

**BỘ GIÁO DỤC
VÀ ĐÀO TẠO**

**VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM**

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



Vũ Thị Trinh

**PHÂN TÍCH BIỂU HIỆN MỘT SỐ GEN LIÊN QUAN ĐẾN
CON ĐƯỜNG SINH TỔNG HỢP GINSENOSE Ở
SÂM NGỌC LINH (*Panax vietnamensis* Ha et Grushv.)**

**LUẬN VĂN THẠC SĨ
SINH HỌC THỰC NGHIỆM**

VŨ THỊ TRINH

**SINH HỌC
THỰC NGHIỆM**

2023

Hà Nội - 2023

**BỘ GIÁO DỤC
VÀ ĐÀO TẠO**

**VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM**

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



Vũ Thị Trinh

**PHÂN TÍCH BIỂU HIỆN MỘT SỐ GEN LIÊN QUAN ĐẾN CON
ĐƯỜNG SINH TỔNG HỢP GINSENOSE Ở
SÂM NGỌC LINH (*Panax vietnamensis* Ha et Grushv.)**

Chuyên ngành: Sinh học thực nghiệm
Mã số: 8420114

**LUẬN VĂN THẠC SĨ
NGÀNH SINH HỌC**

NGƯỜI HƯỚNG DẪN KHOA HỌC:
PGS. TS. Lê Thị Thu Hiền

Hà Nội - 2023

LỜI CAM ĐOAN

Tôi xin cam đoan đề tài nghiên cứu trong luận văn này là công trình nghiên cứu của tôi dựa trên những tài liệu, số liệu do chính tôi tự tìm hiểu và nghiên cứu. Chính vì vậy, các kết quả nghiên cứu đảm bảo trung thực và khách quan nhất. Đồng thời, kết quả này chưa từng xuất hiện trong bất cứ một nghiên cứu nào. Các số liệu, kết quả nêu trong luận văn là trung thực nếu sai tôi hoàn toàn chịu trách nhiệm trước pháp luật.

Tác giả luận văn

Vũ Thị Trinh

LỜI CẢM ƠN

Lời đầu tiên, em xin được bày tỏ lòng biết ơn chân thành đến toàn bộ Quý Thầy Cô của Học viện Khoa học và Công nghệ vì đã truyền tải những kiến thức quý báu, những kinh nghiệm thực tế cho em. Em xin chân thành cảm ơn Ban Lãnh đạo, Phòng Đào tạo, các phòng chức năng của Học viện Khoa học và Công nghệ đã tạo điều kiện thuận lợi cho em trong suốt quá trình học tập.

Với tất cả sự trân trọng và quý mến, em xin gửi lời cảm ơn sâu sắc nhất tới PGS. TS. Lê Thị Thu Hiền, người đã tận tình giúp đỡ, định hướng cách tư duy và làm việc khoa học. Cảm ơn cô đã tận tình chỉ bảo, hướng dẫn và tạo điều kiện thuận lợi cho em trong suốt quá trình học tập, nghiên cứu và thực hiện luận văn tốt nghiệp. Đó là những trợ giúp hết sức quý báu, không chỉ trong quá trình thực hiện luận văn mà còn là hành trang tiếp bước cho em trong quá trình học tập và nghiên cứu sau này.

Em cũng xin được gửi lời cảm ơn tới các thành viên của Phòng Đa dạng sinh học hệ gen, Viện Nghiên cứu hệ gen, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, đặc biệt cảm ơn TS. Huỳnh Thị Thu Huệ, TS. Hà Hồng Hạnh, TS. Vũ Tuấn Nam, ThS. Lưu Hàn Ly, ThS. Nguyễn Nhật Linh, ThS. Đào Quang Hà, CN. Nguyễn Văn Giang và CN. Nguyễn Thị Bích Ngọc, vì đã hỗ trợ em rất nhiều trong quá trình nghiên cứu và học hỏi, để em có thể hoàn thiện luận văn này.

Nghiên cứu này được thực hiện với sự hỗ trợ từ đề tài “Giải trình tự và phân tích hệ phiên mã (transcriptome) ở sâm Ngọc Linh (*Panax vietnamensis* Ha et Grushv.).

Cuối cùng, em vô cùng biết ơn gia đình và bạn bè đã khích lệ, động viên và luôn bên em trong quá trình hoàn thành luận văn này. Em xin chân thành cảm ơn.

MỤC LỤC

DANH MỤC CÁC TỪ VIẾT TẮT.....	i
DANH MỤC BẢNG.....	ii
DANH MỤC HÌNH.....	iii
MỞ ĐẦU.....	1
CHƯƠNG 1. TỔNG QUAN TÀI LIỆU.....	3
1.1 CHI NHÂN SÂM VÀ SÂM NGỌC LINH.....	3
1.1.1 Chi Nhân sâm.....	3
1.1.2 Sâm Ngọc Linh.....	4
1.2 NGHIÊN CỨU GIẢI TRÌNH TỰ HỆ GEN Ở CÁC LOÀI THUỘC CHI <i>PANAX</i> 9	
1.3 NGHIÊN CỨU BIỂU HIỆN CỦA CÁC GEN CHỨC NĂNG Ở CÁC LOÀI THUỘC CHI <i>PANAX</i>	11
CHƯƠNG 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU.....	13
2.1. VẬT LIỆU NGHIÊN CỨU.....	13
2.1.1. Các mẫu sâm nghiên cứu.....	13
2.1.2. Các cặp môi sử dụng.....	13
2.1.3. Hóa chất, thiết bị.....	15
2.2. PHƯƠNG PHÁP NGHIÊN CỨU.....	15
2.2.1. Thu mẫu.....	15
2.2.2. Tách chiết và kiểm tra chất lượng RNA tổng số.....	16
2.2.3. Tổng hợp cDNA.....	17
2.2.4. Xây dựng dữ liệu các gen liên quan đến con đường sinh tổng hợp ginsenoside....	17
2.2.5. Thiết kế các cặp môi.....	18
2.2.6. Nhân gen quan tâm sử dụng phương pháp reverse transcription PCR.....	18
2.2.7. Xác định trình tự gen quan tâm sử dụng phương pháp Sanger.....	18
2.2.8. Nghiên cứu biểu hiện một số gen chức năng bằng realtime PCR.....	20
CHƯƠNG 3. KẾT QUẢ VÀ THẢO LUẬN.....	21
3.1. KẾT QUẢ TÁCH CHIẾT, TINH SẠCH, KIỂM TRA CHẤT LƯỢNG, ĐỊNH LƯỢNG RNA TỔNG SỐ TỪ MÔ LÁ, THÂN RỄ SÂM NGỌC LINH.....	21
3.2. KẾT QUẢ XÂY DỰNG CON ĐƯỜNG SINH TỔNG HỢP GINSENOSE GIẢ ĐỊNH Ở SÂM NGỌC LINH.....	22
3.2.1. Thống kê sơ bộ số lượng unigene gắn với mỗi gen.....	22
3.2.2. Thống kê dữ liệu các unigene cụ thể.....	25

3.3. KẾT QUẢ PHÂN TÍCH BIỂU HIỆN CÁC GEN LIÊN QUAN ĐẾN CON ĐƯỜNG SINH TỔNG HỢP GINSENOSE Ở MẪU LÁ SÂM NGỌC LINH 1 VÀ 4 NĂM TUỔI.....	32
3.4. KẾT QUẢ PHÂN TÍCH BIỂU HIỆN CÁC GEN LIÊN QUAN ĐẾN CON ĐƯỜNG SINH TỔNG HỢP GINSENOSE Ở MẪU THÂN RỄ SÂM NGỌC LINH 1 VÀ 4 NĂM TUỔI.....	36
3.5. KẾT QUẢ SO SÁNH MỨC BIỂU HIỆN CỦA CÁC GEN LIÊN QUAN ĐẾN TỔNG HỢP GINSENOSE GIỮA MÔ LÁ VÀ THÂN RỄ SÂM NGỌC LINH Ở HAI ĐỘ TUỔI	38
DANH MỤC CÔNG TRÌNH CỦA TÁC GIẢ	41
TÀI LIỆU THAM KHẢO	42
PHỤ LỤC	48
Phụ lục 3.1 Ma trận so sánh cDNA của một số gen liên quan đến con đường sinh tổng hợp ginsenoside phân lập từ sâm Ngọc Linh với trình tự tham chiếu và trình tự trích xuất từ dữ liệu transcriptome	49

DANH MỤC CÁC TỪ VIẾT TẮT

Chữ viết tắt	Tên đầy đủ (tiếng Anh)	Tên đầy đủ (tiếng Việt)
Bp	Base pair	Cặp base
cDNA	Complementary DNA	DNA bổ sung
DNA	Deoxyribonucleic acid	Acid deoxyribonucleic
dNTPs	Deoxyribonucleotide triphosphates	Deoxyribonucleotide triphosphates
Đtg	<i>et al.</i>	Đồng tác giả
ESTs	Expressed sequence tags	Các trình tự biểu hiện
IUCN	International Union for Conservation of Nature and Natural Resources	Liên minh Quốc tế Bảo tồn thiên nhiên và Tài nguyên thiên nhiên
Kb	Kilo base	Kilo base (1.000 bp)
LSC	Large single-copy region	Vùng bản đơn lớn
NCBI	National Center for Biotechnology Information	Trung tâm Tin sinh học Quốc gia
ORF	Open reading frame	Khung đọc mở
PCR	Polymerase chain reaction	Phản ứng chuỗi polymerase
qPCR	Quantitative polymerase chain reaction	Phản ứng chuỗi polymerase định lượng
RNA	Ribonucleic Acid	Acid ribonucleic
SSR	Simple sequence repeat	Trình tự lặp lại đơn giản
UTR	Untranslated region	Vùng không dịch mã

DANH MỤC BẢNG

Bảng 2. 1 Các cặp môi phục vụ nhân gen quan tâm sử dụng RT-PCR	13
Bảng 2. 2. Các cặp môi phục vụ nghiên cứu biểu hiện của các gen sử dụng realtime PCR	14
Bảng 3. 1 Nồng độ và độ sạch (A260/A280) của một số mẫu RNA sâm Ngọc Linh sau tách chiết và tinh sạch	22
Bảng 3. 2 Thống kê số lượng và mức độ biểu hiện tương đối của các unigene tham gia vào con đường sinh tổng hợp ginsenoside từ kết quả giải trình tự hệ gen phiên mã sâm Ngọc Linh	23
Bảng 3.3 Dữ liệu unigene của một số gen liên quan đến con đường sinh tổng hợp ginsenoside	25
Bảng 3.4 So sánh mức độ biểu hiện của một số gen liên quan đến con đường sinh tổng hợp ginsenoside ở mô lá và thân rễ sâm Ngọc Linh 1 năm tuổi và 4 năm tuổi.....	34

