa Pa	3.088	Hình cầu	0
92	SP17.3		Sa Pa	2.787		36.67
93	SP15.7		Sa Pa	2.751		16.67
94	SP7.2		Sa Pa	2.781		16.67
95	SP4.1		Sa Pa	3.485		50
96	SP3.3		Sa Pa	3.849		10
97	SP15.3		Sa Pa	3.965		80
98	SP7.4		Sa Pa	3.554		43.33
99	SP6.4		Sa Pa	3.837		0
100	SP3.1		Sa Pa	2.455		0
101	SP12.1		Sa Pa	2.291		80
102	SP22.16	B.t. <i>konkukian</i>	Sa Pa	2.600	Hình cầu	40
103	N4		Sa Pason La	4.002	Hình cầu, không xác định	16.67
104	L3	<i>B. t. aizawai</i>	Lào Cai	3.997	Hình cầu	0
105	LVT2.4	<i>B. t. kurstaki</i>	Lào Cai	4.048	Lưỡng tháp	13.33

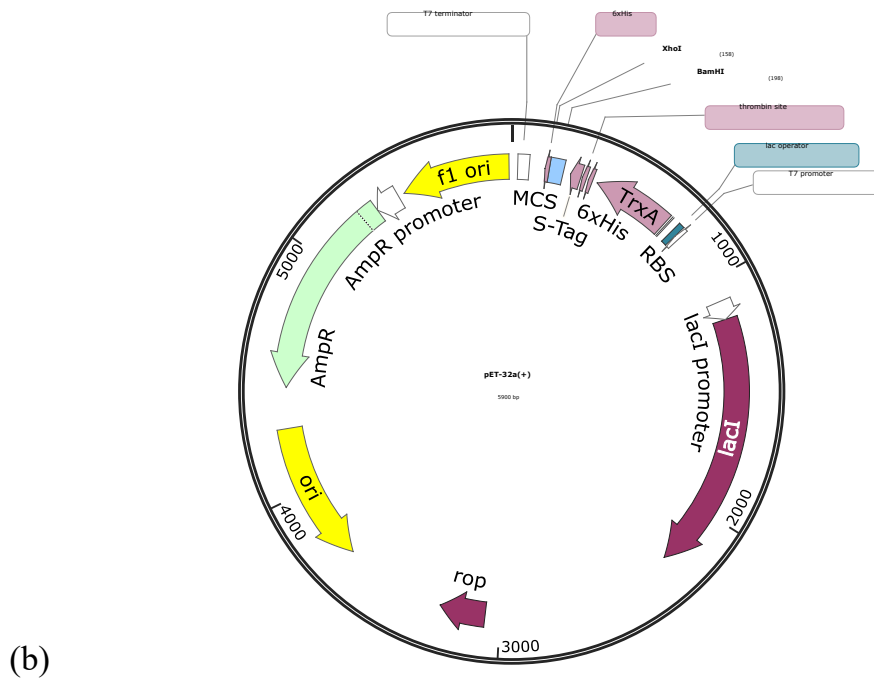
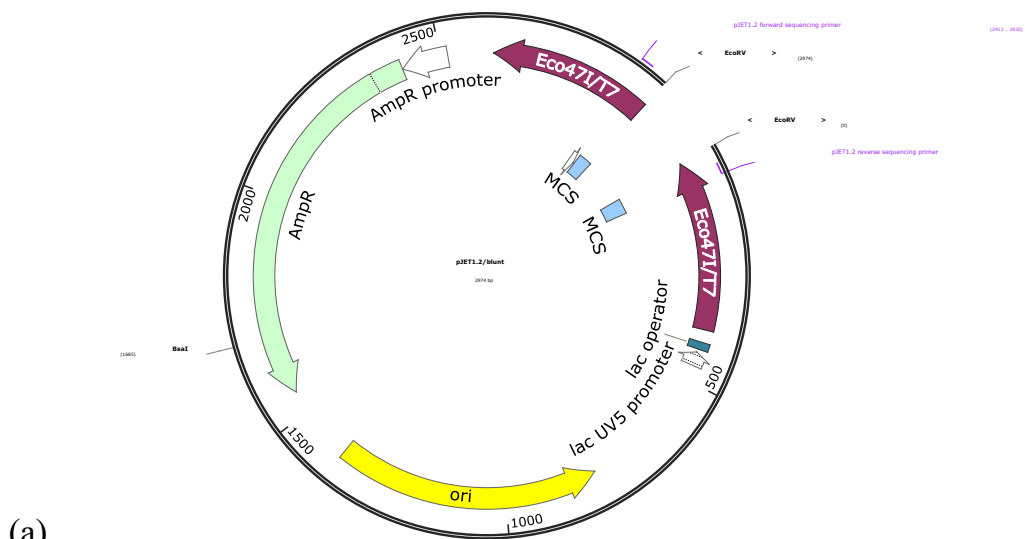
STT	Ký hiệu	Dưới loài	Địa điểm/Nguồn phân lập	OD _{600nm} (72 giờ)	Hình dạng tinh thể	Hoạt tính diệt sâu đục quả đậu tương (%)
106	LVT2.4	<i>B. t. kurstaki</i>	Lào Cai	3.247	Lưỡng thấp	0
107	LVT3.5	<i>B. t. aizawai</i>	Lào Cai	3.457	Lưỡng thấp, không xác định	0
108	Na	<i>B. t. kurstaki</i>	Lào Cai	3.969	Lưỡng thấp	10
109	NỒĐR1	<i>B. t. kurstaki</i>	Lào Cai	2.657	Lưỡng thấp	16.67
110	NT22.3	<i>B.t. israelensis</i>	Lào Cai	2.371		30
111	TC1.3	<i>B.t. israelensis</i>	Lào Cai	2.143		10
112	TC3.6	<i>B.t. israelensis</i>	Lào Cai	2.550		16.67
113	TC9.3	<i>B.t. israelensis</i>	Lào Cai	3.663		30
114	TD5.4	<i>B.t. israelensis</i>	Lào Cai	3.074	Lưỡng thấp, hình cầu	10
115	Yan	<i>B. t. kurstaki</i>	Lào Cai	3.094	Lưỡng thấp	43.33
116	NoĐB7	<i>B.t. israelensis</i>	Lào Cai	3.130	Hình cầu	0
117	CS2.2		Lào Cai	3.066		30
118	CS3.13	<i>B.t. israelensis</i>	Lào Cai	2.424	Lưỡng thấp	43.33
119	PL2		Lào Cai	2.645		50
120	BDI2.2		Lào Cai	2.961		16.67
121	DBB7.4	<i>B.t. israelensis</i>	Lào Cai	2.994	Lưỡng thấp	3.33
122	ĐBI3.5	<i>B.t. israelensis</i>	Lào Cai	3.686	Hình cầu	13.33
123	LC01	<i>B. t. kurstaki</i>	Lào Cai	3.460	Lưỡng thấp	0
124	Đ6.1	<i>B. t. aizawai</i>	Tuyên Quang	2.325	Hình cầu	93,33
125	Đ6.2	<i>B. t. kurstaki</i>	Tuyên Quang	3.490	Hình cầu	100
126	Đ7.1	<i>B. t. kurstaki</i>	Tuyên Quang	2.670	Lưỡng thấp	0
127	Đ7.16	<i>B. t. kurstaki</i>	Tuyên Quang	3.470	Lưỡng thấp	13.33
128	Đ7.17	<i>B. t. morrisoni</i>	Tuyên Quang	3.023	Lưỡng thấp, không xác định	0
129	Đ7.19	<i>B. t. morrisoni</i>	Tuyên Quang	3.729	Hình cầu	0
130	Đ7.7	<i>B. t. kurstaki</i>	Tuyên Quang	3.401	Lưỡng thấp	0
131	Đ8.1	<i>B. t. kurstaki</i>	Tuyên Quang	3.000	Lưỡng thấp	26.67
132	Đ9.8	<i>B. t. kurstaki</i>	Tuyên Quang	3.601	Lưỡng thấp	70
133	ĐC1.6		Tuyên Quang	3.406	Hình cầu	0
134	ĐC2.12		Tuyên Quang	3.306		40
135	ĐC2.3		Tuyên Quang	2.408		46.67
136	ĐC2.5		Tuyên Quang	2.606		0
137	ĐC2.6		Tuyên Quang	2.780		0
138	ĐC2.7		Tuyên Quang	3.421		10
139	ĐC210		Tuyên Quang	3.603	Hình cầu, không xác định	16.67
140	ĐC211		Tuyên Quang	2.980		30
141	ĐC212		Tuyên Quang	3.761		10

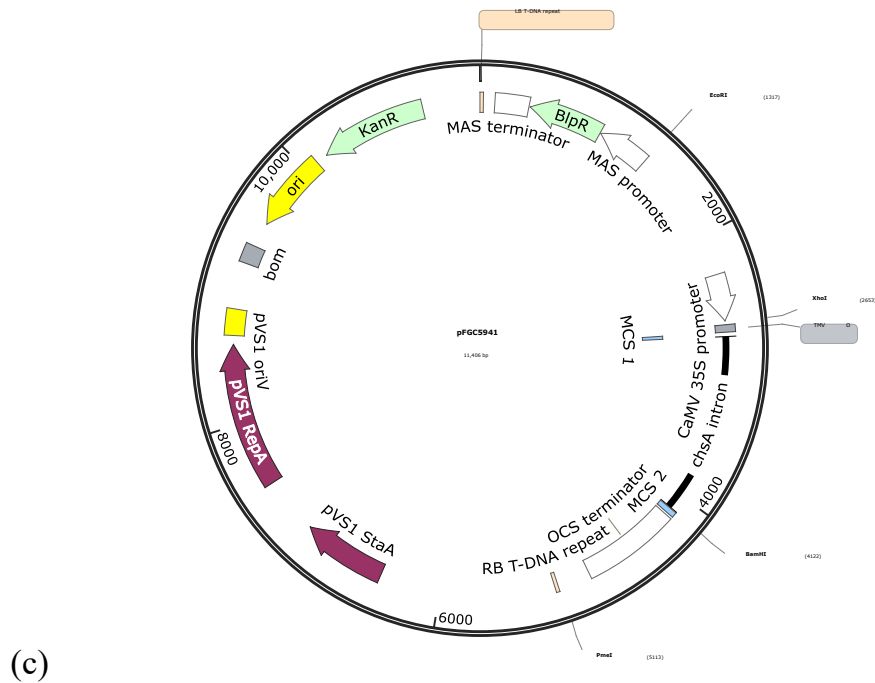
STT	Ký hiệu	Dưới loài	Địa điểm/Nguồn phân lập	OD _{600nm} (72 giờ)	Hình dạng tinh thể	Hoạt tính diệt sâu đục quả đậu tương (%)
142	DHG1.2		Hà Giang	3.523	Hình cầu	0
143	DHG10.2		Hà Giang	3.09	Lưỡng tháp	73.33
144	DHG10.3		Hà Giang	3.021	Lưỡng tháp	3.33
145	DHG10.5		Hà Giang	3.350	Lưỡng tháp	10
146	DHG11.1		Hà Giang	3.851	Hình cầu	26.67
147	DHG12.5		Hà Giang	3.836		0
148	DHG12.6		Hà Giang	3.696		0
149	DHG13.1		Hà Giang	2.278	Hình cầu	0
150	DHG13.3		Hà Giang	2.286	Lưỡng tháp	0
151	DHG13.6		Hà Giang	2.950		0
152	DHG14.1		Hà Giang	3.931		0
153	DHG14.4		Hà Giang	3.833		0
154	DHG14.5		Hà Giang	2.940		0
155	DHG14.6		Hà Giang	3.091		3.33
156	DHG15.1		Hà Giang	2.837		16.67
157	DHG15.2		Hà Giang	2.633		80
158	DHG16.1		Hà Giang	3.932		30
159	DHG16.4		Hà Giang	3.024		50
160	DHG16.5		Hà Giang	3.709		73.33
161	DHG2.1	<i>B.t. novosibisk</i>	Hà Giang	2.131	Hình cầu	0
162	DHG2.4	<i>B. t. kurstaki</i>	Hà Giang	3.348	Lưỡng tháp	33.33
163	DHG3.1		Hà Giang	2.737	Không xác định	16.67
164	DHG3.2	<i>B.t. fukuokaensis</i>	Hà Giang	2.731	Không xác định	56.67
165	DHG3.3		Hà Giang	2.68		0
166	DHG4.1		Hà Giang	3.405		0
167	DHG4.2		Hà Giang	3.769		0
168	DHG4.3		Hà Giang	3.535		3.33
169	DHG4.4		Hà Giang	3.844		16.67
170	DHG5.2		Hà Giang	3.837		66.67
171	DHG5.5		Hà Giang	2.255		0
172	DHG7.1		Hà Giang	2.831		0
173	DHG7.4		Hà Giang	2.800		0
174	DHG7.5	<i>B.t. fukuokaensis</i>	Hà Giang	4.302	Hình cầu	23.33
175	DHG8.2		Hà Giang	3.029		3.33
176	DHG8.3		Hà Giang	3.032		60
177	DHG8.4		Hà Giang	4.462	Lưỡng tháp	26.67
178	DHG8.6	<i>B.t. coreanensis</i>	Hà Giang	3.805		30
179	DHG8.7		Hà Giang	3.259	Lưỡng tháp	23.33

STT	Ký hiệu	Dưới loài	Địa điểm/Nguồn phân lập	OD _{600nm} (72 giờ)	Hình dạng tinh thể	Hoạt tính diệt sâu đục quả đậu tương (%)
180	DHG9.1		Hà Giang	2.836		43.33
181	DHG9.2		Hà Giang	2.937		16.67
182	DHG9.3		Hà Giang	3.932	Hình cầu	26.67
183	DHG9.4	<i>B.t. fukuokaensis</i>	Hà Giang	3.556	Không xác định	60
184	LHG10.2		Hà Giang	3.425		0
185	LHG10.3		Hà Giang	4.746		0
186	LHG13.1		Hà Giang	2.571		0
187	LHG13.2		Hà Giang	4.331		23.33
188	LHG14.2		Hà Giang	2.731		16.67
189	LHG14.3		Hà Giang	3.737		10
190	LHG14.4	<i>B.t. isalensis</i>	Hà Giang	2.950	Hình cầu	56.67
191	LHG14.6		Hà Giang	4.249		0
192	LHG14.7		Hà Giang	3.684		0
193	LHG3.1		Hà Giang	4.112		80
194	LHG3.2		Hà Giang	3.240		26.67
195	LHG3.3		Hà Giang	3.681		0
196	LHG4.2	<i>B.t. isalensis</i>	Hà Giang	2.462	Hình cầu	46.67
197	LHG4.4		Hà Giang	2.940	Hình cầu	0
198	LHG5.2		Hà Giang	2.826	Lưỡng tháp	0
199	LHG5.5		Hà Giang	3.249	Lưỡng tháp	0
200	BD10.8		Điện Biên	3.682	Lưỡng tháp	3.33
201	BD125.13		Điện Biên	4.732	Lưỡng tháp	23.33
202	BD20.7		Điện Biên	3.585	Lưỡng tháp	10
203	BD3.5	<i>B. t. kurstaki</i>	Điện Biên	3.839	Lưỡng tháp	50
204	BD3.5		Điện Biên	2.736	Lưỡng tháp	50
205	BD4.5	<i>B. t. kurstaki</i>	Điện Biên	2.637	Lưỡng tháp	0
206	BD6.2	<i>B. t. kurstaki</i>	Điện Biên	3.362	Lưỡng tháp	66.67
207	BĐ6.2	<i>B. t. kurstaki</i>	Điện Biên	3.886	Lưỡng tháp	83.33
208	BD8.2	<i>B. t. kurstaki</i>	Điện Biên	3.465	Lưỡng tháp	100
209	BĐ8.2		Điện Biên	4.846	Hình cầu	80
210	BD9.1		Điện Biên	2.471	Lưỡng tháp	0
211	BD10.1	<i>B. t. aizawai</i>	Điện Biên	4.341	Lưỡng tháp, Không xác định	0
212	BD3.5	<i>B. t. aizawai</i>	Điện Biên	2.861	Lưỡng tháp, Không xác định	50
213	DB15.4		Điện Biên	2.680	Lưỡng tháp	0
214	ĐB3.5		Điện Biên	4.249	Lưỡng tháp	3.33
215	ĐB4.3		Điện Biên	3.864	Lưỡng tháp	16.67
216	ĐB9.1		Điện Biên	4.012	Lưỡng tháp,	63.33

STT	Ký hiệu	Dưới loài	Địa điểm/Nguồn phân lập	OD _{600nm} (72 giờ)	Hình dạng tinh thể	Hoạt tính diệt sâu đục quả đậu tương (%)
					Hình cầu	
217	ĐBJ2.3		Điện Biên	3.240	Lưỡng thấp	0
218	DDB15.3		Điện Biên	3.571	Lưỡng thấp	0
219	DDB25.2		Điện Biên	2.242	Hình cầu	0
220	DDB17.6		Điện Biên	2.680	Hình cầu, Không xác định	0
221	BD23.7		Điện Biên	2.846	Lưỡng thấp	23.33

Hình phụ lục 1. Các plasmid sử dụng để tách dòng và thiết kế vector





(c)

Hình phụ lục 2. Kết quả tìm kiếm và so sánh trình tự các protein độc tố *Bt* trên AllergenOnline.

(a) protein Cry1Aa; (b) protein Cry1Ac; (c) protein Cry1Be; (d) protein Cry1Na; (e) protein Cry2Ab; (f) protein Cry2Ah

```
# fasta36.exe -q -B -m 9i -w 80 -E 1 -d 20 C:\Windows\Temp\allFDF2.tmp version2136.fasta
FASTA searches a protein or DNA sequence data bank
version 36.3.8g Oct, 2018
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query: C:\Windows\Temp\allFDF2.tmp
1>>>query - 1176 aa
Library: version2136.fasta
540227 residues in 2233 sequences

Statistics: Altschul/Gish params: n0: 1176 Lambda: 0.158 K: 0.019 H: 0.100
statistics sampled from 288 (288) to 288 sequences
Algorithm: FASTA (3.8 Nov 2011) [optimized]
Parameters: BL50 matrix (15:-5), open/ext: -10/-2
ktup: 2, E-join: 1 (0.408), E-opt: 0.2 (0.129), width: 16
Scan time: 0.000
!! No sequences with E() < 1
>>>///
```

(a)

```
# fasta36.exe -q -B -m 9i -w 80 -E 1 -d 20 C:\Windows\Temp\allD1EA.tmp version2136.fasta
FASTA searches a protein or DNA sequence data bank
version 36.3.8g Oct, 2018
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query: C:\Windows\Temp\allD1EA.tmp
1>>>query - 1178 aa
Library: version2136.fasta
540227 residues in 2233 sequences

Statistics: Altschul/Gish params: n0: 1178 Lambda: 0.158 K: 0.019 H: 0.100
statistics sampled from 293 (293) to 293 sequences
Algorithm: FASTA (3.8 Nov 2011) [optimized]
Parameters: BL50 matrix (15:-5), open/ext: -10/-2
ktup: 2, E-join: 1 (0.454), E-opt: 0.2 (0.131), width: 16
Scan time: 0.000
!! No sequences with E() < 1
>>>///
```

(b)

```
# fasta36.exe -q -B -m 9i -w 80 -E 1 -d 20 C:\Windows\Temp\all17A.tmp version2136.fasta
FASTA searches a protein or DNA sequence data bank
version 36.3.8g Oct, 2018
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query: C:\Windows\Temp\all17A.tmp
1>>>query - 1227 aa
Library: version2136.fasta
540227 residues in 2233 sequences

Statistics: Altschul/Gish params: n0: 1227 Lambda: 0.158 K: 0.019 H: 0.100
statistics sampled from 238 (238) to 238 sequences
Algorithm: FASTA (3.8 Nov 2011) [optimized]
Parameters: BL50 matrix (15:-5), open/ext: -10/-2
ktup: 2, E-join: 1 (0.341), E-opt: 0.2 (0.107), width: 16
Scan time: 0.000
!! No sequences with E() < 1
>>>///
```

(c)

```
>>gi|46486948|gid|1121|allergen Pen m 2 [Fenneropenaeus chinensis] (53 aa)
initn: 44 init1: 44 opt: 71 Z-score: 114.9 bits: 27.8 E(2233): 0.31
Smith-Waterman score: 71; 28.8% identity (67.3% similar) in 52 aa overlap (550-601:2-50)
```

```
      510      520      530      540      550      560      570      580
query  APLSQRYRVRIRYASTTDLQFVTSINGTTINIGNFKTINNLTGSEGVRTVSFSTPFSFNAQSIIFRLGIQAFSGVQE
      ... ..: . . . .: . . . .:
notag|  YNLQVRGTRGEHTEAEGGIYDISNKR---RMGLTEFQAVKE
      10      20      30

      590      600
query  VYVDKIEFIPVE
      . . .: .:
notag|  MQDGILELIKMEKEM
      40      50
```

(d)


```
# fasta36.exe -q -B -m 9i -w 80 -E 1 -d 20 C:\Windows\Temp\allA70C.tmp version2136.fasta
FASTA searches a protein or DNA sequence data bank
version 36.3.8g Oct, 2018
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448
```

```
Query: C:\Windows\Temp\allA70C.tmp
1>>>query - 633 aa
Library: version2136.fasta
540227 residues in 2233 sequences
```

```
Statistics: Altschul/Gish params: n0: 633 Lambda: 0.158 K: 0.019 H: 0.100
statistics sampled from 231 (231) to 231 sequences
Algorithm: FASTA (3.8 Nov 2011) [optimized]
Parameters: BL50 matrix (15:-5), open/ext: -10/-2
ktup: 2, E-join: 1 (0.356), E-opt: 0.2 (0.103), width: 16
Scan time: 0.000
!! No sequences with E() < 1
>>>///
```