DANH MỤC HÌNH

Hình 1. 1 Các loài tự nhiên thuộc chi <i>Panax</i> ở Việt Nam	5
Hình 1. 2 Sơ đồ giản lược quá trình sinh tổng hợp các ginsenoside ở thực vật	8
Hình 2. 1 Giải trình tự DNA bằng phương pháp Sanger	19
Hình 3. 1 Kiểm tra kết quả tách chiết và tinh sạch RNA tổng số trên gel agarose từ các mẫu lá (A) và thân rễ sâm Ngọc Linh (B).....	21
Hình 3. 2 Tổng quan con đường sinh tổng hợp ginsenoside giả định ở loài sâm Ngọc Linh	24
Hình 3. 3 Kết quả khuếch đại một số gen liên quan đến sinh tổng hợp ginsenoside ở sâm Ngọc Linh	29
Hình 3. 4 Tối ưu hóa mỗi khuếch đại các gen <i>EF</i> , <i>AACT</i> , <i>IDI</i> , <i>SS</i> , <i>SE</i>	34
Hình 3. 5 Biểu hiện của các gen liên quan đến sinh tổng hợp ginsenoside ở mô lá sâm Ngọc Linh	36
Hình 3. 6 Biểu hiện của các gen liên quan đến sinh tổng hợp ginsenoside ở mô thân rễ sâm Ngọc Linh	38
Hình 3. 7 So sánh mức biểu hiện của các gen liên quan đến sinh tổng hợp ginsenoside ở mô lá và thân rễ sâm Ngọc Linh 1 và 4 năm tuổi	39

MỞ ĐẦU

Sâm Ngọc Linh là loài dược liệu quý thuộc chi Nhân sâm (*Panax L.*) có hàm lượng saponin khung dammaran cao nhất (khoảng 12-15%) và hàm lượng triterpene nhiều nhất so với các loài khác thuộc chi *Panax* trên thế giới. Sâm Ngọc Linh có đến 52 loại saponin, 18 amino acid, 20 chất khoáng vi lượng, 0,1% tinh dầu. Sâm Ngọc Linh có nhiều công dụng đặc biệt như chống stress, chống trầm cảm, giảm lo âu, chống oxy hóa, kích thích hệ miễn dịch, tăng khả năng hồi phục máu, tăng khả năng vận động, chống suy nhược cơ thể [1-2]. Từ năm 1985 đến nay, đã có nhiều công trình nghiên cứu công bố về hóa học, hoạt tính sinh học và cấu trúc hóa học của các hợp chất tiêu biểu phân lập từ sâm Ngọc Linh [3-7]. Những kết quả nghiên cứu này khẳng định giá trị rất quý của sâm Ngọc Linh so với các loại nhân sâm khác trên thị trường. Từ những năm 1990, sâm Ngọc Linh trong tự nhiên hầu như đã bị khai thác cạn kiệt và có nguy cơ tuyệt chủng, trở thành một loài thực vật có tên trong sách đỏ Việt Nam [5]. Theo tiêu chuẩn phân hạng trong Danh lục Đỏ của IUCN (1994 và 2001), sâm Ngọc Linh ở trong tình trạng “Cực kỳ nguy cấp”. Trước tình hình đó, Việt Nam đã có những chủ trương, chính sách đúng đắn nhằm khôi phục, bảo tồn và phát triển cây sâm Ngọc Linh. Nhiều công trình nghiên cứu nhằm điều tra, phân loại, xác định vùng phân bố, nghiên cứu thành phần hóa học và dược tính, đánh giá đa dạng di truyền, tìm kiếm mã vạch DNA, xây dựng mối quan hệ phát sinh chủng loại, bảo tồn, nhân giống, sản xuất giá thể, kỹ thuật canh tác, gieo trồng, bảo vệ thực vật, chế biến, bảo quản sâm Ngọc Linh đã và đang được thực hiện. So với các loài khác trong chi Nhân sâm, các nghiên cứu ở mức độ phân tử và *in silico* liên quan đến hệ gen phiên mã cũng như các gen liên quan đến sinh tổng hợp ginsenoside vẫn còn rất ít.

Trong nghiên cứu này dữ liệu hệ gen phiên mã của sâm Ngọc Linh được phân tích và trên cơ sở các thông tin thu được, phương pháp real time PCR được sử dụng để đánh giá mức độ biểu hiện của các gen quan tâm có liên quan đến quá trình sinh tổng hợp ginsenoside ở lá và thân rễ của cây sâm Ngọc Linh. Kết quả góp phần cung cấp cơ sở khoa học đánh giá vai trò của yếu tố di truyền đến sự sinh tổng hợp ginsenoside ở cây sâm Ngọc Linh. Nghiên cứu **“Phân tích biểu hiện một số gen liên quan đến con đường sinh**

tổng hợp ginsenoside ở sâm Ngọc Linh (*Panax vietnamensis* Ha et Grushv.) ” đã được thực hiện với mục tiêu và những nội dung cụ thể như sau:

1. MỤC TIÊU:

Đề tài đặt mục tiêu xác định được các gen chính tham gia vào con đường sinh tổng hợp ginsenoside ở sâm Ngọc Linh và đánh giá được mức độ biểu hiện của một số gen này ở mô lá, thân rễ của sâm Ngọc Linh 1 và 4 năm tuổi.

2. NỘI DUNG:

a) Xác định các gen liên quan đến con đường sinh tổng hợp ginsenoside ở sâm Ngọc Linh thông qua phân tích dữ liệu hệ gen phiên mã thu được từ giải trình tự gen thế hệ mới;

b) Xây dựng con đường sinh tổng hợp ginsenoside giả định ở sâm Ngọc Linh;

c) Nhân bản và xác định trình tự một số gen liên quan đến con đường sinh tổng hợp ginsenoside ở sâm Ngọc Linh sử dụng phương pháp Sanger;

d) Phân tích biểu hiện các gen liên quan đến con đường sinh tổng hợp ginsenoside ở mẫu lá sâm Ngọc Linh 1 và 4 năm tuổi;

e) Phân tích biểu hiện các gen liên quan đến con đường sinh tổng hợp ginsenoside ở mẫu thân rễ sâm Ngọc Linh 1 và 4 năm tuổi;

f) So sánh mức biểu hiện của các gen liên quan đến tổng hợp ginsenoside giữa mô lá và thân rễ sâm Ngọc Linh ở hai độ tuổi.

CHƯƠNG 1. TỔNG QUAN TÀI LIỆU

1.1 CHI NHÂN SÂM VÀ SÂM NGỌC LINH

1.1.1 Chi Nhân sâm

Chi Nhân sâm (*Panax* L.) là loài cây lâu năm có củ, phát triển rất chậm thuộc họ Ngũ gia bì (Araliaceae), bộ Apiales (Hoa tán), chủ yếu mọc ở Bắc bán cầu trong khu vực Đông Á (phần lớn ở Triều Tiên - Cao Ly), Trung Quốc, Đông Siberi và Bắc Mỹ, đặc trưng là xứ lạnh. Trong đó ở Bắc Mỹ chỉ có 1 loài, còn lại tập trung ở Đông Á (kể cả Himalaya). Tất cả các loài đều có giá trị làm thuốc, một số có giá trị kinh tế rất cao, mang lại nguồn thu hàng trăm triệu USD như sâm Triều Tiên (*Panax ginseng*), sâm Mỹ (*Panax quinquefolius*), sâm Nhật (*Panax japonicus*) và Tam thất (*Panax notoginseng*) [9]. Gần đây, 12 loài với 2 thứ thuộc chi này đã được mô tả [10-11]. Các loài chính bao gồm: *P. quinquefolius* L., *P. trifolium* L., *P. ginseng* C.A.Mey., *P. pseudoginseng* Wall., *P. japonicus* (T.Nees) C.A.Mey., *P. notoginseng* (Burkill) F.H.Chen, *P. stipuleanatus* H.T.Tsai et K.M.Feng, *P. bipinnatifidus* Seem., *P. vietnamensis* Ha et Grushv. [4, 11]. Một số loài được gọi tên theo tên quốc gia, chỉ nguồn gốc xuất xứ của sâm như *P. japonicus*, *P. notoginseng*, *P. quinquefolius* và *P. vietnamensis* lần lượt được gọi là sâm Nhật, sâm Trung Quốc, sâm Mỹ và sâm Việt Nam [11].

Nhiều loài *Panax* chứa các hợp chất tự nhiên triterpene saponin hay còn gọi là các ginsenoside có cấu tạo phân tử khá phức tạp, độc đáo, có hoạt tính tốt và có tác dụng tăng cường thể lực, chống oxy hóa, ung thư, đái tháo đường, bảo vệ hệ thần kinh [1-2]. Nhiều công trình trên thế giới cũng như ở Việt Nam đã lựa chọn các loài thuộc chi *Panax* làm đối tượng nghiên cứu về thành phần hóa học và dược tính, đánh giá đa dạng di truyền, xây dựng mối quan hệ phát sinh chủng loại, nhân giống, gieo trồng. Từ năm 1959, khi Petkov lần đầu công bố về dược tính của chiết xuất từ *P. ginseng*, đến nay đã có trên 6.000 công trình về các phương thức sử dụng, thành phần hóa học, các đặc tính sinh học và dược học của sâm [11].

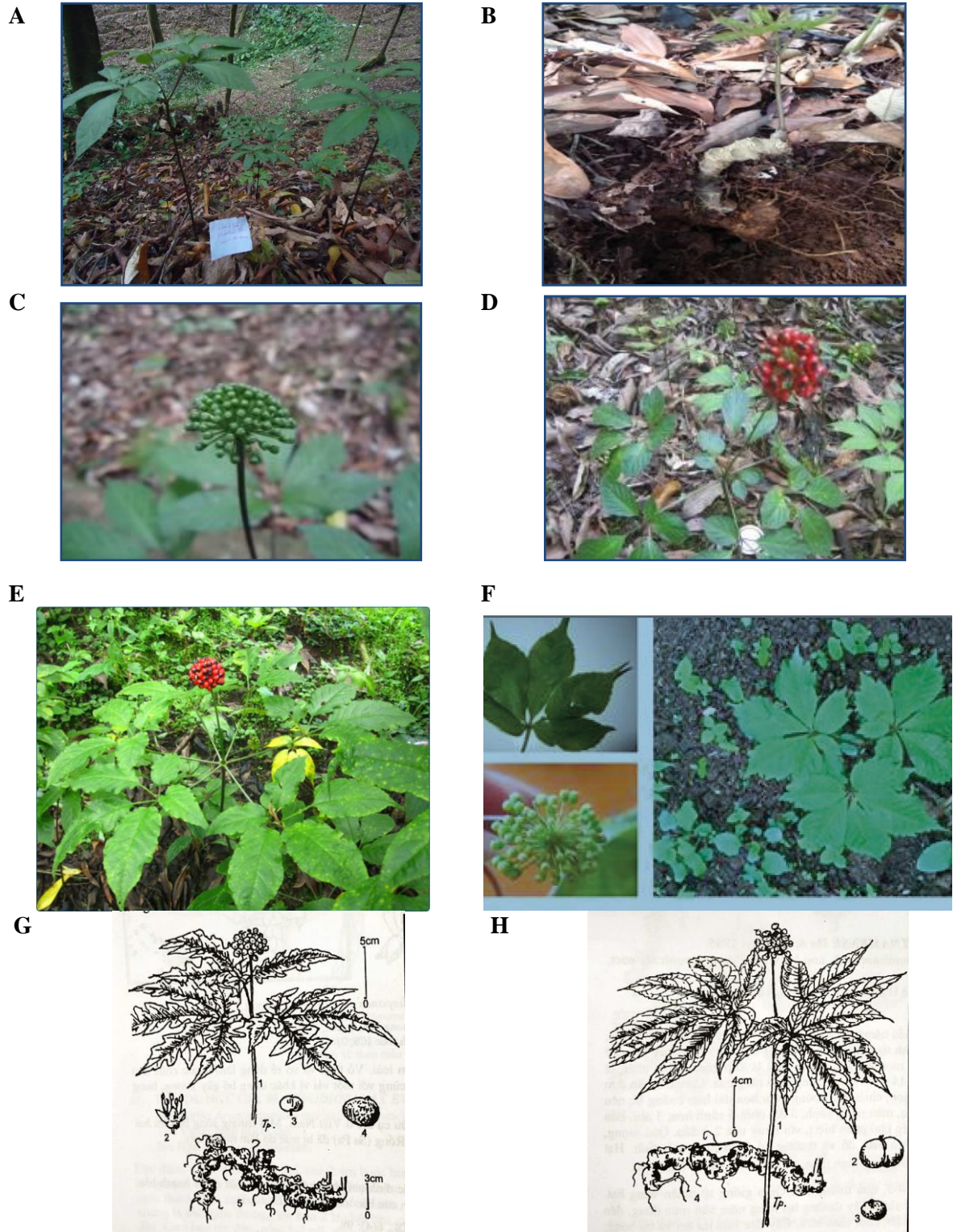
Do vùng phân bố hạn chế và giá trị kinh tế cùng nhu cầu tiêu dùng rất cao nên nhiều loài thuộc chi *Panax* đã bị khai thác quá mức và trở nên rất