(e)

```
# fasta36.exe -q -B -m 9i -w 80 -E 1 -d 20 C:\Windows\Temp\allBB76.tmp version2136.fasta
FASTA searches a protein or DNA sequence data bank
version 36.3.8g Oct, 2018
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448
```

```
Query: C:\Windows\Temp\allBB76.tmp
1>>>query - 632 aa
Library: version2136.fasta
540227 residues in 2233 sequences
```

```
Statistics: Altschul/Gish params: n0: 632 Lambda: 0.158 K: 0.019 H: 0.100
statistics sampled from 248 (248) to 248 sequences
Algorithm: FASTA (3.8 Nov 2011) [optimized]
Parameters: BL50 matrix (15:-5), open/ext: -10/-2
ktup: 2, E-join: 1 (0.372), E-opt: 0.2 (0.111), width: 16
Scan time: 0.000
!! No sequences with E() < 1
>>>///
```

(f)

Hình phụ lục 3. So sánh trình tự gen *cry2Ab39* và gen tham chiếu *cry2Ab3*

```

      10      20      30      40      50      60      70      80
cry2Ab3  ATGAATAGTGTATTGAATAGCGGAAGAAGTACTACTATTTGTGATGCGTATAATGTAGCGGCTCATGATCCATTTAGTTTTCA
cry2Ab39 M N S V L N S G R T T I C D A Y N V A A H D P F S F Q
      90     100     110     120     130     140     150     160
cry2Ab3  ACACAAATCATTAGATACCGTACAAAAGGAATGGACGGAGTGGAAAAAATAATCATAGTTTATACCTAGATCCTATTG
cry2Ab39 H K S L D T V Q K E W T E W K K N N H S L Y L D P I
      170     180     190     200     210     220     230     240
cry2Ab3  TTGGAAGTGTGGCTAGTTTTCTGTTAAAGAAAGTGGGGAGTCTTGTGGAAAAAGGATACTAAGTGAGTTACGGAATTTA
cry2Ab39 V G T V A S F L L K K V G S L V G K R I L S E L R N L
V G T V A S F L L K K V G S L V G K R I L S E L R N L
```



```

1130      1140      1150      1160      1170      1180      1190      1200
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
TTAGGAGTTGGCTAGATTTCAGGTTTCAGATCGGGAGGGCGTTGCCACCGTTACAAATGGCAAACAGAATCCTTTGAGACA
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
V R S W L D S G S D R E G V A T V T N W Q T E S F E T

      1210      1220      1230      1240      1250      1260      1270      1280
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
ACTTTAGGGTTAAGGAGTGGTGCTTTTACAGCTCGCGGTAATTCAAACTATTCCAGATTATTTATTCGTAATATTC
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T L G L R S G A F T A R G N S N Y F P D Y F I R N I S

      1290      1300      1310      1320      1330      1340      1350      1360
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
TGGAGTTCCTTTAGTTGTTAGAAAATGAAGATTTAAGAAGACCGTTTACACTATAATGAAATAAGAAATATAGCAAGTCCTT
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
G V P L V V R N E D L R R P L H Y N E I R N I A S P

      1370      1380      1390      1400      1410      1420      1430      1440
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
CAGGAACACCTGGTGGAGCAGAGCTTATATGGTATCTGTGCATAACAGAAAAATAATATCCATGCTGTTTCATGAAAT
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
S G T P G G A R A Y M V S V H N R K N N I H A V H E N

      1450      1460      1470      1480      1490      1500      1510      1520
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
GCTTCTATGATTCAATTTAGCGCCAAATGACTATACAGGATTTACTATTTCGCCGATACATGCAACTCAAGTGAATAATCA
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
G S M I H L A P N D Y T G F T I S P I H A T Q V N N Q

      1530      1540      1550      1560      1570      1580      1590      1600
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
AACACGAACATTTATTTCTGAAAATTTGGAATCAAGGTGATTCCTTTAAGGTTTGAACAAAACAACAGCAGCTCGTTT
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T R T F I S E K F G N Q G D S L R F E Q N N T T A R

      1610      1620      1630      1640      1650      1660      1670      1680
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
ATACGCTTAGAGGGAAATGGAAATAGTTACAATCTTTATTTAAGAGTTTCTTCAATAGGAAATCCACTATTTCGAGTTACT
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Y T L R G N G N S Y N L Y L R V S S I G N S T I R V T

      1690      1700      1710      1720      1730      1740      1750      1760
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
ATAAACGGTAGGGTATATACTGCTACAATGTTAATACTACTACAAATAACGATGGAGTTAATGATAATGGAGCTCGTTT
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
I N G R V Y T A T N V N T T T N N D G V N D N G A R F

      1770      1780      1790      1800      1810      1820      1830      1840
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
 TTCAGATATTAATATCGGTAATGTAGTAGCAAGTAGTAATTCGATGTACCATTAGATATAAATGTAACATTAAACTCCG
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
S D I N I G N V V A S S N S D V P L D I N V T L N S

      1850      1860      1870      1880      1890      1900
cry2Ab3  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
GTACTCAATTTGATCTTATGAATATTATGCTTGTACCAACTAATATTCACCACTTTATTA
cry2Ab39 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
G T Q F D L M N I M L V P T N I S P L Y *

```

Hình phụ lục 4. So sánh trình tự gen *cry1Be* và gen tham chiếu *cry1Be1*

```

      10      20      30      40      50      60      70      80
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTGACTTCAAATAGGAAAATGAGAATGAAATATAAATGCTTTATCGATTCCAGCTGTATCGAATCATTCGGCACAAAT
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
M T S N R K N E N E I I N A L S I P A V S N H S A Q M

      90      100     110     120     130     140     150     160
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
GAATCTATCAACCGATGCTCGTATTGAGGATAGCTTGTGTATAGCCGAGGGGAACAATATCGATCCATTTGTTAGCGCAT
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
N L S T D A R I E D S L C I A E G N N I D P F V S A

      170     180     190     200     210     220     230     240
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
CAACAGTCCAACGGGTATTAACATAGCTGGTAGAATACTAGGTGTATTAGCGGTACCGTTTGCTGGACAATAGCTAGT
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S T V Q T G I N I A G R I L G V L G V P F A G Q I A S

      250     260     270     280     290     300     310     320
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTTATAGTTTTCTTGTGGTGAATTATGGCCCGCGGCAGAGATCCTTGGGAAATTTCTTAGAACATGTCGACAACCT
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
F Y S F L V G E L W P R G R D P W E I F L E H V E Q L

      330     340     350     360     370     380     390     400
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TATAAGACAACAAGTAACAGAAAATACTAGGGATACGGCTCTTGTCTCGATTACAAGGTTTAGGAAATTCCTTTAGAGCCT
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
I R Q Q V T E N T R D T A L A R L Q G L G N S F R A

      410     420     430     440     450     460     470     480
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ATCAACAGTCACTTGAAGATTGGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTATACCCAATATATAGCC
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Y Q Q S L E D W L E N R D D A R T R S V L Y T Q Y I A

      490     500     510     520     530     540     550     560
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
TTAGAACTTGATTTTCTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAAGTTCATTATTAATGGTATATGCTCA
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
L E L D F L N A M P L F A I R N Q E V P L L M V Y A Q

      570     580     590     600     610     620     630     640
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGCTGCAAATTTACACCTATTATTATTGAGAGATGCCCTCTCTTTTGGTAGTGAATTTGGGCTTACATCCAAGAAATTC
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
A A N L H L L L L R D A S L F G S E F G L T S Q E I

      650     660     670     680     690     700     710     720
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AACGTTATTATGAGCGCCAAGTGGAAAAACGAGAGAATATTCTGATTATTGCGCAAGATGGTATAATACGGGTTTAAAT
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Q R Y Y E R Q V E K T R E Y S D Y C A R W Y N T G L N

      730     740     750     760     770     780     790     800
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AATTTGAGAGGGACAATGCTGAAAGTTGGTTGCGATATAATCAATTCGAGACTTAAACGCTAGGAGTATTAGATCT
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
N L R G T N A E S W L R Y N Q F R R D L T L G V L D L

      810     820     830     840     850     860     870     880
cry1Be1  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
AGTGGCACTATTCCAAGCTATGACACGCGTGTATTATCCAATGAATACCAAGTGTCAATTAACAAGAGAAATTTATACAG
cry1Be   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
V A L F P S Y D T R V Y P M N T S A Q L T R E I Y T

```

V A L F P S Y D T R V Y P M N T S A Q L T R E I Y T

890 900 910 920 930 940 950 960

cry1Be1 **ATCCAATTGGGAGAACAATGCACCTTCAGGATTTGCAAGTACGAATTGGTTTAAATAAATGCACCATCGTTTTCTGCC**

cry1Be D P I G R T N A P S G F A S T N W F N N N A P S F S A

890 900 910 920 930 940 950 960

cry1Be1 **ATAGAGGCTGCCGTTATTAGGCCTCCGCATCTACTTGAATTTCCAGAACAGCTTACAATTTTCAGCGTATTAAGTCGATG**

cry1Be I E A A V I R P P H L L D F P E Q L T I F S V L S R W

890 900 910 920 930 940 950 960

cry1Be1 **I E A A V I R P P H L L D F P E Q L T I F S V L S R W**

cry1Be I E A A V I R P P H L L D F P E Q L T I F S V L S R W

1050 1060 1070 1080 1090 1100 1110 1120

cry1Be1 **GAGTAATACTCAATATATGAATTACTGGGTGGGACATAGACTTGAATCGCAACAATAAGGGGGTCATTAAGTACCTCGA**

cry1Be S N T Q Y M N Y W V G H R L E S R T I R G S L S T S

1050 1060 1070 1080 1090 1100 1110 1120

cry1Be1 **S N T Q Y M N Y W V G H R L E S R T I R G S L S T S**

cry1Be S N T Q Y M N Y W V G H R L E S R T I R G S L S T S

1130 1140 1150 1160 1170 1180 1190 1200

cry1Be1 **CACACGGAAATACCAACTTCTATTAATCCTGTAACATTACAGTTCACATCTCGAGACGTTTATAGAACAGAATCATT**

cry1Be T H G N T N T S I N P V T L Q F T S R D V Y R T E S F

1130 1140 1150 1160 1170 1180 1190 1200

cry1Be1 **T H G N T N T S I N P V T L Q F T S R D V Y R T E S F**

cry1Be T H G N T N T S I N P V T L Q F T S R D V Y R T E S F

1210 1220 1230 1240 1250 1260 1270 1280

cry1Be1 **GCAGGGATAAATATACTTCTAACTACTCCTGTGAATGGAGTACCTTGGGCTAGATTTAATTTGGAGAAATCCCTGAATTC**

cry1Be A G I N I L L T T P V N G V P W A R F N W R N P L N S

1210 1220 1230 1240 1250 1260 1270 1280

cry1Be1 **A G I N I L L T T P V N G V P W A R F N W R N P L N S**

cry1Be A G I N I L L T T P V N G V P W A R F N W R N P L N S

1290 1300 1310 1320 1330 1340 1350 1360

cry1Be1 **TCTTAGAGGTAGCCTTCTATACTATAGGGTACTGGAGTGGGACACAACATTTGATTTCAGAACTGAATTACCAC**

cry1Be L R G S L L Y T I G Y T G V G T Q L F D S E T E L P

1290 1300 1310 1320 1330 1340 1350 1360

cry1Be1 **L R G S L L Y T I G Y T G V G T Q L F D S E T E L P**

cry1Be L R G S L L Y T I G Y T G V G T Q L F D S E T E L P

1370 1380 1390 1400 1410 1420 1430 1440

cry1Be1 **CAGAAACAGAACGACCAATATGAATCTTACAGTCATAGATTATCTAATAAAGACTAATATCAGGAAACACTTTG**

cry1Be P E T T E R P N Y E S Y S H R L S N I R L I S G N T L

1370 1380 1390 1400 1410 1420 1430 1440

cry1Be1 **P E T T E R P N Y E S Y S H R L S N I R L I S G N T L**

cry1Be P E T T E R P N Y E S Y S H R L S N I R L I S G N T L

1450 1460 1470 1480 1490 1500 1510 1520

cry1Be1 **AGAGCACCAGTATATCTTGGACGCACCGTAGTGCAGATCGTACAATACCATTAGTTCAGATAGCATAACACAAATACC**

cry1Be R A P V Y S W T H R S A D R T N T I S S D S I T Q I P

1450 1460 1470 1480 1490 1500 1510 1520

cry1Be1 **R A P V Y S W T H R S A D R T N T I S S D S I T Q I P**

cry1Be R A P V Y S W T H R S A D R T N T I S S D S I T Q I P

1530 1540 1550 1560 1570 1580 1590 1600

cry1Be1 **ATFGGTAATAATCATTCAACCTTAATTCAGGTACTCTGTAGTCAAGTGGCCAGGATTTACAGGAGGGGATATAATCCGAA**

cry1Be L V K S F N L N S G T S V V S G P G F T G G D I I R

1530 1540 1550 1560 1570 1580 1590 1600

cry1Be1 **L V K S F N L N S G T S V V S G P G F T G G D I I R**

cry1Be L V K S F N L N S G T S V V S G P G F T G G D I I R

1610 1620 1630 1640 1650 1660 1670 1680

cry1Be1 **CTAACGTTAATGGTAGTGTACTAAGTATGGGCTTAATTTTAAATAATACATCATTACAGCGGTATCGCGTGAGAGTTCGT**

cry1Be T N V N G S V L S M G L N F N N T S L Q R Y R V R V R

1610 1620 1630 1640 1650 1660 1670 1680

cry1Be1 **T N V N G S V L S M G L N F N N T S L Q R Y R V R V R**

cry1Be T N V N G S V L S M G L N F N N T S L Q R Y R V R V R

1690 1700 1710 1720 1730 1740 1750 1760

cry1Be1 **TATGCTGCTTCTCAAACAATGGTCTGAGGGTAACTGTGCGAGGGAGTACTACTTTTGATCAAGGATTCCCTAGTACTAT**

cry1Be Y A A S Q T M V L R V T V G G S T T F D Q G F P S T M

1690 1700 1710 1720 1730 1740 1750 1760

cry1Be1 **Y A A S Q T M V L R V T V G G S T T F D Q G F P S T M**

cry1Be Y A A S Q T M V L R V T V G G S T T F D Q G F P S T M

```

      1770      1780      1790      1800      1810      1820      1830      1840
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
GAGTGCAAATGAGTCTTTGACATCTCAATCATTAGATTTGCAGAATTCCTGTAGGTATTAGTGCATCTGGCAGTCAA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
S A N E S L T S Q S F R F A E F P V G I S A S G S Q
      1850      1860      1870      1880      1890      1900      1910      1920
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
CTGCTGGAATAAGTATAAGTAATAATGCAGGTAGACAAACGGTTTCACTTTGATAAAATGAATTCATTCCAATTACTGCA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T A G I S I S N N A G R Q T F H F D K I E F I P I T A
      1930      1940      1950      1960      1970      1980      1990      2000
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
ACCTTCGAAGCAGAATATGATTTAGAAAAGAGCGCAAGAGCGGTGAATGCTCTGTTTACTAATACGAATCCAAGAAGGTT
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T F E A E Y D L E R A Q E A V N A L F T N T N P R R L
      2010      2020      2030      2040      2050      2060      2070      2080
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
GAAACAGGTGTGACAGATTATCATATTGATGAAGTATCCAATTTAGTGGCGTGTATTCCGGATGAATCTGCCTTGATG
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
K T G V T D Y H I D E V S N L V A C L S D E F C L D
      2090      2100      2110      2120      2130      2140      2150      2160
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
AAAAGAGAGAATTACTTGAGAAAGTGAATATGCGAAAGACTCAGTGATGAAGAAACTTACTCCAAGATCCAACTTC
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
E K R E L L E K V K Y A K R L S D E R N L L Q D P N F
      2170      2180      2190      2200      2210      2220      2230      2240
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
ACATCCATCAATAAGCAACCAGACTTTCATATCTACTAATGAGCAATCGAATTTACATCTATCCATGAACAATCTGAACA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T S I N K Q P D F I S T N E Q S N F T S I H E Q S E H
      2250      2260      2270      2280      2290      2300      2310      2320
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
TGGATGGTGGGAAGTGAGAACATTACAATCCAGAAAGAAATGACGTATTTAAGAGAAATTACGTCATACTACCGGGTA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
G W W G S E N I T I Q E G N D V F K E N Y V I L P G
      2330      2340      2350      2360      2370      2380      2390      2400
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
CTTTTAATGAGTGTATCCGACGTATTATATCAAAAAATAGGGGAGCGGAAATAAAAGCTTACTCGCTACCAATTA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T F N E C Y P T Y L Y Q K I G E A E L K A Y T R Y Q L
      2410      2420      2430      2440      2450      2460      2470      2480
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
AGTGGCTATATGAAGATAGTCAAGATTGAGATATATTTGATTCGTTACAATGCGAAACATGAACATTGGATGTTCC
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
S G Y I E D S Q D L E I Y L I R Y N A K H E T L D V P
      2490      2500      2510      2520      2530      2540      2550      2560
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
AGGTACCGAGTCCGTATGGCCGCTTTCAAGTTGAAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGCACCACATT
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
G T E S V W P L S V E S P I G R C G E P N R C A P H
      2570      2580      2590      2600      2610      2620      2630      2640
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
TTGAATGGAATCCTGATCTAGATTGTTCTGCAGAGATGGAGAAAATGTGCGCATCATTCCCATCATTCTCTTTGGAT
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
F E W N P D L D C S C R D G E K C A H H S H H F S L D

```

```

                2650      2660      2670      2680      2690      2700      2710      2720
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
ATTGATGTTGGATGCATAGACTTGCATGAGAACCTAGGCGTGTGGGTGGTATTCAAGATTAAAGACGCAGGAAGGTCATGC
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
I D V G C I D L H E N L G V W V V F K I K T Q E G H A
                2730      2740      2750      2760      2770      2780      2790      2800
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
AAGACTAGGGAACCTGGAATTTATTGAAGAGAACCATTATTAGGAGAAGCACGTCTCGTGTGAAGAGAGCAGAGAAAA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
R L G N L E F I E E K P L L G E A L S R V K R A E K
                2810      2820      2830      2840      2850      2860      2870      2880
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
AATGGAGAGACAAACGTGAAAACTACAATTGGAAACAAAACGAGTATATACAGAGGCAAAAGAAGCTGTGGATGCTTTA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
K W R D K R E K L Q L E T K R V Y T E A K E A V D A L
                2890      2900      2910      2920      2930      2940      2950      2960
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
TTGTAGATTCTCAATATGATAGATTACAAGCGGATACAACATTTGGCATGATTCATGGCGCAGATAAAGCTTGTTCATCG
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
F V D S Q Y D R L Q A D T N I G M I H A A D K L V H R
                2970      2980      2990      3000      3010      3020      3030      3040
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
AATTCGAGAGGCGTATCTTTCAGAATTTCTGTATTCCAGGTTAAATGCGGAAATTTTGAAGAATTAGAAGTTCGCA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
I R E A Y L S E L S V I P G V N A E I F E E L E G R
                3050      3060      3070      3080      3090      3100      3110      3120
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
TTATCACTGCAATCTCCCTATACGATGCGAGAATGTCTGTTAAAAATGGTGATTTTAAATAATGGATTAGCATGCTGGAAT
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
I I T A I S L Y D A R N V V K N G D F N N G L A C W N
                3130      3140      3150      3160      3170      3180      3190      3200
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
GTAAAAGGGCATGTAGATGTACAAACAGAGCCATCACCGTTCTGTCCCTTGTATCCAGAAATGGGAAGCAGAAGTTCACA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
V K G H V D V Q Q S H H R S V L V I P E W E A E V S Q
                3210      3220      3230      3240      3250      3260      3270      3280
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
AGCAGTTCGCGTCTGTCCGGGCGTGGCTATATCTCCGTGTCACAGCGTACAAGAGGGATATGGAGAGGGTGTGTAA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
A V R V C P G R G Y I L R V T A Y K E G Y G E G C V
                3290      3300      3310      3320      3330      3340      3350      3360
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
CGATCCATGAAATCGAGAACAATACAGACGAACTAAAAATTTAAAAACTGTGAAGAAGGAAGTGTATCCAACGGATACA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T I H E I E N N T D E L K F K N C E E E E V Y P T D T
                3370      3380      3390      3400      3410      3420      3430      3440
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
GGAACGTGTAATGATTATACTGCACCAAGGTACAGCAGCATGTAATCCCGTAATGCTGGATATGAGGATGCATATGA
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
G T C N D Y T A H Q G T A A C N S R N A G Y E D A Y E
                3450      3460      3470      3480      3490      3500      3510      3520
cry1Be1 .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
AGTTGATACTACAGCATCTGTTAATTACAACCGACTTATGAAGAAGAACGTATACAGATGTACGAAGAGATAATCAT
cry1Be .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
V D T T A S V N Y K P T Y E E E T Y T D V R R D N H
                3530      3540      3550      3560      3570      3580      3590      3600

```

```

cry1Be1 .....|
GTGAATATGACAGAGGGTATGTGAATTATCCACCAGTACCAGCTGGTTATATGACAAAAGAATTAGAATACTTCCAGAA
C E Y D R G Y V N Y P P V P A G Y M T K E L E Y F P E
cry1Be .....C.....G.....G.....T.....
C E Y D R G Y V N Y P P L P A G Y V T K E L E Y F P E

          3610      3620      3630      3640      3650      3660      3670      3680
cry1Be1 .....|
ACCGATAAGGTATGGATTGAGATTGGAGAAACGGAAGGGAAGTTTATTGTAGACAGCGTGAATTACTCCTTATGGAGGA
T D K V W I E I G E T E G K F I V D S V E L L L M E E
cry1Be .....C..G.....
T D K V W I E I G E T E G K F I V D S V E L L L M E E

cry1Be1 ....
ATAG
cry1Be *
cry1Be ....
*
```