hiếm trong tự nhiên [12,13]. Nhiều quốc gia đã triển khai các chương trình bảo tồn, khai thác và sử dụng bền vững các nguồn gen đặc biệt quý hiếm này, như sâm Triều Tiên/ Hàn Quốc, sâm Trung Quốc và sâm Mỹ được trồng thương mại và phổ biến nhất [11]. Gần đây, sâm Việt Nam cũng được ưu tiên phát triển thành sản phẩm quốc gia.

1.1.2 Sâm Ngọc Linh

Sâm Ngọc Linh (*Panax vietnamensis* Ha et Grushv.) thuộc chi Nhân sâm (*Panax* L.) là một trong những loài cây dược liệu quý hiếm và đặc hữu của Việt Nam, được sử dụng như thảo dược giúp phục hồi và tăng cường sức khỏe. Ở Việt Nam, các nghiên cứu về phân loại và xác định thành phần loài của chi *Panax* cho thấy chi này gồm 3 loài mọc tự nhiên có giá trị làm thuốc bao gồm *P. bipinnatifidus* Seem. (sâm Vũ diệp), *P. stipuleanatus* Tsai et Feng (Tam thất hoang), *P. vietnamensis* Ha et Grushv. (sâm Ngọc Linh) [3-7]. Sâm Ngọc Linh còn có các tên gọi khác là sâm Việt Nam, sâm khu Năm (K5), sâm trúc (sâm đốt trúc, trúc tiết sâm), củ Ngải rơm con hay cây Thuốc giầu. Sâm Ngọc Linh được phát hiện vào năm 1973 và đến năm 1985 mới được công bố là hoàn toàn mới đối với khoa học [3]. Sâm Ngọc Linh mọc dày thành đám dưới tán rừng dọc theo các suối ẩm trên đất nhiều mùn, mọc tập trung chủ yếu ở chân núi Ngọc Linh, thuộc các huyện Trà My, Trà Lĩnh, Trà Giang, tỉnh Quảng Nam. Hai thứ khác của *P. vietnamensis* được phát hiện gần đây là *P. vietnamensis* var. *fuscidiscus* K.Komatsu, S.Zhu et S.Q.Cai (sâm Lai Châu) phân bố ở Lai Châu [7], *P. vietnamensis* Ha et Grushvitzky var. *langbianensis* N.V.Duy, V.T.Tran & L.N.Trieu (sâm Lang Bian) được phát hiện ở vùng núi Langbian, Lâm Đồng. Ngoài ra, *Panax* sp. (sau đây gọi là *Panax* sp. Puxailaileng) phân bố ở vùng núi Puxailaileng thuộc tỉnh Nghệ An được xác định là một thứ của sâm Ngọc Linh và cần các nghiên cứu sâu hơn [14]. Gần đây, phân tích cây phát sinh chủng loại dựa trên vùng ITS-rDNA và *matK*, Pham *et al.* (2020) đã xác định sâm Puxailaileng là *P. vietnamensis* var. *fuscidiscus* [15]. Tương tự như nhiều loài thuộc chi *Panax*, sâm Ngọc Linh trở nên cực kỳ quý hiếm trong tự nhiên được đưa vào Danh lục đỏ của IUCN (2003) và Sách đỏ Việt Nam (2007) do phạm vi phân bố hạn chế và việc khai thác quá mức. Hình 1.1 Các loài tự nhiên thuộc chi *Panax* ở Việt Nam mô tả một số đặc điểm hình thái của các loài/ thứ thuộc chi *Panax* ở Việt Nam (bao

gồm sâm Ngọc Linh, sâm Lai Châu, sâm Nghệ An, sâm Vũ diệp, Tam thất hoang) sử dụng trong định loại hình thái [4-5, 7, 14].



Hình 1. 1 Các loài tự nhiên thuộc chi *Panax* ở Việt Nam

A: Hình dạng cây *Panax vietnamensis* Ha et Grushv. 5 năm tuổi; B: Thân rễ *P.*

vietnamensis 7 năm tuổi; C: Đặc thù hoa *P. vietnamensis*; D: Đặc thù quả *P. vietnamensis* [16]; E: Hình ảnh cây mọc tự nhiên của *Panax vietnamensis* var. *fuscidicus* [17]; F: Hình ảnh lá, hoa và cây mọc tự nhiên của *Panax* sp. Puxailaileng [14]; G: Hình vẽ *Panax bipinnatifidus* Seem. theo mẫu VDL 3265: 1. Cành mang hoa; 2. Hoa; 3. Quả; 4. Hạt; 5. Thân rễ (củ) [5]; H: Hình vẽ *Panax stipuleanatus* Tsai et Feng theo mẫu VDL 3270: 1. Cành mang hoa; 2. Quả; 3. Hạt; 4. Thân rễ (củ) [5].

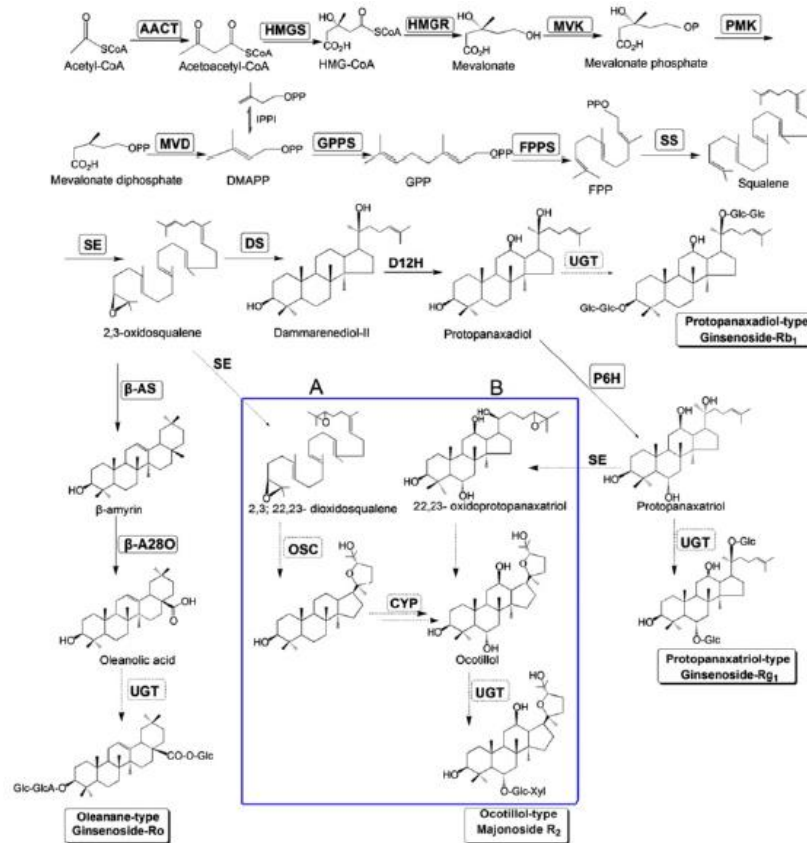
Sâm Ngọc Linh chứa các hợp chất ginsenoside với nhiều giá trị dược liệu quan trọng. Các ginsenoside có bản chất là triterpene saponin với cấu trúc bao gồm 1 khung squalene 30 carbon được gắn với các gốc đường ở nhiều vị trí khác nhau trên khung. Tên của các triterpene saponin được đặt theo chi thực vật mà chúng được phát hiện lần đầu tiên (astragaloside, cucurbitacin, ginsenoside...) hoặc Rx với x là vị trí tương đối hợp chất trong thí nghiệm sắc ký lớp mỏng (Ro, Ra1, Rb1, Rg1...). Hiện nay, hơn 30 loại ginsenoside đã được tách chiết từ chi *Panax* và đã được phân làm 2 nhóm chính là dammarane ginsenoside và oleanane ginsenoside. Phần lớn các ginsenoside thuộc nhóm dammarane ginsenoside, bao gồm Rb₁, Rc, Rd, Re, và Rg₁... trong khi nhóm oleanane ginsenoside chỉ có 1 đại diện duy nhất là Ro.

Các con đường sinh tổng hợp ginsenoside được chia thành hai phần: (1) acetyl-CoA trải qua tác dụng của mevalonate (MVA) để tạo thành 3-isoprene pyrophosphate (IPP) và dimethylallyl pyrophosphate (DMAPP), sau đó IPP và DMAPP được xúc tác bởi các enzym khác nhau để tạo thành 2,3-oxidosqualene; và (2) các loại ginsenoside được hình thành nhờ các phản ứng vòng hóa, hydroxyl hóa và glycosyl hóa hợp chất 2,3-oxidosqualene. Con đường sinh tổng hợp ginsenoside là một chuỗi hơn 20 phản ứng xúc tác liên tục, với sự tham gia của hàng loạt enzyme như 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR), farnesyl pyrophosphate synthase (FPS), squalene synthase (SS), squalene epoxidase (SQE), dammarenediol-II synthase (DS), β -amyrin synthase (AS), cytochrome P450 (CYP450) và UDP-glycosyltransferase (UGT)...[18].