Hình phụ lục 5. So sánh trình tự gen cry2Ah và gen tham chiếu cry2Ah1

```

          10      20      30      40      50      60      70      80
cry2Ah1 .....|
ATGAATAATGTATGAAATAGCGGAAGAGCTACTAATGGTGATGCGTATAATGTAGTGGCTCATGATCCATTTAGTTTCA
M N N V L N S G R A T N G D A Y N V V A H D P F S F Q
cry2Ah .....
M N N V L N S G R A T N G D A Y N V V A H D P F S F Q

          90      100     110     120     130     140     150     160
cry2Ah1 .....|
ACATAAATCATTAGATACCATACAAGAAGAATGGATGGAGTGGAAAAAGATAATCATATTTTATATGTAGATCCTATTG
H K S L D T I Q E E W M E W K K D N H I L Y V D P I
cry2Ah .....
H K S L D T I Q E E W M E W K K D N H I L Y V D P I

          170     180     190     200     210     220     230     240
cry2Ah1 .....|
TTGGAAC TGTGGCTAGCTTTCTTTTAAAGAAAGTGGGGAGTCTTGTTGAAAAAGAATATTAAGTGAGTTACGGAATTTA
V G T V A S F L L K K V G S L V E K R I L S E L R N L
cry2Ah .....G.....
V G T V A S F L L K K V G S L V G K R I L S E L R N L

          250     260     270     280     290     300     310     320
cry2Ah1 .....|
ATATTTCCTAGTGGCAGTACAAATCTAATGCAAGATAATTTAAGAGAGACAGAAAAATTCCTGAATCAAAGACTTAATAC
I F P S G S T N L M Q D I L R E T E K F L N Q R L N T
cry2Ah .....
I F P S G S T N L M Q D I L R E T E K F L N Q R L N T

          330     340     350     360     370     380     390     400
cry2Ah1 .....|
AGACACTCTTGCCCGTGTAAATGCGGAATGACAGGGCTGCAAGCAAATGTAGAAGAGTTTAATCGACAAGTAGATAATT
D T L A R V N A E L T G L Q A N V E E F N R Q V D N
cry2Ah .....
D T L A R V N A E L T G L Q A N V E E F N R Q V D N

          410     420     430     440     450     460     470     480
cry2Ah1 .....|
TTTGAACCCCTAACCGAAATGCTGTTCCTTTATCAATAACTTCTTCAGTTAATACAATGCAGCAATTATTTCTAAATAGA
F L N P N R N A V P L S I T S S V N T M Q Q L F L N R
cry2Ah .....
F L N P N R N A V P L S I T S S V N T M Q Q L F L N R

          490     500     510     520     530     540     550     560
cry2Ah1 .....|
TTACCCCAGTTTTCAGATGCAAGGATACCAATTTATTATTACCTTTATTGGCACAGGCAGCCAAATTTACATCTTTCTTT
L P Q F Q M Q G Y Q L L L L P L F A Q A A N L H L S F
cry2Ah .....
L P Q F Q M Q G Y Q L L L L P L F A Q A A N L H L S F

          570     580     590     600     610     620     630     640
cry2Ah1 .....|
TATTAGAGATGTATTCTTAATGCAGATGAATGGGGAATTCAGCAGCAACATTACGTACGTATCAAATCACCTGAGAA
I R D V I L N A D E W G I S A A T L R T Y Q N H L R
cry2Ah .....
I R D V I L N A D E W G I S A A T L R T Y Q N H L R
```