Các ginsenoside được sinh tổng hợp trong tế bào thông qua con đường isoprenoid với tiền chất ban đầu là 2,3-oxidosqualene có nhóm chức sterol

(Hình 1.2) [17]. Trong đó, 2,3-oxidosqualene được tạo ra từ acetyl-CoA thông qua các biến đổi phổ biến ở nhiều loài thực vật, sau đó 2,3-oxidosqualene được đóng vòng bởi các 2,3-oxidosqualene cyclase (OSCs) như β -amyrin synthase (AS) hoặc dammarenediol-II synthase (DS) nhằm tạo ra các khung tương ứng với oleanane ginsenoside hoặc dammarane ginsenoside. Để tạo hoạt tính cho hợp chất, các khung này sau đó sẽ được tương tác với các cytochrome P450 (CYP450) và các glycotranferase (UGT) đặc hiệu để lắp ghép các nhóm đường, từ đó tạo các ginsenoside hoàn chỉnh [19-21]. Sâm Ngọc Linh lần đầu tiên được phát hiện ở vùng Tây Nguyên, miền Trung Việt Nam vào năm 1973. Phát hiện này đã đóng góp cho chi *Panax* một loài mới, đồng thời khởi đầu cho các nghiên cứu về hóa học, hoạt tính sinh học, nâng cao giá trị của loài *Panax* đặc hữu ở Việt Nam. Sâm Ngọc Linh được xác định là một cây thuốc quý của Việt Nam, tuy nhiên, cho đến nay số nhóm nghiên cứu về thành phần hóa học của loài này còn khá ít (<https://www.cas.org/products/scifinder>). Các nghiên cứu đã khái quát cho thấy sâm Ngọc Linh khá phong phú về số lượng các hợp chất dammaran saponin (với 2 loại khung aglycone chính protopanaxadiol và protopanaxatriol) và ocotillol saponin. Trong đó, các ocotillol saponin được ước tính chiếm 36-75% về khối lượng saponin tổng số. Đặc biệt, ocotillol saponin majonoside-R₂ (M-R₂) được đánh giá là hợp chất chính chiếm hàm lượng khoảng 5,3% về khối lượng mẫu dược liệu khô [22, 23].

Để khẳng định giá trị đặc biệt của sâm Ngọc Linh, nhiều nhà nghiên cứu đã tiến hành đánh giá, so sánh các hoạt chất của sâm Ngọc Linh và các loài sâm khác như *P. ginseng*, *P. quinquefolium* và *P. notoginseng*. Kết quả so sánh thành phần các saponin chính cũng khẳng định giá trị sử dụng của sâm Ngọc Linh là rất quý [24-25]. Ngoài ra, sâm Việt Nam còn chứa 18 loại amino acid và 20 nguyên tố vi lượng. Trong đó, đủ 8 loại amino acid cần thiết cho cơ thể, đặc biệt Arginin được tìm thấy với hàm lượng khá lớn (chiếm 46,66% tổng lượng amino acid) và một số nguyên tố vi lượng quan trọng như K, Na, Mg, Fe...[26].



Hình 1. 2 Sơ đồ giản lược quá trình sinh tổng hợp các ginsenoside ở thực vật

AACT: acetyl-CoA acetyltransferase; AS: β -amyrin synthase; CAS: cycloartenol synthase; DS: dammarenediol-II synthase; FPS: farnesyl diphosphate synthase; HMGR: HMG-CoA reductase; HMGS: HMG-CoA synthase; IDI: isopentenyl diphosphate isomerase; MVD: mevalonate diphosphate decarboxylase; MVK: mevalonate kinase; P450: cytochrome P450; PMK: phosphomevalonate kinase; Rb1: ginsenoside Rb1; Rg1: ginsenoside Rg1; Ro: ginsenoside Ro; SE: squalene epoxidase; SS: squalene synthase và UGT: glycosyltransferase.

Từ năm 1985 đến nay, đã có nhiều công trình nghiên cứu công bố về hóa học, hoạt tính sinh học và cấu trúc hóa học của các hợp chất tiêu biểu phân lập từ sâm Ngọc Linh [3, 6, 28]. Những kết quả nghiên cứu này khẳng định giá trị rất quý của sâm Ngọc Linh so với các loại nhân sâm khác trên thị trường. Từ những năm 1990, sâm Ngọc Linh trong tự nhiên hầu như đã bị khai thác cạn kiệt và có nguy cơ tuyệt chủng, trở thành một loài thực vật có tên trong sách đỏ Việt Nam [8]. Theo tiêu chuẩn phân hạng trong Danh lục

Đỏ của IUCN (1994 và 2001), sâm Ngọc Linh ở trong tình trạng “Cực kỳ nguy cấp”. Trước tình hình đó, Việt Nam đã có những chủ trương, chính sách đúng đắn nhằm khôi phục, bảo tồn và phát triển cây sâm Ngọc Linh. Nhiều công trình nghiên cứu nhằm điều tra, phân loại, xác định vùng phân bố, nghiên cứu thành phần hóa học và dược tính, đánh giá đa dạng di truyền, tìm kiếm mã vạch DNA, xây dựng mối quan hệ phát sinh chủng loại, bảo tồn, nhân giống, sản xuất giá thể, kỹ thuật canh tác, gieo trồng, bảo vệ thực vật, chế biến, bảo quản sâm Ngọc Linh đã và đang được thực hiện.

1.2 NGHIÊN CỨU GIẢI TRÌNH TỰ HỆ GEN Ở CÁC LOÀI THUỘC CHI *PANAX*

Các nghiên cứu giải trình tự gen ở các loài thuộc chi *Panax* thường tập trung vào hệ gen phiên mã, hệ gen lục lạp, xây dựng thư viện của các trình tự biểu hiện (Expressed sequence tags - ESTs) phục vụ các nghiên cứu đánh giá đa dạng di truyền nguồn gen, tìm hiểu vai trò của các gen chức năng, đặc biệt là các gen liên quan đến con đường sinh tổng hợp các hoạt chất có giá trị ở sâm, tìm kiếm các chỉ thị phân tử, xác định mối quan hệ tiến hóa...

Năm 2010, Sun và đtg đã công bố nghiên cứu đầu tiên về giải trình tự hệ gen phiên mã của *P. quinquefolius*, xác định được hầu hết các gen mã hóa enzyme đặc hiệu tham gia trực tiếp vào quá trình sinh tổng hợp khung của các ginsenoside [29]. Cũng trên đối tượng này, Wu và đtg (2013) đã tiến hành giải trình tự các ESTs ở rễ sâm *P. quinquefolius* các độ tuổi khác nhau, trong đó có 3.955 trình tự mã hóa cho các gen tham gia vào tổng hợp các saponin [30].

Tương tự nghiên cứu của Sun và đtg (2010), nhóm đã tìm ra được 11 gen mã hóa cho enzyme tham gia sinh tổng hợp khung của triterpene saponin, 174 trình tự mã hóa các cytochrome P450 và 242 UGT từ mô rễ *P. notoginseng*. Cũng trên đối tượng này, Liu và đtg (2015) đã giải trình tự hệ gen phiên mã sử dụng hệ thống giải trình tự thế hệ mới. Sau khi lắp ráp các trình tự đọc, nhóm tác giả thu được 107.340 unigene, trong đó 9.908 unigene liên quan đến con đường sinh tổng hợp các hợp chất hóa học, với 270 unigene tham gia sinh tổng hợp triterpene saponin [31].

Các nghiên cứu giải trình tự hệ gen phiên mã của *P. ginseng* cũng được tiến hành. Năm 2011, Chen và đtg công bố thu được gần 32 nghìn ESTs dựa trên việc phân tích trình tự của thư viện cDNA ở rễ sâm 11 năm tuổi bằng hệ thống giải trình tự GS FLX Titanium. Nhóm còn xác định được 9 gen đặc hiệu tham gia vào quá trình sinh tổng hợp ginsenoside, 133 gen mã hóa các cytochrome P450 và 235 gen mã hóa các glycosyltransferase [32]. Năm 2013, nghiên cứu của Li và đtg đã phát hiện tổng cộng 233 gen mã hóa enzyme tham gia vào con đường sinh tổng hợp ginsenoside, 326 gen mã hóa cytochrome P450 và 129 gen mã hóa các glycosyltransferase. Ngoài ra, nhóm còn xác định được 14 trình tự mã hóa các microRNA có thể điều hòa việc tổng hợp protein của khoảng 100 gen đích và hơn 13.000 trình tự SSR [33]. Năm 2013, Mathiyalagan và đtg tại Trung tâm Nghiên cứu sâm Hàn Quốc đã công bố nghiên cứu về 69 miRNAs bảo thủ dựa trên ngân hàng trình tự ESTs xây dựng từ thư viện cDNA của hoa, lá và rễ ở *P. ginseng* [34].

Năm 2016, Rai và đtg đã tiến hành nghiên cứu giải trình tự RNA bằng hệ thống Illumina trên đối tượng sâm Nhật Bản *P. japonicus* thu được 135.235 unigene, trong đó xác định được 24 unigene mã hóa cho cytochrome P450 và 48 unigene mã hóa cho các enzyme chuyển hóa glycosyltransferase. Nghiên cứu này cũng cho thấy các unigene trong *P. japonicus* được bảo tồn trình tự như tất cả các loài sâm khác và đồng biểu hiện với các gen khác có liên quan đến con đường sinh tổng hợp khung triterpenoid [35].

Trên đối tượng *P. vietnamensis*, nhóm nghiên cứu tại Hàn Quốc đã sử dụng hệ thống Illumina MiSeq để giải trình tự *de novo* hệ gen lục lạp từ DNA tổng số mẫu sâm Ngọc Linh thu từ vườn nhà kính tại thành phố Đà Lạt (Lâm Đồng). Kết quả của nghiên cứu đưa ra trình tự hệ gen lục lạp với kích thước 155.992 bp với 86.177 bp vùng LSC, 17.935 bp vùng SSC, 25.940 bp thuộc hai vùng IRa và IRb [36]. Manzanilla và đtg (2018) đã sử dụng phương pháp phân tách các phân đoạn DNA methyl hóa và không methyl hóa ở CpG kết hợp với giải trình tự gen thế hệ mới để giải trình tự thành công toàn bộ hệ gen lục lạp của các loài thuộc chi *Panax* phân bố ở Việt Nam. Zhang và đtg (2015) đã giải trình tự hệ gen phiên mã của *P. vietnamensis* var. *fuscidicus* sử dụng hệ thống Illumina HiSeq2000 và thu được 114.703.210 đoạn đọc với 126.758 unigene. Trình tự cDNA hoàn chỉnh của 3 gen mã hóa SE đã được

phân lập sử dụng RT-PCR và 10 gen đã được nghiên cứu biểu hiện thông qua real-time PCR (qPCR) [17]. Dữ liệu hệ gen của loài là cơ sở để nghiên cứu các gen liên quan đến con đường sinh tổng hợp saponin. Các chỉ thị SSR nhận được trong nghiên cứu cho thấy sự đa dạng di truyền của nguồn gen và là chỉ thị có giá trị trong chọn tạo giống.