I R D V I L N A D E W G I S A A T L R T Y Q N H L R

650 660 670 680 690 700 710 720

cry2Ah1 **ATTATACAAGAGAGTACTCTAATTATTGTATACTACGTATCAAACCTGCGTTTAGAGGTTTAAACACCCGTTTACACGAT**

cry2Ah N Y T R E Y S N Y C I T T Y Q T A F R G L N T R L H D

730 740 750 760 770 780 790 800

cry2Ah1 **ATGTTAGAATTTAGAACATATATGTTTTAAATGATTTGAATATGTATCTATCTGGTCGTTGTTAAATATCAAGCCT**

cry2Ah M L E F R T Y M F L N V F E Y V S I W S L F K Y Q S L

810 820 830 840 850 860 870 880

cry2Ah1 **TCTAGTATCTTCTGGCGCTAATTTATATGCAAGTGGTAGTGGACCACAGCAGACCCAATCATTACTTCAACAAGACTGGC**

cry2Ah L V S S G A N L Y A S G S G P Q Q T Q S F T S Q D W

890 900 910 920 930 940 950 960

cry2Ah1 **CATTTTTATATTCTTTTCCAAGTTAATTCGAATTAATGTTTAAATGGCTTTAGTGGCGCTAGACTTACGCAGACTTTC**

cry2Ah P F L Y S L F Q V N S N Y V L N G F S G A R L T Q T F

970 980 990 1000 1010 1020 1030 1040

cry2Ah1 **CCTAATATGTTGGTTTACCTGGTACTACTACACTCAGCATTGCTTCTGCAAGGGTCAATATACAGTGGAGGAGTTTC**

cry2Ah P N I V G L P G T T T T H A L L A A R V N Y S G G V S

1050 1060 1070 1080 1090 1100 1110 1120

cry2Ah1 **GTCTGGTGATATAGGCGCTGTGTTTAAATCAAAATTTAGTGTAGTACATTTCTCCACCTTTGTTAACACCATTGTTA**

cry2Ah S G D I G A V F N Q N F S C S T F L P P L L T P F V

1130 1140 1150 1160 1170 1180 1190 1200

cry2Ah1 **GAAGTTGGCTAGATTTCAGGTTTCAGATCGGGGGGATTAAATACCGTTACCAATGGCAACAGAAATCCTTTGAGACAAC**

cry2Ah R S W L D S G S D R G G I N T V T N W Q T E S F E T T

1210 1220 1230 1240 1250 1260 1270 1280

cry2Ah1 **TTAGGTTTAAAGAGTGGTGTCTTTACAGCTCGAGGTAATTCAAACTATTTCCAGATTATTTTCCGTAATATTTCTGG**

cry2Ah L G L R S G A F T A R G N S N Y F P D Y F I R N I S G

1290 1300 1310 1320 1330 1340 1350 1360

cry2Ah1 **AGTTCCTTAGTGTGTAGAAATGAAGATTTAAGAAGACCGTTACTATAATCAAATAAGAAATATAGAAAGTCTTCAG**

cry2Ah V P L V V R N E D L R R P L H Y N Q I R N I E S P S

1370 1380 1390 1400 1410 1420 1430 1440

cry2Ah1 **GAACACCTGGTGGATTACGAGCTTATATGGTATCTGTGCATAACAGAAAAATAATATCTATGCCGTTTCATGAAAATGGT**

cry2Ah G T P G G L R A Y M V S V H N R K N N I Y A V H E N G

1450 1460 1470 1480 1490 1500 1510 1520

cry2Ah1 **ACTATGATTCATTAGCGCCGGAAGATTATACAGGATTTACTATATCGCGATACATGCAACTCAAGTGAATAATCAAAC**

cry2Ah T M I H L A P E D Y T G F T I S P I H A T Q V N N Q T

T K I E G Y A K N K A L S E L A G I Q R N L E T Y I

410 420 430 440 450 460 470 480

cry1Na3 **AATTACGTAATGAATGGGAAATGATATTGAAACTCAAGGCTCAAGGTAAGGTAGCTAATTACTATGAAAGCTTTGAG**

cry1Na Q L R N E W E N D I E N S K A Q G K V A N Y Y E S L E

490 500 510 520 530 540 550 560

cry1Na3 **CAGGCGTTGAAAGGAGTATGCCTCAATTTGCAGTGGGGAATTTGAAGTACCCTTTAACTGTCTATGTGCAAGCTGC**

cry1Na Q A V E R S M P Q F A V G N F E V P L L T V Y V Q A A

570 580 590 600 610 620 630 640

cry1Na3 **TAATCTTCATTTATTATTATTAAGAGATGTTTTAGTTTATGGAAAGCGTTGGGGATGGTCGGAGCAGAAAATTAATAATTT**

cry1Na N L H L L L L R D V S V Y G K R W G W S E Q K I K I

650 660 670 680 690 700 710 720

cry1Na3 **AATTATGATAAACAGATTAAGTATACCCATGAATACACAATCATTGTGTAAATTTGGTATAATAAAGGACTTGAGAGATTA**

cry1Na Y Y D K Q I K Y T H E Y T N H C V N W Y N K G L E R L

730 740 750 760 770 780 790 800

cry1Na3 **AAAAATAAAGGTTCTTCTTATCAAGATTGGTACAATTAATCGTTTCCGTAGAGAAATGACTGTACTGTTTTAGATAT**

cry1Na K N K G S S Y Q D W Y N Y N R F R R E M T L T V L D I

810 820 830 840 850 860 870 880

cry1Na3 **CGTTGCTTTATTTCCCGCACTATGATGTACAACCTTATCCAATAACAACCGTTGCTCAGCTAACAGGGAGTTTATACGG**

cry1Na V A L F P H Y D V Q T Y P I T T V A Q L T R E V Y T

890 900 910 920 930 940 950 960

cry1Na3 **ATCCTTTACTTAATTTTAAATCCTAAATACATTCTGTGTCTCAATTACCTAGTTTTAGTGACATGGAAAATGCAACAATT**

cry1Na D P L L N F N P K L H S V S Q L P S F S D M E N A T I

970 980 990 1000 1010 1020 1030 1040

cry1Na3 **AGAACTCCACATCTGATGGAATTTTAAAGATGCTAACAATTTATACAGATTGGTATAGTGTGGGAAGAACTATTATTG**

cry1Na R T P H L M E F L R M L T I Y T D W Y S V G R N Y Y W

1050 1060 1070 1080 1090 1100 1110 1120

cry1Na3 **GGGAGGACATCGCGTGACGTCTTACCATGTAGGAGGAGAGAATATAAGATCACCTCTATATGGTAGAGAGGCAAAATCAAG**

cry1Na G G H R V T S Y H V G G E N I R S P L Y G R E A N Q

1130 1140 1150 1160 1170 1180 1190 1200

cry1Na3 **AGGTTCCTAGAGATTTTTATTTTTATGGACCGTTTTTAAAGACGTTATCAAGCCGACTCTAAGACCATTACAGCAGCCT**

cry1Na E V P R D F Y F Y G P V F K T L S K P T L R P L Q Q P

1210 1220 1230 1240 1250 1260 1270 1280

cry1Na3 **GCACCAGTCCCTCCTTTTAAATTTACGTAGCTTAGAGGGAGTAGAATTCACACTCCTACAGGTAGTTTTATGTATCGTGA**

cry1Na A P A P P F N L R S L E G V E F H T P T G S F M Y R E

```

      1290      1300      1310      1320      1330      1340      1350      1360
cry1Na3 AAGAGGATCGGTAGATTCTTTTAAATGAGTTACCGCCTTTTAAATCCAGTTGGGTACCTCATAAGGTATACAGTCACCGTT
      R G S V D S F N E L P P F N P V G L P H K V Y S H R
cry1Na .....
      R G S V D S F N E L P P F N P V G L P H K V Y S H R

      1370      1380      1390      1400      1410      1420      1430      1440
cry1Na3 TATGTCATGCAACGTTTGTTCGTAAATCTGGGACCCCTTATTTAACAACAGGTGCCATCTTTTCTTGACACATCGTAGT
      L C H A T F V R K S G T P Y L T T G A I F S W T H R S
cry1Na .....
      L C H A T F V R K S G T P Y L T T G A I F S W T H R S

      1450      1460      1470      1480      1490      1500      1510      1520
cry1Na3 GCTGAAGAAACCAATACAAATGAATCAAAATATTATTACGCAAAATCCCGTTAGTAAAAGCATATCAAATGGGTGAGGCAC
      A E E T N T I E S N I I T Q I P L V K A Y Q I G S G T
cry1Na .....
      A E E T N T I E S N I I T Q I P L V K A Y Q I G S G T

      1530      1540      1550      1560      1570      1580      1590      1600
cry1Na3 TACTGTAAAGGAAGGACCAGGATTCACAGGAGGGGATATACTTCGAAGAACAGGTCCTGGAACATTTGGAGATATGAGAA
      T V R K G P G F T G G D I L R R T G P G T F G D M R
cry1Na .....
      T V R K G P G F T G G D I L R R T G P G T F G D M R

      1610      1620      1630      1640      1650      1660      1670      1680
cry1Na3 TAAATATTAATGCACCATTATCTCAAAGATATCGTGTAGGATTCGTTATGCTTCTACGACAGATTTACAATTTGTCAGC
      I N I N A P L S Q R Y R V R I R Y A S T T D L Q F V T
cry1Na .....
      I N I N A P L S Q R Y R V R I R Y A S T T D L Q F V T

      1690      1700      1710      1720      1730      1740      1750      1760
cry1Na3 AGTATTAATGGGACCACCATTAATATTGGTAAC TTC CCAAAAATATTAATAATCTAAATACTTTAGGTTCTGAGGGCTA
      S I N G T T I N I G N F P K T I N N L N T L G S E G Y
cry1Na .....
      S I N G T T I N I G N F P K T I N N L N T L G S E G Y

      1770      1780      1790      1800      1810      1820      1830      1840
cry1Na3 TAGAACAGTATCGTTTAGTACTCCATTTAGTTTCTCAAATGCACAAAGCATATTTAGATTAGGTATACAAGCATTCTCTG
      R T V S F S T P F S F S N A Q S I F R L G I Q A F S
cry1Na .....
      R T V S F S T P F S F S N A Q S I F R L G I Q A F S

      1850      1860      1870      1880      1890
cry1Na3 GAGTTCAAGAAGTTTATGTGGATAAAATGAATTTATTCCTGTGAATAG
      G V Q E V Y V D K I E F I P V E *
cry1Na .....
      G V Q E V Y V D K I E F I P V E *

```

Hình phục lục 7. Bằng độc quyền sáng chế của chủng *Bt* SP14.2 mang gen mã hóa độc tố Cry2Ab39



Hình phụ lục 8. Dữ liệu đăng kí trên GenBank của các gen *cry* mã hóa độc tố diệt sâu mới phân lập được từ các chủng *Bt* bản địa

❖ Dữ liệu gen và độc tố *Bt Cry2Ab39* của chủng SP14.2:

Database	BestMatchFinder	Analysis	Submit your sequence for naming	Links	Feedback	Cart ()
<input type="checkbox"/>	Cry2Ab34	Cry2Ab34		ANN87794	2016	Download
<input type="checkbox"/>	Cry2Ab35	Cry2Ab35		ART33428	2017	Download
<input type="checkbox"/>	Cry2Ab36	Cry2Ab36		QBO24621	2019	Download
<input type="checkbox"/>	Cry2Ab37	Cry2Ab37		QIN91189	2020	Download
<input type="checkbox"/>	Cry2Ab38			MT199207	2020	Download
<input checked="" type="checkbox"/>	Cry2Ab39	Cry2Ab39		QIQ19560	2020	Download
<input type="checkbox"/>	Cry2Ab40	Cry2Ab40		MT941864	2020	Download
*	Cry2Ab41			MW238547	2020	
<input type="checkbox"/>	Cry2Ac1	Cry2Ac1	CryIIc	CAA40536	1991	Download
<input type="checkbox"/>	Cry2Ac2	Cry2Ac2		AAG35410	2000	Download
<input type="checkbox"/>	Cry2Ac3	Cry2Ac3		AAQ52385	2003	Download
<input type="checkbox"/>	Cry2Ac4	Cry2Ac4		ABC95997	2006	Download
<input type="checkbox"/>	Cry2Ac5	Cry2Ac5		ABC74969	2005	Download
<input type="checkbox"/>	Cry2Ac6	Cry2Ac6		ABC74793	2006	Download

Bacillus thuringiensis strain SP142 Cry toxin (cry2Ab) gene, partial cds

GenBank: MN319700.1

[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS MN319700 1899 bp DNA linear BCT 29-MAR-2020

DEFINITION Bacillus thuringiensis strain SP142 Cry toxin (cry2Ab) gene, partial cds.

ACCESSION MN319700

VERSION MN319700.1

KEYWORDS .

SOURCE Bacillus thuringiensis

ORGANISM [Bacillus thuringiensis](#)
Bacteria; Bacillota; Bacilli; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE 1 (bases 1 to 1899)

AUTHORS Le,N.T., Pham,N.B., Hoang,H., Le,L.T., Tran,L.H. and Chu,H.H.

TITLE identification and expression of a novel cry2Ab gene isolated from Bacillus thuringiensis strain SP142

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1899)

AUTHORS Le,N.T., Pham,N.B., Hoang,H., Le,L.T., Tran,L.H. and Chu,H.H.

TITLE Direct Submission

JOURNAL Submitted (15-AUG-2019) Plant cell Biotechnology, Institute of
Biotechnology, 18 Hoang Quoc Viet, Ha Noi 10000, Vietnam

COMMENT ##Assembly-Data-START##
Sequencing Technology :: Sanger dideoxy sequencing
##Assembly-Data-END##

FEATURES Location/Qualifiers

source 1..1899
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/strain="SP142"
/isolation_source="soil"
/db_xref="taxon:[1428](#)"
/country="Viet Nam"

[gene](#) 1..>1899
/gene="cry2Ab"

[CDS](#) 1..>1899
/gene="cry2Ab"
/codon_start=1
/transl_table=[11](#)
/product="Cry toxin"
/protein_id="[QIQ19560.1](#)"
/translation="MNSVLNSGRTTICDAYNVAAHDFSFQHKSLDTVQKEWTEWKKN
NHSLYLDPIVGTVASFLLKKVGSVVGKRISELRLNLIFFSGSTNLMQDILRETEKFLN
QRLNTDTLARVNAELTGLQANVEEFNRQVDNFLNPNRNAVPLSITSSVNTMQQLFLNR
LPQFQMQGYQLLLLPLFAQAANLHLSFIRDVILNADEWGISAATLRITYRDYLNKYTRD
YSNYCINTYQSAFRGLNTRLHDMLEFRTYMFLNVFEYVSIWSLFKYQSLLVSSGANLY
ASGSGPQQTQSFTSQDWPFLYSLFQVNSNYVLNGFSGARLTQTFPNIVGLPGTTTTHA
LLAARVNYSGGVSSGDIGASPFNQNFNCSTFLPPLLTPFVRSWLDGSDREGVATVTN
WQTESFETTLGLRSGAFTARGNSNYFPDYFIRNISGVPLVVRNEDLRRPLHYNEIRNI
ASPSGTPGGARAYMVSVHNRKNNIHAVHENGSMIHLAPNDYTGFTISPIHATQVNNQT
RTFISEKFGNQGDSLRFEQNNTTARYTLRGNGNSYNLYLRVSSIGNSTIRVTINGRVY
TATNVNTTTNNDGVNDNGARFSDINIGNVASSNSDVPLDINVTLNSGTQFDLMNIML
VPTNISPLY"

ORIGIN

1 atgaatagtg tattgaatag cggaagaact actatattgtg atgcgataa tgtagcggct
61 catgatccat ttagttttca acacaaatca ttagataccg tacaaaagga atggacggag
121 tggaaaaaaa ataatcatag tttataccta gatcctattg ttggaactgt ggctagtttt
181 ctgttaaaga aagtggggag tcttgttgga aaaaggatac taagtgagtt acggaattta
241 atatttccta gtgtagtac aaatctaag caagatattt taagagagac agaaaaatc
301 ctgaatcaaa gacttaatac agacactcct gcccggtgaa atgcggaatt gacagggctg
361 caagcaaatg tagaagagtt taatcgacaa gtagataatt ttttgaacce taaccgaaac
421 gctgttcctt tatcaataac ttcttcagtt aatacaatgc aacaattatt tctaaataga
481 ttaccccagt tccagatgca aggataccea ctggtattat tacctttatt tgcacaggca
541 gccaatctac atctttcttt tattagagat gtattctaa atgcagatga atggggaatt
601 tcagcagcaa cattacgtac gtatcgagat tacttgaaaa attatacaag agattactct

```

661 aactattgta taaatacgta tcaaagtgcg tttagagggt taaacacccg tttacacgat
721 atgttagaat ttagaacata tatgttttta aatgtatttg aatatgtatc tatctggctc
781 ttgtttaaat atcaaagcct tctagtatct tctggcgcta atttatatgc aagtggtagt
841 ggaccacagc agaccaatc atttacttca caagactggc catttttata ttctcttttc
901 caagttaatt caaattatgt gttaaattggc tttagtggcg ctgacttac gcagactttc
961 cctaataattg ttggtttacc tggactact acaactcacg cattgcttgc tgcaagggtc
1021 aattacagtg gaggagtttc gtctgggtgat ataggtgcat ctccgtttaa tcaaaatfff
1081 aattgtagca catttctccc ccattgtta acgccatttg ttaggagttg gctagattca
1141 ggttcagatc gggagggcgt tgccaccggt acaaattggc aaacagaatc ctttgagaca
1201 actttagggg taaggagtgg tgcttttaca gctcgcggtta attcaaacta tttcccagat
1261 tattttattc gtaatatffc tggagtctct ttagttgtta gaaatgaaga tttaagaaga
1321 ccgttacact ataatgaaat aagaaatata gcaagtcctt caggaacacc tggtgaggca
1381 cgagcttata tggatctgt gcataacaga aaaaataata tccatgctgt tcatgaaat
1441 ggttctatga ttcatttagc gccaaatgac tatacaggat ttactatttc gccgatacat
1501 gcaactcaag tgaataatca aacacgaaca tttatttctg aaaaatttgg aaatcaaggt
1561 gattctttaa ggtttgaaca aaacaacacg acagctcgtt atacgcttag agggaatgga
1621 aatagttaca atctttatff aagagtffc tcaataggaa attccactat tcgagttact
1681 ataaacggta ggtatatac tgctacaaat gttaatacta ctacaaataa cgatggagtt
1741 aatgataatg gagctcgttt ttcagatatt aatatcggta atgtagtagc aagtagtaat
1801 tctgatgtac cattagatat aaatgtaaca ttaaaactccg gtactcaatt tgatcttatg
1861 aatattatgc ttgtaccaac taatatttca ccactttat

```

//

❖ ***Dữ liệu gen cry1Na của chủng BD8.2:***

Bacillus thuringiensis serovar kurstaki strain BD8.2 insecticidal crystal protein Cry1Na (cry1Na) gene, partial cds

GenBank: MN725076.1

[FASTA Graphics](#)

[Go to:](#)

```

LOCUS      MN725076                1920 bp    DNA     linear   BCT 06-SEP-2020
DEFINITION Bacillus thuringiensis serovar kurstaki strain BD8.2 insecticidal
            crystal protein Cry1Na (cry1Na) gene, partial cds.
ACCESSION  MN725076
VERSION    MN725076.1
KEYWORDS   .
SOURCE     Bacillus thuringiensis serovar kurstaki
  ORGANISM Bacillus thuringiensis serovar kurstaki
            Bacteria; Bacillota; Bacilli; Bacillales; Bacillaceae; Bacillus;
            Bacillus cereus group.
REFERENCE  1 (bases 1 to 1920)
  AUTHORS  Thanh,L.T.M., Ngoc,L.T., Ha,T.T.T., Ngoc,N.B., Quyen,D.V.,
            Binh,N.D., Phuoc,M.H., Ha,C.H., Dong,N.V. and Hang,D.T.
  TITLE    Screening and indentification of genes coding for potential
            insecticidal crystal proteins against Etiella zinckenella larvae

```


from Vietnamese *Bacillus thuringiensis* strains

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1920)

AUTHORS Thanh, L.T.M.