1.3 NGHIÊN CỨU BIỂU HIỆN CỦA CÁC GEN CHỨC NĂNG Ở CÁC LOÀI THUỘC CHI *PANAX*

Phương pháp qPCR được sử dụng rộng rãi như là một công cụ hữu hiệu trong việc định lượng mức độ biểu hiện của gen nhờ độ nhạy cao, tính đặc hiệu và khả năng lặp lại [37-38]. Đặc trưng của qPCR là phát hiện sản phẩm khuếch đại trong quá trình chạy PCR khi sản phẩm khuếch đại từ DNA đích được nhân bản đủ số lượng để làm ống phản ứng phát được huỳnh quang khi nhận được nguồn sáng kích thích. Nhờ những tính chất trên, qPCR được sử dụng để đánh giá mức độ biểu hiện các gen liên quan đến con đường sinh tổng hợp ginsenoside ở các mô khác nhau của nhiều loài sâm. Năm 2008, He và đtg đã nghiên cứu mức độ biểu hiện của enzyme SE, một trong những enzyme hiếm trong con đường sinh tổng hợp các triterpene saponin ở *P. notoginseng*. Về trình tự, gen này giống 98% với *P. ginseng* và biểu hiện cao trong mô rễ, đặc biệt là rễ 3 năm tuổi [39]. Năm 2010, dựa trên các trình tự hệ gen phiên mã thu được, Sun và đtg đã nghiên cứu mức độ biểu hiện của 1 CYP450 và 4 UGT trên các mô của *P. quinquefolius* sử dụng qPCR [29]. Luo & đtg (2011) đã đánh giá mức độ biểu hiện của DS, enzyme quan trọng của quá trình sinh tổng hợp các triterpenes saponin và kết quả cho thấy hoạt chất này biểu hiện mạnh nhất ở rễ *P. notoginseng* 4 năm tuổi [40]. Li và đtg (2013) đã xác định mức độ biểu hiện khác biệt của các gen ở các mô và thời điểm sinh trưởng khác nhau của sâm *P. Ginseng* [33]. Liu *et al.* (2015) đã tiến hành phân tích hệ gen phiên mã của *P. notoginseng* để so sánh sự biểu hiện các gen liên quan đến quá trình sinh tổng hợp ginsenoside và alkaloid ở mô lá, rễ và hoa. Kết quả cho thấy cytochrome P450 CYP716A53v2, enzyme tham gia vào quá trình chuyển hóa protopanaxatriol từ protopanaxadiol trong *P. ginseng*, có mức độ biểu hiện ở mô rễ cao hơn rất nhiều so với mô lá và hoa. Park và đtg (2016) đã nghiên cứu ảnh hưởng của sự biểu hiện gen

CYP716A53v2 lên mức độ tích lũy protopanaxatriol trong rễ *P. ginseng*. Sau khi tiến hành tăng biểu hiện gen, mức độ phiên mã của gen này trong rễ chứa nồng độ các ginsenoside nhóm protopanaxatriol (gồm Rg1, Re và Rf) cao và nồng độ các ginsenoside nhóm protopanaxadiol (Rb1, Rc, Rb2, và Rd) thấp. Ngược lại, ức chế biểu hiện gen *CYP716A53v2* bằng RNAi và kiểm tra bằng qPCR cho kết quả nồng độ các hợp chất protopanaxatriol thấp và nồng độ các hợp chất nhóm protopanaxadiol cao. Điều này cho thấy hàm lượng các hợp chất ginsenoside nhóm protopanaxadiol và protopanaxatriol có thể được thay đổi thông qua điều chỉnh mức độ biểu hiện gen *CYP716A53v2* trong cây chuyển gen *P. Ginseng* [41].

CHƯƠNG 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. VẬT LIỆU NGHIÊN CỨU

2.1.1. Các mẫu sâm nghiên cứu

Tổng 12 mẫu sâm Ngọc Linh (bao gồm 3 mẫu thân rễ và 3 mẫu lá tương ứng với mỗi độ tuổi 1 năm tuổi và 4 năm tuổi) được thu thập phục vụ xác định và đánh giá sự biểu hiện của các gen chức năng liên quan đến con đường sinh tổng hợp ginsenoside. Các mẫu được định loại hình thái và xác định độ tuổi bởi PGS. TS. Nguyễn Tập (nguyên cán bộ Viện Dược liệu, Bộ Y tế), được định loại phân tử tại Phòng Đa dạng sinh học hệ gen, Viện Nghiên cứu hệ gen. Các mẫu được bảo quản tươi cho đến khi về phòng thí nghiệm, sau đó được bảo quản lạnh trong -80°C và sử dụng cho các thí nghiệm tiếp theo.

Địa điểm thu mẫu: Trại sâm gốc Tắc Ngo (Trung tâm Sâm Ngọc Linh, huyện Nam Trà My) và Trạm Dược liệu Trà Linh (Trung tâm Phát triển sâm Ngọc Linh và dược liệu Quảng Nam).

Dữ liệu hệ gen phiên mã của mô lá và mô thân rễ sâm Ngọc Linh 1 năm tuổi và 4 năm tuổi thu được sử dụng phương pháp giải trình tự gen thế hệ mới do Phòng Đa dạng sinh học hệ gen, Viện Nghiên cứu hệ gen cung cấp.

2.1.2. Các cặp mồi sử dụng

Các cặp mồi phục vụ nhân gen quan tâm sử dụng RT-PCR: Bốn cặp mồi được thiết kế để khuếch đại 4 gen liên quan đến con đường sinh tổng hợp ginsenoside của sâm Ngọc Linh) (Bảng 2.1).

Bảng 2. 1 Các cặp mồi phục vụ nhân gen quan tâm sử dụng RT-PCR

TT	Vùng gen	Tên mồi	Trình tự mồi (5'-3')	Kích thước sản phẩm (bp)
1	AACT	PvAACT_RT_F1	GATACACTTCTCCCAGCTTCCG	1364
		PvAACT_RT_R1	CTGGAGCAATGAAACTACCAAGT	
2	IDI	PvIDI_F	CCGTAGTATAATGTCAACCAC	962

TT	Vùng gen	Tên mồi	Trình tự mồi (5'-3')	Kích thước sản phẩm (bp)
		PvIDI_R	CGGTCTGAAGTACAAATCAAG	
3	SS	PvSS_RT_F1	GGGAAGTTTGGGGGCAATTCTG	1229
		PvSS_RT_R1	GTAGGTTTGAAGATAGATATGC	
4	SE	PvSE3_RT_F1	CAACCAACACCATGAATTC	1655
		PvSE3_RT_R1	GCTTGATAACAAAATGAAGGC	

Các cặp mồi phục vụ nghiên cứu biểu hiện của các gen sử dụng real-time PCR: Năm cặp mồi được thiết kế để đánh giá sự biểu hiện của 4 gen liên quan đến sinh tổng hợp ginsenoside) ở mô lá và thân rễ sâm Ngọc Linh ở các độ tuổi nghiên cứu (Bảng 2.2).

Bảng 2. 2. Các cặp mồi phục vụ nghiên cứu biểu hiện của các gen sử dụng realtime PCR

TT	Vùng gen	Tên mồi	Trình tự mồi (5'-3')	Kích thước sản phẩm (bp)
Gen sử dụng làm đối chứng nội				
1	EF	qPvEF-1g_F	ATCGCATTAAAGAGAGCACTAGG	186
		qPvEF-1g_R	CATGGTCCAAAAATATCTCTCTACG	
Các gen liên quan đến con đường sinh tổng hợp ginsenoside				
2	AACT	qPvAACT_F	CTATTCAGAGCTTTGAGCGTGG	145
		qPvAACT_R	GCAGCATCGAACTTTCCTAGACC	
3	IDI	qPvIDI_F	ACTTGTTGCAGCCATCCACT	110
		qPvIDI_R	GCAGGAATACCCAGTTCATCC	
4	SS	qPvSS_F	AGTGCAGTGCCTGAATGACA	122
		qPvSS_R	GCAGGATCTCGCAAATCAGA	
5	SE	qPvSE_F	CGTGGCTATCTATGGCGTTG	175
		qPvSE_R	GGGAGCTCTGTAGTAAGCAG	

2.1.3. Hóa chất, thiết bị

Hóa chất: Các hóa chất sử dụng đều đảm bảo độ tinh khiết cho nghiên cứu sinh học phân tử và được nhập từ các hãng Thermo Fisher Scientific (Hoa Kỳ), Applied Biosystems (Hoa Kỳ), Bioline (Hoa Kỳ), Eppendorf (Đức), Roche (Hoa Kỳ)... Các hóa chất chính bao gồm: *Taq* DNA polymerase, thang DNA chuẩn 1kb, dNTPs, SensiFAST cDNA Synthesis Kit, GeneJET™ Gel Extraction Kit, BigDye Terminator v3.1 Cycle Sequencing Kit, bộ kit SYBR Green qRT PCR master mix, SuperScript™ III One-Step RT-PCR System, các cặp môi, phenol, chloroform, EDTA, ethanol, agarose, 2-mercaptoethanol, sodium acetate, isopropanol, guanidinium thiocyanate, N-laurosylsarcosine, isoamyl alcohol, EtBr, nước khử ion vô trùng...

Thiết bị: Các máy móc, thiết bị chính được sử dụng thuộc Viện Hàn lâm Khoa học và Công nghệ Việt Nam bao gồm: Máy ly tâm Eppendorf 5415C (Eppendorf, Đức), máy quang phổ Biospectrometer (Eppendorf, Đức), máy PCR Mastercycler Pro (Eppendorf, Đức), hệ thống điện di (Bio-Rad, Hoa Kỳ), máy chụp ảnh GEL-DOC (Pharmacia Biotech, Thụy Điển), tủ lạnh -20°C, -80°C (Sanyo, Nhật Bản), cân phân tích 10⁻⁴g (Mettler Toledo, Thụy Sĩ), cân phân tích 5 số lẻ (Precisa, Thụy Sĩ), máy đo pH (Mettler, Thụy Sĩ), máy cô quay (Buchi, Đức), bể ổn nhiệt, nồi khử trùng, tủ sấy 60-200°C, máy khuấy từ, đèn soi UV-VIS, máy ảnh, hệ thống giải trình tự ABI 3500 Genetic Analyzer, hệ thống Cyclor 96. Ngoài ra, nghiên cứu sử dụng các chương trình phần mềm chuyên dụng miễn phí hoặc được cấp phép tại Viện Nghiên cứu hệ gen, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

2.2. PHƯƠNG PHÁP NGHIÊN CỨU

2.2.1. Thu mẫu

Mẫu lá và thân rễ của cây sâm Ngọc Linh 1 năm tuổi và 4 năm tuổi được thu và bảo quản riêng biệt. Các cây được rửa sạch bằng nước cất đã khử trùng, thấm khô, cắt nhỏ và bảo quản trong dung dịch RNAlater tự pha. Mẫu mô trong RNAlater được giữ ở 4°C trong 24 giờ trước khi bảo quản thời gian dài trong tủ lạnh sâu -80°C.