TITLE Direct Submission

JOURNAL Submitted (21-NOV-2019) Center for Culture Collection and Genetic Resource Conservation of Microorganisms, Institute of Biotechnology, Hoang Quoc Viet, Ha Noi, Ha Noi 100000, Viet Nam

COMMENT ##Assembly-Data-START##
 Assembly Method :: BioEdit software v. BioEdit 7.2
 Sequencing Technology :: AEI PRISM@ 3700 Genetic Analyzer
 ##Assembly-Data-END##

FEATURES Location/Qualifiers

source 1..1920
 /organism="Bacillus thuringiensis serovar kurstaki"
 /mol_type="genomic DNA"
 /strain="BD8.2"
 /serovar="kurstaki"
 /isolation_source="soil"
 /db_xref="taxon:29339"
 /country="Viet Nam: Dien Bien province"

gene 34..>1920
 /gene="cry1Na"

CDS 34..>1920
 /gene="cry1Na"
 /note="toxic to *Etiella zinckenella* larvae"
 /codon_start=1
 /transl_table=11
 /product="insecticidal crystal protein Cry1Na"
 /protein_id="QNN26304.1"
 /translation="MSELKGNFKKSTNRTCCLLKIINIGGRGMNSKEHDYLVKVCNDLS
 DANINMERFDKNDALIEIGMSIVSELIGMIPGGTALQFVFNQLWSRLGDSGWNAMFEHV
 EELIDTKIEGYAKNKALSELAGIQRNLETYIQLRNEWENDIENSKAQGKVANYYESLE
 QAVERSMPQFAVGNFEVPLLTVYVQAANLHLLLRDVSVYGKRWGSEQKIKIYYDKQ
 IKYTHEYTNHCVNWNKGLERLKNKGSSYQDWYNYNRFREMTLTVLDIVALFPHYDV
 QTYPIITVAQLTREVVYTDPLLNFNPKLHSVSQLPSFSDMENATIRTPHLMFLRMLTI
 YTDWYSVGRNYYWGGHRVTSYHVGGENIRSPLYGREANQEVPRDFYFYGPVFKTSLKP
 TLRPLQQPAPAPPFNLRSLGVEFHTPTGSFMYRERGSVDSFNELPFNPVGLPHKVY
 SHRLCHATFVRKSGTPYLTTGAIFSWTHRSAEETNTIESNIITQIPLVKAYQIGSGTT
 VRKGPFGFTGGDILRRTGPGTFGDMRININAPLSQRYRVRIRYASTTDLQFVTSINGTT
 INIGNFPKTINNLTGSEGYRTVSFSTPFSSNAQSIFRLGIQAFSGVQEVYVDKIE
 FIPVE"

ORIGIN

1 agcctcattt taaaattatt ccttttaaat gtaatgagtg aattgaaggg gaattttaag
 61 aaaagtacta atcgaacttg ttgtttgcta aaaataataa atataggagg aagaggtatg

```

121 aattcaaagg aacatgatta tctaaaagtt tgtaatgatt taagtgacgc caatattaat
181 atggagcggg ttgataagaa tgatgactg gaaattggta tgtccattgt atctgaactt
241 attggtatga ttccaggcgg aacagctttg caatttgtgt ttaatcaatt gtgggtctcgt
301 ttagggtgatt ctggatggaa tgcgttcagtg gaacatgtgg aggaattaat tgatactaaa
361 atagaagggg atgcaaaaaa taaagcctta tctgaattag caggatataca aagaaacctt
421 gaaacatata tacaattacg taatgaatgg gaaaatgata ttgaaaactc aaagggtcaa
481 ggtaaggtag ctaattacta tgaagtctt gagcaggcgg ttgaaaggag tatgcctcaa
541 tttgcagtgg ggaatthttga agtaccactt ttaactgtct atgtgcaagc tgctaactct
601 catttattat tattaagaga tgthtcagtt tatggaaagc gttggggatg gtcggagcag
661 aaaattaaaa tttattatga taaacagatt aagtataccc atgaatacac aaatcattgt
721 gtaaatgggt ataataaagg acttgagaga ttaaaaaata aaggttcttc ttatcaagat
781 tggtaacaatt ataatcgttt ccgtagagaa atgactctta ctgtthttaga tatcgttgct
841 ttattcccgc actatgatgt acaaaccttat ccaataacaa ccgttgctca gctaacaagg
901 gaagthttata cggatccttt acttaatthtt aatcctaaat tacattctgt gtctcaatta
961 cctagthttta gtgacatgga aaatgcaaca attagaactc cacatctgat ggaatthttta
1021 agaatgctaa caatthtatac agatthgtat agtgthggaa gaaactatta thggggagga
1081 catcgctgta cgtcttacca tgtaggagga gagaatataa gatcacctct atatggtaga
1141 gaggcaaatc aagagthtcc tagagatthtt taththttatg gaccgthttt taagacgtta
1201 tcaaagccga ctctaagacc attacagcag cctgcaccag ctctccttht taatthacgt
1261 agcttagagg gagtagaatt ccacactcct acaggtagtt ttatgtatcg tgaaagagga
1321 tccgtagatt cththtaatga gttaccgctt thtaatccag thgggttacc tcataagga
1381 tacagtcacc gthttatgtca tgcaacgtht gthctgtaaat ctgggacccc thatthtaaca
1441 acaggtgcc a thththcttg gacacatcgt agtgctgaag aaaccaatac aatthaatca
1501 aatattatta cgcaaatccc gthtagtaaaa gcatatcaaa thgggtcagg cactactgta
1561 aggaaaggac caggatthac aggagthggat atactthcga gaacaggtcc tggaacatth
1621 ggagatatga gaataaatat taatgcacca thtatctcaa gatatcgtgt aaggtatcgt
1681 tatgctthta cgacagattht acaatthgtc acgagtatta atgggaccac cattaatatt
1741 ggtaactthc caaaaactat taataatcta aatactthtag gthctgaggg ctatagaaca
1801 gtatcgttht gtactccatt tagththctca aatgcacaaa gcatatthtag attaggtata
1861 caagcattht ctggagthta agaagthttat gthggataaaa thgaattht tctcgtthga

```

//

❖ *Dữ liệu gen cry2Ah của chủng Đ6.1:*

Bacillus thuringiensis serovar aizawai strain D6.1 insecticidal crystal protein Cry2Ah1 (cry2Ah1) gene, complete cds

GenBank: MN725074.1

[FASTA Graphics](#)

[Go to:](#)

```

LOCUS      MN725074                1899 bp    DNA        linear    BCT 06-SEP-2020
DEFINITION Bacillus thuringiensis serovar aizawai strain D6.1 insecticidal
            crystal protein Cry2Ah1 (cry2Ah1) gene, complete cds.
ACCESSION  MN725074
VERSION    MN725074.1
KEYWORDS   .

```

SOURCE Bacillus thuringiensis serovar aizawai

ORGANISM [Bacillus thuringiensis serovar aizawai](#)
 Bacteria; Bacillota; Bacilli; Bacillales; Bacillaceae; Bacillus;
 Bacillus cereus group.

REFERENCE 1 (bases 1 to 1899)

AUTHORS Thanh,L.T.M., Ngoc,L.T., Ha,T.T.T., Ngoc,N.B., Quyen,D.V.,
 Binh,N.D., Ha,C.H., Dong,N.V. and Hang,D.T.

TITLE Detection of two potential genes encoding insecticidal crystal
 proteins toxic to Etiella zinckenella larvae from Bacillus
 thuringiensis serovar aizawai D6.1 strain in Vietnam

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1899)

AUTHORS Thanh,L.T.M.

TITLE Direct Submission

JOURNAL Submitted (21-NOV-2019) Center for Culture Collection and Genetic
 Resource Conservation of Microorganisms, Institute of
 Biotechnology, Hoang Quoc Viet, Ha Noi, Ha Noi 100000, Viet Nam

COMMENT ##Assembly-Data-START##
 Assembly Method :: BioEdit software v. BioEdit 7.2
 Sequencing Technology :: AEI PRISM@ 3700 Genetic Analyzer
 ##Assembly-Data-END##

FEATURES Location/Qualifiers

source 1..1899
 /organism="Bacillus thuringiensis serovar aizawai"
 /mol_type="genomic DNA"
 /strain="D6.1"
 /serovar="aizawai"
 /isolation_source="soil"
 /db_xref="taxon:1433"
 /country="Viet Nam: Tuyen Quang province"

[gene](#) 1..1899
 /gene="cry2Ah1"

[CDS](#) 1..1899
 /gene="cry2Ah1"
 /note="toxic to Etiella zinckenella larvae"
 /codon_start=1
 /transl_table=11
 /product="insecticidal crystal protein Cry2Ah1"
 /protein_id="QNN26302.1"
 /translation="MNNVLNSGRATNGDAYNVVAHDPFSFQHKSLDTIQEEMWKKD
 NHILYVDPIVGTVASFLLKKVGSIVGKRILSELRNLIFFPSGSTNLMQDILRETEKFLN
 QRLNTDTLARVNAELTGLQANVEEFNRQVDNFLNPNRNAVPLSITSSVNTMQQLFLNR
 LPQFQMQGYQLLLLPLFAQAANLHLSFIRDVILNADEWGISAATLRTYQNHLRNYTRE
 YSNYCITTYQTAFRGLNTRLHDMLEFRTYMFLNVFEYVSIWSLFKYQSLLVSSGANLY
 ASGSGPQQTQSFTSQDWPFLYSLFQVNSNYVLNGFSGARLTQTFPNIVGLPGTTTTHA

LLAARVNYSGGVSSGDIGAVFNQNFSCSTFLPPLLPFVRSWLDSGSDRGGINTVTNW
 QTESFETTLGLRSGAFTARGNSNYFPDYFIRNISGVPLVVRNEDLRRPLHYNQIRNIE
 SPSGTPGGLRAYMVSVHNRKNNIYAVHENGTMIHLPEDYTGFTISPIHATQVNNQTR
 TFISEKFGNQGDSLRFEQSNTTARYTLRGNGNSYNLYLRVSSIGNSTIRVTINGRVYT
 ASNVNTTTNNDGVNDNGARFSDINIGNVVASDNTNVPLDINVTLNSGTQFELMNIMEV
 PTNISPLY"