2.2.2. Tách chiết và kiểm tra chất lượng RNA tổng số

Tách chiết RNA tổng số của mẫu lá: RNA tổng số của các mẫu lá sâm Ngọc Linh 1 năm tuổi và 4 năm tuổi được tách chiết và tinh sạch theo quy trình của Chomczynski và Sacchi (2006) với các bước chính: (1) Cân 100 mg mô lá tươi được nghiền mịn trong nitơ lỏng, tránh để mô bị rã đông trong khi nghiền; Bổ sung 1 ml dung dịch biến tính có thành phần 4 M guanidinium thiocyanate, 25 mM sodium citrate, pH 7,0, 0,5% (wt/vol) N-lauroylsarcosine (sarkosyl) và 0,1 M 2-mercaptoethanol vào bột; (2) Chuyển phần dịch trên vào ống eppendorf; Bổ sung vào 1 ml dịch chiết 0,1 ml 2 M sodium acetate, pH 4,0, đảo ống để trộn kỹ; 1 ml phenol bão hòa, đảo ống để trộn kỹ; 0,2 ml chloroform : isoamyl alcohol (49 : 1), lắc mạnh bằng tay trong 10 giây; (3) Làm lạnh hỗn hợp trong đá trong 15 phút; (4) Ly tâm 20 phút ở 10.000 v/p, 4°C; (5) Nhẹ nhàng hút phần dịch trong (phần chủ yếu chứa RNA) vào eppendorf mới; (6) Bổ sung 1 ml isopropanol vào để tủa RNA; Ủ mẫu ít nhất 1 giờ ở -20°C; (7) Ly tâm 20 phút ở 10.000 v/p, 4°C và loại bỏ dịch nổi. Phần tủa RNA sẽ không nhìn thấy được trước khi ly tâm. Tuy nhiên, sau khi ly tâm sẽ hình thành tủa dạng gel; (8) Rửa tủa RNA bằng 0,5-1 ml 75% ethanol và vortex vài giây; Ủ mẫu ở nhiệt độ phòng trong 10-15 phút để loại bỏ phần vết của guanidium; (9) Ly tâm 5 phút ở 10.000 v/p, 4°C và loại bỏ dịch nổi; (10) Làm khô tủa RNA ở nhiệt độ phòng trong 5-10 phút; (11) Hòa tủa RNA bằng 50 µl nước đã xử lý bằng DEPC; Ủ trong 10-15 phút ở 60°C để hòa tan hoàn toàn tủa [42].

Tách chiết RNA tổng số của mẫu thân rễ: RNA tổng số của các mẫu thân rễ sâm Ngọc Linh 1 năm tuổi và 4 năm tuổi được tách chiết và tinh sạch theo quy trình có thay đổi so với tách chiết RNA tổng số của mẫu lá với các bước chính như sau: (1) Cân 200 mg mô thân rễ tươi được nghiền mịn thành bột trong nitơ lỏng; Chuyển ngay phần bột vào eppendorf có chứa sẵn 1 ml dung dịch biến tính, vortex kỹ để trộn đều; (2) Ủ mẫu ở nhiệt độ phòng trong 10 phút để ly giải hoàn toàn mô củ; (3) Ly tâm 10 phút ở 12.000 v/p, 4°C, loại bỏ phần tủa, thu phần dịch nổi vào 1 eppendorf mới; (4) Bổ sung 200 µl chloroform, 50 µl sodium acetate (2M, pH 4,2), 15 µl 2-mercaptoethanol, 2% PVP (20 µl), vortex kỹ, ủ trên đá 10 phút; (5) Ly tâm 15 phút với tốc độ 12.000 v/p ở 4°C, nhẹ nhàng hút phần dịch trong - là phần chứa chủ yếu RNA

vào eppendorf mới; (6) Bổ sung 1 ml isopropanol và 200 μ l sodium acetate (3M) vào để tủa RNA. Ủ mẫu ít nhất 1 giờ ở -20°C ; (7) Ly tâm 10 phút ở 12.000 v/p, 4°C và loại bỏ dịch nổi; (8) - (11): Tương tự như tách chiết RNA tổng số của mẫu lá.

Kiểm tra chất lượng của RNA tổng số: RNA tổng số của các mẫu lá và thân rễ sâm Ngọc Linh được kiểm tra, đánh giá chất lượng bằng điện di trên gel agarose 0,8%. Để xác định nồng độ và độ sạch, các mẫu RNA được phân tích bằng máy quang phổ Biospectrometer (Eppendorf, Đức). RNA sau tách chiết và tinh sạch được bảo quản ở -80°C cho đến khi được sử dụng cho các thí nghiệm tiếp theo.

2.2.3. Tổng hợp cDNA

Các đoạn mRNA thu được từ mô lá và mô thân rễ sâm Ngọc Linh ở các độ tuổi sẽ được phiên mã ngược tạo thành các đoạn cDNA sử dụng SensiFAST cDNA Synthesis Kit (Bioline, Hoa Kỳ) theo hướng dẫn của nhà sản xuất. Tổng hợp cDNA được thực hiện trong tổng thể tích 20 μ l, 14 μ l nước, 4 μ l 5x TransAmp Buffer, 1 μ l Reverse Transcriptase, 1 μ l RNA với chương trình sau: 25°C -10 phút; 42°C -15 phút; 85°C -5 phút; sau đó giữ ở 4°C . Sau tổng hợp, cDNA được sử dụng để nhân bản, xác định trình tự và phân tích biểu hiện một số gen liên quan đến sinh tổng hợp ginsenoside bằng phương pháp realtime PCR.

2.2.4. Xây dựng dữ liệu các gen liên quan đến con đường sinh tổng hợp ginsenoside

Trên cơ sở dữ liệu hệ gen phiên mã của mô lá và thân rễ sâm Ngọc Linh 1 năm tuổi và 4 năm tuổi đã được chú giải sử dụng các cơ sở dữ liệu GO, EggNOG, NT và NR của NCBI, KEGG, UniProt và Pfam, các gen chính liên quan đến quá trình sinh tổng hợp ginsenoside đã được xác định. Để tìm kiếm được các gen này trong dữ liệu hệ gen phiên mã sâm Ngọc Linh đã chú giải chức năng, các công cụ tìm kiếm tin sinh cơ bản với các từ khóa là các enzyme tham gia vào con đường sinh tổng hợp ginsenoside được sử dụng. Các unigene xuất hiện từ khóa ít nhất ở một trong các mục chú giải tương ứng với mỗi cơ sở dữ liệu được chọn lọc và sử dụng cho các nghiên cứu tiếp theo.

2.2.5. Thiết kế các cặp mồi

Thiết kế các cặp mồi nhân gen bằng RT-PCR: Sử dụng phần mềm Primer 3 kết hợp với công cụ thiết kế mồi OligoAnalyzer™ Tool (<https://www.idtdna.com/pages/tools/oligoanalyzer>), các cặp mồi được thiết kế trên cơ sở khai thác dữ liệu trình tự tham khảo của gen quan tâm và các trình tự gen của chi *Panax* đã công bố trên Ngân hàng gen quốc tế - GenBank (<http://www.ncbi.nlm.nih.gov/genbank>), cũng như trích xuất dữ liệu hệ gen phiên mã của *P. vietnamensis* thu được trong nghiên cứu. Trình tự các cặp mồi được trình bày trên Bảng 2.1 và kích thước dự kiến của các sản phẩm PCR là 1364, 962, 1229 và 1655 bp, bao gồm vùng mang mã và một phần vùng 5', 3' của các gen *AACT*, *IDI*, *SS*, *SE*.

Thiết kế các cặp mồi đánh giá sự biểu hiện gen bằng realtime PCR: Sử dụng các dữ liệu thu được từ GenBank, các cặp mồi đánh giá sự biểu hiện của các gen liên quan đến sinh tổng hợp ginsenoside được thiết kế sử dụng chương trình <http://frodo.wi.mit.edu/primer3/>. Vùng gen được lựa chọn là trình tự mang tính duy nhất đại diện cho gen đó để có thể đáp ứng yêu cầu đặc hiệu cao của mồi được thiết kế. Ngoài ra, các tiêu chuẩn khác cần đáp ứng tương tự như thiết kế mồi cho RT-PCR. Các trình tự mồi sau khi thiết kế được kiểm tra lại bằng công cụ OligoAnalyzer <https://www.idtdna.com/calc/analyzer/> (Bảng 2.2).

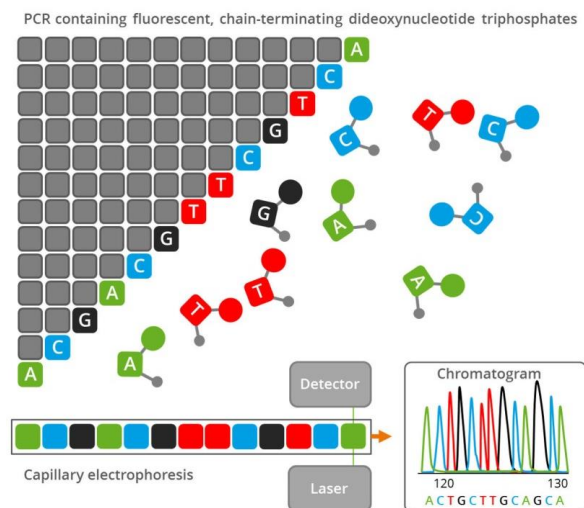
2.2.6. Nhân gen quan tâm sử dụng phương pháp reverse transcription PCR

RT-PCR được thực hiện sử dụng các cặp mồi đặc hiệu cho các gen cần xác định trình tự. PCR được thực hiện trong tổng thể tích 25 μ l, 0,5 μ M mồi mỗi loại, 200 μ M dNTPs, 5 μ l reaction buffer, 0,25 μ l DNA polymerase, 1 μ l cDNA khuôn, với chương trình sau: 94°C-3 phút; 35 chu kỳ (94°C-1 phút; 59°C*-1 phút; 72°C-5 phút; 72°C-10 phút; sau đó giữ ở 4°C. (*) Nhiệt độ gắn mồi tùy thuộc vào cặp mồi tương ứng. Sản phẩm PCR được điện di trên gel agarose và xác định trình tự [17].

2.2.7. Xác định trình tự gen quan tâm sử dụng phương pháp Sanger

Thôi gel và tinh sạch sản phẩm PCR: GeneJET™ Gel Extraction Kit (Thermo Fisher Scientific, Hoa Kỳ) và quy trình kèm theo được sử dụng để tinh sạch sản phẩm PCR với các bước chính: (1) Sản phẩm PCR được kiểm

tra trên gel agarose 1%; (2) Sau thời gian khoảng 1 giờ 30 phút ($U=110\text{ V}$, $I=100\text{ mA}$), bản gel được nhuộm EtBr $0,5\ \mu\text{g/ml}$ trong 15 phút, rửa sạch bằng nước, soi UV để cắt vạch gel (thực hiện bước thôi gel); (3) Băng DNA thu từ bản gel agarose với dao nhỏ sắc nhọn đã vô trùng, được chuyển vào ống eppendorf 1,5 ml, cân kiểm tra trọng lượng và thêm 1 thể tích dung dịch gắn (binding buffer), ủ ở $50\text{-}60^\circ\text{C}$ trong 10 phút, mỗi 2 phút đảo ống một lần cho đến khi gel tan hoàn toàn; (4) Mẫu được chuyển sang cột GeneJET™ Gel Extraction, ly tâm ở 13.000 v/p trong 1 phút, loại bỏ dịch; (5) Mẫu được bổ sung $700\ \mu\text{l}$ dung dịch rửa, ly tâm ở 13.000 v/p trong 1 phút, loại bỏ dịch; (6) Mẫu được ly tâm lại cột ở 13.000 v/p trong 1 phút, chuyển cột sang ống eppendorf mới, vô trùng; (7) Mẫu được làm khô và bổ sung $20\ \mu\text{l H}_2\text{O}$ khử ion vô trùng, để trong khoảng 2 phút, ly tâm ở 13.000 v/p , trong 3 phút. Mẫu được lưu trữ để thực hiện giải trình tự.



Hình 2. 1 Giải trình tự DNA bằng phương pháp Sanger

Giải trình tự các đoạn DNA bằng phương pháp Sanger: Trình tự của các gen quan tâm được xác định trên hệ thống giải trình tự ABI 3500 Genetic Analyzer sử dụng BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Hoa Kỳ). Thành phần của PCR đọc trình tự gồm: $3,2\ \mu\text{M}$ mỗi, $200\ \text{ng}$ sản phẩm PCR tinh sạch, BigDye, đệm tương ứng trong tổng thể tích $15\ \mu\text{l}$ với chu trình nhiệt trên máy Mastercycler Pro (Eppendorf, Đức) như sau: 96°C -1 phút; 25 chu kỳ (96°C -10 giây, 50°C -5 giây, 60°C -4 phút); giữ ở 4°C . Sản phẩm PCR được tinh sạch bằng phương pháp tủa ethanol/EDTA: (1) Bổ sung $5\ \mu\text{l}$ EDTA $125\ \text{mM}$, $60\ \mu\text{l}$ ethanol 100%, để ở nhiệt độ phòng trong 15 phút; (2) Ly tâm 12.000 v/p trong 15 phút để thu tủa

DNA và loại bỏ ethanol; (3) Bổ sung 60 μl ethanol 70% và ly tâm 10.000 v/p trong 10 phút, làm khô tua DNA; (4) Bổ sung 10 μl Hi-Di formamide và biến tính tại 95°C trong 5 phút; (5) Các mẫu được cho vào các giếng của khay đựng mẫu, sau đó điện di trong ống mao quản 80 cm x 50 μm với polymer POP-7 (Applied Biosystems, Hoa Kỳ). Trình tự nucleotide của mỗi mẫu được xác định từ 2 chiều xuôi và ngược.

Phân tích, so sánh, đánh giá các vùng trình tự nhận được với các trình tự đã công bố trên Ngân hàng gen quốc tế: Kết quả xác định trình tự của các đoạn gen được đối chiếu và so sánh với trình tự tương ứng của các loài có quan hệ gần gũi đã được công bố trên GenBank nhằm loại bỏ các vị trí lỗi do đọc trình tự và phát hiện các đa hình nucleotide đơn. Việc phân tích trình tự được hỗ trợ bằng các phần mềm chuyên dụng như Bioedit 7.0.9.0.

2.2.8. Nghiên cứu biểu hiện một số gen chức năng bằng realtime PCR

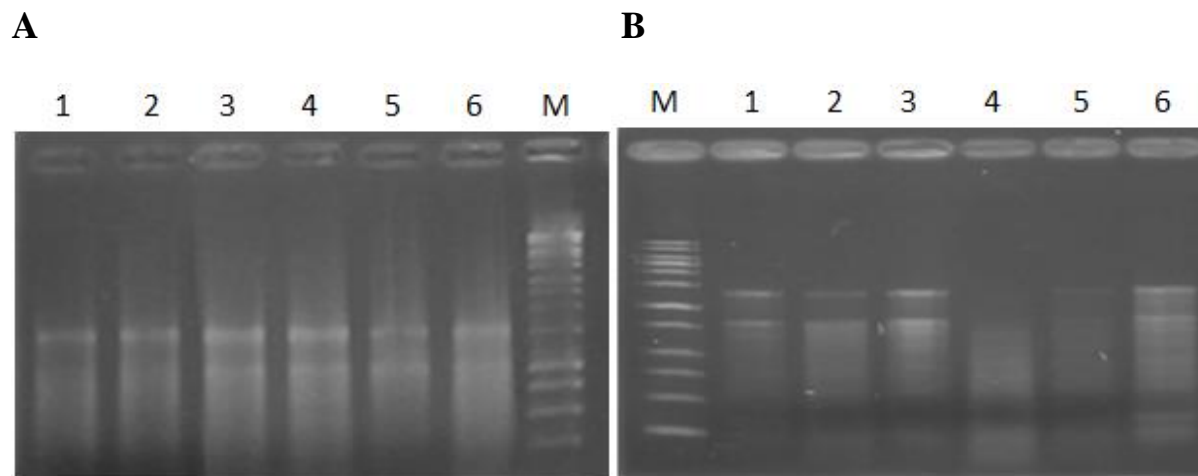
Các gen chức năng quan trọng tham gia vào quá trình tổng hợp và chuyển hóa ginsenoside được kiểm tra mức độ biểu hiện bằng phương pháp qPCR với các cặp mồi đặc hiệu cho từng gen (Bảng 2.2). Vùng gen được lựa chọn là trình tự mang tính duy nhất đại diện cho gen đó để có thể đáp ứng yêu cầu đặc hiệu cao của mồi được thiết kế.

Phản ứng realtime PCR được thực hiện với bộ kit SYBR Green qRT PCR master mix (Thermo Scientific, Hoa Kỳ) trong hệ thống Cycloer 96 (Roche, Hoa Kỳ). Các gen quan tâm được đánh giá mức độ biểu hiện bằng qPCR với các cặp mồi đặc hiệu cho từng gen với mỗi phản ứng chứa 20 μL chứa 2x SYBR™ Green PCR Master Mix (Thermo Fisher Scientific, Hoa Kỳ), 0,4 μM mồi xuôi, 0,4 μM mồi ngược và 2 μL cDNA khuôn. Điều kiện phản ứng PCR bao gồm giai đoạn biến tính ở 95°C-60 giây, sau đó là 45 chu kỳ nhiệt với 95°C-20 giây, 56°C-20 giây và 72°C-20 giây. Kết thúc phản ứng có thêm giai đoạn “high resolution melting stage” với chu trình 95°C-5 giây, 65°C-60 giây và 97°C-15 giây. Mỗi phản ứng được lặp lại 3 lần. Gen Elongation factor 1 (EF_1) được sử dụng làm đối chứng nội. Sản phẩm khuếch đại được phân tích nhiệt độ biến tính để kiểm tra độ đặc hiệu của phản ứng. Mức độ biểu hiện gen được định lượng thông qua giá trị $2^{-\Delta\Delta\text{Ct}}$ được công bố bởi Livak và Schmittgen (2001) [43]. Các số liệu nghiên cứu được biểu thị dưới dạng giá trị trung bình \pm độ lệch chuẩn và được xử lý theo các thuật toán thống kê Student *t*-test. Sự khác biệt có ý nghĩa khi giá trị $p < 0,05$.

CHƯƠNG 3. KẾT QUẢ VÀ THẢO LUẬN

3.1. KẾT QUẢ TÁCH CHIẾT, TINH SẠCH, KIỂM TRA CHẤT LƯỢNG, ĐỊNH LƯỢNG RNA TỔNG SỐ TỪ MÔ LÁ, THÂN RỄ SÂM NGỌC LINH

Để chuẩn bị nguyên liệu tổng hợp cDNA phục vụ xác định trình tự gen quan tâm bằng RT-PCR và nghiên cứu mức độ biểu hiện của các gen bằng realtime PCR, RNA tổng số ở các mẫu lá và thân rễ của sâm Ngọc Linh ở các độ tuổi được tách chiết, tinh sạch và kiểm tra, đánh giá chất lượng bằng điện di trên gel agarose 1,5% (Hình 3.1). Kết quả cho thấy, các mẫu RNA thu được có chất lượng tương đối tốt, có sự xuất hiện của các băng ribosome RNA với kích thước khoảng 1,5-2 kb. Đối với các mẫu lá, các mẫu RNA thu được khá sạch và không bị nhiễm DNA (Hình 3.1). Đối với mẫu thân rễ sâm, RNA tách chiết được có băng nhạt màu hơn so với các mẫu RNA tách chiết từ lá (Hình 3.1B). Vệt smear cho thấy một phần RNA bị đứt gãy và không toàn vẹn. Để tinh sạch một số mẫu RNA, DNase được sử dụng nhằm loại bỏ DNA tổng số.



Hình 3. 1 Kiểm tra kết quả tách chiết và tinh sạch RNA tổng số trên gel agarose từ các mẫu lá (A) và thân rễ sâm Ngọc Linh (B)

M: Thang chuẩn DNA 1 kb; A: 1-6: L1.1-L1.3, L4.1-L4.3: Tương ứng là 03 mẫu lá của sâm Ngọc Linh 1 hoặc 4 năm tuổi; B: 1-6: C1.1-C1.3, C4.1-C4.3: Tương ứng là 03 mẫu thân rễ của sâm Ngọc Linh 1 hoặc 4 năm tuổi.

Để đánh giá chất lượng RNA thu được, các mẫu này được phân tích bằng máy quang phổ Biospectrometer (Eppendorf, Đức). Kết quả kiểm tra nồng độ và độ sạch của RNA (Bảng 3.1) cho thấy, các sản phẩm có chỉ số

A260/A280 thể hiện sự tinh sạch của mẫu dao động từ 1,90-2,15 so với yêu cầu phải đạt từ 1,8~2,0 chứng tỏ các mẫu RNA tách chiết được có chất lượng cao, không bị tạp nhiễm protein, đủ điều kiện để tiến hành những thí nghiệm tiếp theo. Đặc biệt, nồng độ của RNA thu được ở các mẫu tương đối cao.

Bảng 3. 1 Nồng độ và độ sạch (A260/A280) của một số mẫu RNA sâm Ngọc Linh sau tách chiết và tinh sạch

STT	Tên mẫu	Nồng độ (ng/ μ l)	A260/A280
1	L1.1	1.979,6	2,12
2	L1.2	2.348,1	2,15
3	L1.3	2.734,6	2,01
4	L4.1	1.010,0	1,96
5	L4.2	792,20	2,05
6	L4.3	1.468,3	2,08
7	C1.1	854,00	2,07
8	C1.2	365,70	1,95
9	C1.3	105,20	2,03
10	C4.1	207,10	2,11
11	C4.2	482,40	2,11
12	C4.3	470,80	1,90

3.2. KẾT QUẢ XÂY DỰNG CON ĐƯỜNG SINH TỔNG HỢP GINSENOSE GIẢ ĐỊNH Ở SÂM NGỌC LINH

3.2.1. Thống kê sơ bộ số lượng unigene gắn với mỗi gen

Dựa vào kết quả chú giải các hệ gen phiên mã nhận được từ giải trình tự hệ gen phiên mã, các unigene của các mẫu mô lá và thân rễ sâm Ngọc Linh được rà soát theo các từ khóa liên quan tới các gen tham gia vào con đường sinh tổng hợp ginsenoside. Kết quả sàng lọc cho thấy tất cả các gen đều được chú giải cho ít nhất 1 unigene (Bảng 3.2). Nhiều unigene được chú giải thuộc 2 họ gen lớn *CYP450* và *UGT*, tương ứng là 196 và 113 unigene. Trong đó, 2 gen thuộc phân họ *CYP716A* và 3 gen *UGT1*, *UGT74AE2*, *UGT94Q2* đã được