ORIGIN

1 atgaataatg tattgaatag cggaagagct actaatggtg atgcgtataa tgtagtggct
 61 catgatccat ttagttttca acataaatca ttagatacca tacaagaaga atggatggag
 121 tggaaaaaag ataatacatat tttatatgta gatcctattg ttggaactgt ggctagcttt
 181 cttttaaaga aagtggggag tcttggttga aaaagaatat taagtgagtt acggaattta
 241 atatttccta gtggcagtac aaatctaag caagatattt taagagagac agaaaaattc
 301 ctgaatcaaa gacttaatac agacactcct gcccggtgaa atgcggaatt gacagggctg
 361 caagcaaag tagaagagtt taatcgacaa gtagataatt ttttgaacce taaccgaaat
 421 gctgttcctt tatcaataac ttcttcagtt aatacaatgc agcaattatt tctaaataga
 481 ttaccccagt ttcagatgca aggataccaa ttggttattat tacctttatt tgcacaggca
 541 gccaatttac atctttcttt tattagagat gttattctta atgcagatga atgggggaatt
 601 tcagcagcaa cattacgtac gtatcaaaat cacctgagaa attatacaag agagtactct
 661 aattattgta taactacgta tcaaactgcg tttagagggt taaacaccg tttacacgat
 721 atgttagaat ttagaacata tatgttttta aatgtatttg aatatgtatc tatctggtcg
 781 ttgtttaaat atcaaagcct tctagtatct tctggcgcta atttatatgc aagtggtagt
 841 ggaccacagc agaccaatc atttacttca caagactggc catttttata ttctcttttc
 901 caagttaatt caaattatgt gttaaatggc tttagtgggc ctgacttac gcgactttc
 961 cctaataattg ttggtttacc tggactact acaactcacg cattgcttgc tgcaagggtc
 1021 aattacagtg gaggagtttc gctcgggtgat ataggcgctg tgtttaatca aaattttagt
 1081 tgtagtacat ttctcccacc tttgttaaca ccatttgta gaagttggct agattcaggt
 1141 tcagatcggg gggggattaa taccgttacc aattggcaaa cagaatcctt tgagacaact
 1201 ttaggtttaa ggagtggcgc ttttacagct cgaggtaatt caaactatctt cccagattat
 1261 tttatccgta atatctctgg agttccttta gttgtagaa atgaagattt aagaagaccg
 1321 ttacactata atcaaataag aaatatagaa agtccttcag gaacacctgg tggattacga
 1381 gcttatatgg tatctgtgca taacagaaaa aataatatct atgccgttca tgaaaatggt
 1441 actatgattc atttagcgcc ggaagattat acaggattta ctatatcgcc gatacatgca
 1501 actcaagtga ataatacaac gcgaacattt atttctgaaa aatttggaat tcaaggtgat
 1561 tccttaagat ttgaacaaag caacacgaca gctcgttata cccttagagg gaatggaaat
 1621 agttacaatc tttatttaag agtatcttca ataggaaatt ccaactattcg agttactata
 1681 aacggtagag tttatactgc ttcaaagtt aataactacta caaataacga tggagttaat
 1741 gataatggag ctcgtttttc agatattaat atcggtaatg tagtagcaag tgataatact
 1801 aatgtaccgt tagatataaa tgtgacatta aattcgggta ctcaatttga gcttatgaat
 1861 attatggttg ttccaactaa tatttcacca ctttattaa

//

Hình phụ lục 9. So sánh trình tự gen *cry2Ab39* trước và sau khi cải biến mã

```

      10      20      30      40      50      60      70      80
cry2Ab39  -----ATGAATAGTGTATT
cry2Ab39opt  M N S V L
      90      100     110     120     130     140     150     160
cry2Ab39  GAATAGCGGAAGAACTACTATTGTGATGCGGTATAATGTAGCGGCTCATGATCCATTAGTTTCAACACAATCATTAG
cry2Ab39opt  N S G R T T I C D A Y N V A A H D P F S F Q H K S L
      170     180     190     200     210     220     230     240
cry2Ab39  ATACCGTACAAAAGGAATGGACGGAGTGGAAAAAATAATCATAGTTTATACCTAGATCCTATTGTTGGAACGTGGGCT
cry2Ab39opt  D T V Q K E W T E W K K N N H S L Y L D P I V G T V A
      250     260     270     280     290     300     310     320
cry2Ab39  AGTTTCTGTAAAGAAAGTGGGGAGTCTTGTGGAAAAGGATACTAAGTGAGTTACGGAAATTTAATATTCCTAGTGG
cry2Ab39opt  S F L L K K V G S L V G K R I L S E L R N L I F P S G
      330     340     350     360     370     380     390     400
cry2Ab39  TAGTACAAATCTAATGCAAGATATTTAAGAGAGACAGAAAAATCCTGAATCAAAGACTTAATACAGACACTCTTGCCC
cry2Ab39opt  S T N L M Q D I L R E T E K F L N Q R L N T D T L A
      410     420     430     440     450     460     470     480
cry2Ab39  GTGTAATGCGGAATTGACAGGGCTGCAAGCAATGTAGAAGAGTTAATCGACAAGTAGATAAATTTTGAACCCTAAC
cry2Ab39opt  R V N A E L T G L Q A N V E E F N R Q V D N F L N P N
      490     500     510     520     530     540     550     560
cry2Ab39  CGAAACGCTGTTCTTTATCAATACTTCTTCAAGTAAATACATGCAACAATTTATCTAAATAGATATCCCCAGTCCA
cry2Ab39opt  R N A V P L S I T S S V N T M Q Q L F L N R L P Q F Q
      570     580     590     600     610     620     630     640
cry2Ab39  GATGCAAGGATACCAACTGTTATTATTACCTTTATTGACAGGCAGCCAATTTACATCTTTCTTTATTAGAGATGTTA
cry2Ab39opt  M Q G Y Q L L L L P L F A Q A A N L H L S F I R D V
      650     660     670     680     690     700     710     720
cry2Ab39  TTCTAAATGAGATGAATGGGAATTTCAAGCAACATTACGTACGTATCGAGATTACTTGAAAAATTATACAAGAGAT
cry2Ab39opt  I L N A D E W G I S A A T L R T Y R D Y L K N Y T R D
      730     740     750     760     770     780     790     800
cry2Ab39  TACTCTAATGTTGATAAATACGTATCAAAGTGCCTTTAGAGGTTAAACACCCGTTTACACGATATGTTAGAATTTAG
cry2Ab39opt  Y S N Y C I N T Y Q S A F R G L N T R L H D M L E F R
      810     820     830     840     850     860     870     880
cry2Ab39  AACATATATGTTTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG
cry2Ab39opt  T Y M F L N V F E Y V S I W S L F K Y Q S L L V S S
      890     900     910     920     930     940     950     960
cry2Ab39  GCGCTAATTTATATGCAAGTGGTAGTGGACCCAGCAGACCCAATCATTACTTTCACAAGACTGGCCATTTTATATTTCT
cry2Ab39opt  G A N L Y A S G S G P Q Q T Q S F T S Q D W P F L Y S
      970     980     990     1000  1010  1020  1030  1040
cry2Ab39  CTTTTCAAGTTAATTCAAATTATGTGTTAAATGGCTTTAGTGGCGCTAGACTTACGCAGACTTTCCCTAATATTGTTGG

```

```

L F Q V N S N Y V L N G F S G A R L T Q T F P N I V G
cry2Ab39opt .....G..C..T..C.....G..C..A...TC...A....GT.G..A.....C..C..G..
L F Q V N S N Y V L N G F S G A R L T Q T F P N I V G
      1050      1060      1070      1080      1090      1100      1110      1120
cry2Ab39 TTTACCTGGTACTACTACAACCTCAGCATTGCTTGTCTGCAAGGGTCAATTACAGTGGAGGAGTTTCGTCTGGTGATATAG
cry2Ab39opt L P G T T T T H A L L A A R V N Y S G G V S S G D I
      1130      1140      1150      1160      1170      1180      1190      1200
cry2Ab39 GTGCATCTCCGTTTAAACAAAATTTAATTGTAGCACATTTCTCCCCCATTGTTAACGCCATTGTTAGGAGTTGGCTA
cry2Ab39opt G A S P F N Q N F N C S T F L P P L L T P F V R S W L
      1210      1220      1230      1240      1250      1260      1270      1280
cry2Ab39 GATTCAGGTTTCAGATCGGGAGGGCGTTGCCACCGTTACAATTTGGCAAACAGAATCCTTTGAGACAACCTTTAGGTTAAG
cry2Ab39opt .D.S.G.S.D.R.E.G.V.A.T.V.T.N.W.Q.T.E.S.F.E.T.T.L.G.L.R
      1290      1300      1310      1320      1330      1340      1350      1360
cry2Ab39 GAGTGGTGTCTTTTACAGCTCGCGTAATTCAAACTATTTCAGATTATTTTATTCGTAATATTCTGGAGTTCCCTTTAG
cry2Ab39opt S G A F T A R G N S N Y F P D Y F I R N I A S P S G T P G
      1370      1380      1390      1400      1410      1420      1430      1440
cry2Ab39 TTGTAGAAATGAAGATTTAAGAAGACCGTTACACTATAATGAAATAAGAAATATAGCAAGTCCTCAGGAACACCTGGT
cry2Ab39opt V V R N E D L R R P L H Y N E I R N I A S P S G T P G
      1450      1460      1470      1480      1490      1500      1510      1520
cry2Ab39 GGAGCAGGCTTATATGGTATCTGTGCATAACAGAAAAATAATATCCATGCTGTTCATGAAATGGTTCTATGATTCA
cry2Ab39opt G A R A Y M V S V H N R K N N I H A V H E N G S M I H
      1530      1540      1550      1560      1570      1580      1590      1600
cry2Ab39 TTTAGCGCCAAATGACTATACAGGATTTACTATTTTCGCGGATACATGCAACTCAAGTGAATAATCAACACGAAACATTTA
cry2Ab39opt L A P N D Y T G F T I S P I H A T Q V N N Q T R T F
      1610      1620      1630      1640      1650      1660      1670      1680
cry2Ab39 TTTCTGAAAATTTGGAATCAAGGTGATTTCTTAAAGTTTGAACAAAACAACACGACGCTCGTTATACGCTTAGAGGG
cry2Ab39opt I S E K F G N Q G D S L R F E Q N N T T A R Y T L R G
      1690      1700      1710      1720      1730      1740      1750      1760
cry2Ab39 AATGGAATAGTTACAATCTTTATTTAAGATTTCTTCAATAGGAAATCCACTATTCGAGTTACTATAACCGTAGGGT
cry2Ab39opt N G N S Y N L Y L R V S S I G N S T I R V T I N G R V
      1770      1780      1790      1800      1810      1820      1830      1840
cry2Ab39 ATATACTGTACAAATGTTAATACTACTACAATAACGATGGAGTTAATGATAATGGAGCTCGTTTTTCAGATATTAATA
cry2Ab39opt Y T A T N V N T T T N N D G V N D N G A R F S D I N
      1850      1860      1870      1880      1890      1900      1910      1920
cry2Ab39 TCGGTAATGTAGTAGCAAGTAGTAATCTGTATGATGACCATTAGATATAAATGTAACATAAACTCCGGTACTCAATTTGAT
cry2Ab39opt I G N V V A S S N S D V P L D I N V T L N S G T Q F D
      1930      1940      1950      1960      1970      1980      1990      2000
cry2Ab39 CTTATGAATATTATGCTTGTACCACTAATATTTACCACCTTTATTAA-----
cry2Ab39opt L M N I M L V P T N I S P L Y *
      GCTTTGA

```