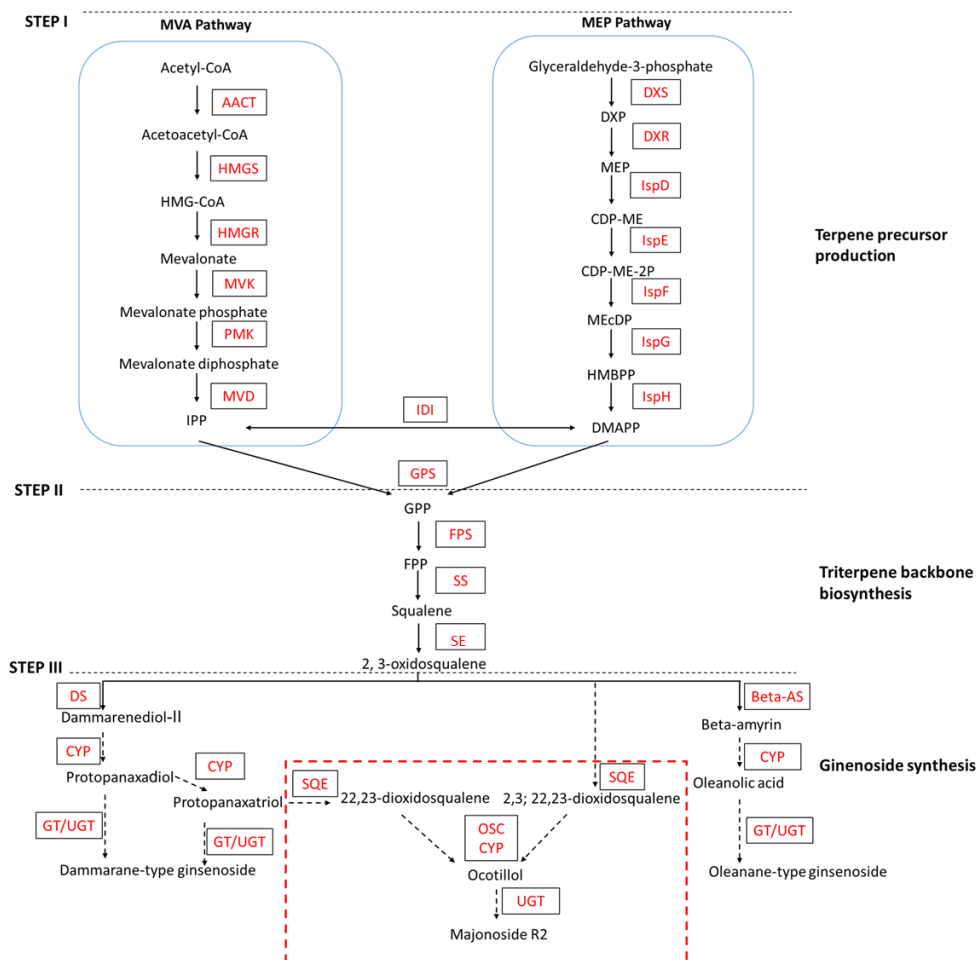
lựa chọn để xây dựng dữ liệu do có hoạt tính trực tiếp trong con đường sinh tổng hợp ginsenoside, cụ thể là ginsenoside thuộc phân nhóm protopanaxadiol và protopanaxatriol [44-45].

Bảng 3. 2 Thống kê số lượng và mức độ biểu hiện tương đối của các unigene tham gia vào con đường sinh tổng hợp ginsenoside từ kết quả giải trình tự hệ gen phiên mã sâm Ngọc Linh

	Tên gen	Kí hiệu	Số lượng unigene
Con đường MVA	Acetyl-CoA acetyltransferase	AACT	4
	Hydroxymethylglutaryl CoA synthase	HMGS	3
	Hydroxymethylglutaryl CoA reductase	HMGR	10
	Mevalonate kinase	MVK	2
	Phosphomevalonate kinase	PMK	2
	Mevalonate diphosphate decarboxylase	MVD	2
	Isopentenyl diphosphate isomerase	IDI	2
	Farnesyl diphosphate synthase	FPS	1
	Geranylgeranyl diphosphate synthase	GGPS	13
	Squalene synthase	SS	4
	Squalene epoxidase	SE	6
	Dammarenediol II synthase	DS	2
	β -amyrin synthase	β -AS	4
	Cycloartenol synthase	CAS	3
	Cytochrome P450	CYP716A52v 2 CYP716A53v 2	1 1
UDP-glycosyl transferase	UGT1 UGT74AE2 UGT94Q2	1 1 2	
Con đường MEP	1-deoxy-D-xylulose-5-phosphate	DXS	9
	1-deoxy-D-xylulose-5-phosphate reductoisomerase	DXR	1
	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase	IspD	1
	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	IspE	2
	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	IspF	1
	4-hydroxy-3-methylbut-2-enyl diphosphate synthase	IspG	1
	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	IspH	7

Từ kết quả sàng lọc này, chúng tôi đặt ra giả thiết rằng con đường sinh tổng hợp ginsenoside của sâm Ngọc Linh có sự tương đồng cao so với các loài cùng chi *Panax*. Sử dụng dữ liệu đã sàng lọc và các nguồn tham khảo [17, 29, 31, 40, 46], chúng tôi đã xây dựng mô hình con đường sinh tổng hợp ginsenoside ở loài sâm Ngọc Linh (Hình 3.2).

Đối với từng gen cụ thể, trình tự của các unigene được phân tích và so sánh với các trình tự tham chiếu của gen tương ứng ở các loài cùng chi *Panax* đã công bố trên GenBank. Unigene có chứa đoạn nucleotide tương đồng trên 90% với trình tự tham chiếu được xem là có mức độ tương đồng cao và sử dụng để hiệu chỉnh trình tự gen bằng kỹ thuật giải trình tự Sanger.



Hình 3. 2 Tổng quan con đường sinh tổng hợp ginsenoside giả định ở loài sâm Ngọc Linh

Kí hiệu: AACT, acetyl-CoA acetyltransferase; beta-AS, β -amyrin synthase; CDP-ME, 4-diphosphocytidyl-2-C-methyl-D-erythritol; CYP, cytochrome P450; DMAPP, dimethylallyl diphosphate; DS, dammarenediol-II synthase; DXS, 1-deoxy-D-xylulose-5-phosphate; DXR, 1-deoxy-D-xylulose-5-phosphate

reductoisomerase; DXS, 1-deoxy-D-xylulose-5-phosphate synthase; FPP, farnesyl diphosphate; FPPS, farnesyl diphosphate synthase; GPP, geranyl pyrophosphate; GPS, geranylgeranyl pyrophosphate synthase; GT, glycosyltransferase; HMBPP, 4-hydroxy-3-methylbut-2-enyl diphosphate; HMG-CoA, 3-hydroxy-3-methylglutaryl coenzyme A; HMGR, HMG-CoA reductase; HMGS, HMG-CoA synthase; IPP, isopentenyl diphosphate; IDI, IPP isomerase; IspD, 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase; IspE, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; IspF, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; IspG, 4-hydroxy-3-methylbut-2-enyl diphosphate synthase; IspH, 4-hydroxy-3-methyl but-2-enyl diphosphate reductase; MVA, mevalonic acid; MEcDP, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate; MEP, 2-C-methyl-D-erythritol 4-phosphate; MVD, mevalonate diphosphate decarboxylase; MVK, mevalonate kinase; OSC, oxidosqualene cyclase; PMK, phosphomevalonate kinase; SQE, squalene epoxidase; SS, squalene synthase; UGT, UDP-glycosyltransferase.

3.2.2. Thống kê dữ liệu các unigene cụ thể

Trong khuôn khổ của đề tài, chúng tôi tập trung nghiên cứu 4 gen liên quan đến sinh tổng hợp ginsenoside theo con đường MVA đó là các gen *AACT*, *IDI*, *SS* và *SE*. Con đường MVA được cho là con đường chủ yếu để tổng hợp ginsenoside và một số gen liên quan đã được nghiên cứu trên các loài thuộc chi *Panax*.

Bảng 3.3 Dữ liệu unigene của một số gen liên quan đến con đường sinh tổng hợp ginsenoside

Gen	Unigene	Chiều dài (bp)	% GC	Dự đoán ORF	Tỉ lệ tương đồng
<i>AACT</i>	c790622_g1_i3	1.536	33,98%	Full ORF	89% AUZ98386.1 <i>Trachyspermum ammi</i>
	c790622_g5_i1	1.061	42,41%	5' partial	89% AUZ98386.1 <i>T. ammi</i>
	c801217_g6_i1	351	33,90%	Internal	dzi:111275290
	c801217_g7_i1	2.430	43,33%	Full ORF	99% KJ804173 <i>P. notoginseng</i>
	c730395_g1_i1	683	41,29%	5' partial	97,52% KJ804170 <i>P. notoginseng</i>

Gen	Unigene	Chiều dài (bp)	% GC	Dự đoán ORF	Tỉ lệ tương đồng
	c803150_g5_i5	1.587	36,93%	Full ORF	77,73% XM_017397983 <i>Daucus carota</i> subsp. <i>Sativus</i>
IDI	c800937_g4_i1	1.361	42,00%	5' partial	92,01% KJ804168 <i>P. notoginseng</i>
	800937_g5_i1	1.337	42,00%	Full ORF	95,83% KJ804168 <i>P. notoginseng</i>
SS	c669272_g1_i1	326	32,82%	Internal	99,09% KP890782 <i>P. japonicus</i>
	c800646_g6_i2	826	31,84%	Internal	100% GU354313 <i>Aralia elata</i>
	c802035_g3_i1	1.665	38,08%	Full ORF	98,77% AM182456 <i>P. quinquefolius</i>
	c802035_g5_i1	1.507	37,43%	Full ORF	96,94% GU997681 <i>P. quinquefolius</i>
SE	c739748_g1_i1	787	42,44%	Internal	91,30% XM_023896099 <i>Lactuca sativa</i>
	c746527_g1_i1	1.200	42,25%	3' partial	87,55% XM_017386662 <i>D. carota</i> subsp. <i>Sativus</i>
	c763721_g3_i1	460	31,96%	Internal	99,28% KJ946468 <i>P. vietnamensis</i> var. <i>fuscidiscus</i>
	c779029_g1_i1	1.192	42,70%	3' partial	86,52% XM_021796019 <i>Hevea brasiliensis</i>
	c780573_g1_i4	1.968	40,65%	Full ORF	98,19% AB122078 <i>P. ginseng</i>

Dữ liệu unigene của gen AACT: Gen *AACT* mã cho enzyme Acetyl-CoA acetyltransferase, là enzyme khởi đầu của quá trình sinh tổng hợp ginsenoside. Enzyme này xúc tác cho phản ứng kết hợp hai phân tử acetyl-CoA để tạo thành acetoacetyl CoA. Ngoài ra, enzyme *AACT* còn có vai trò trong một số hoạt động của sinh vật như phân giải benzoate, tổng hợp butanoate hay tổng hợp tryptophan. Ở các loài thuộc chi *Panax*, gen *AACT*

được phát hiện trong rễ của *P. ginseng* [47], *P. quinquefolium* [48] và *P. notoginseng* [49]. Từ dữ liệu transcriptome, 4 unigene được chú giải là gen *AACT*. Hai trên 4 unigene này mang trình tự ORF đầy đủ của gen *AACT*. Trong các trình tự tham khảo trên các cơ sở dữ liệu, unigene c790622_g1_i3 có độ tương đồng cao nhất (89%) với trình tự mang mã số AUZ98386.1 có nguồn gốc từ *T. ammi*. Unigene chứa ORF đầy đủ còn lại dài 2430 bp, tương đồng 99% với trình tự gen *AACT* từ loài *P. notoginseng* (mã số KJ804173).

Dữ liệu unigene của gen *IDI*: Enzyme *IDI* xúc tác sự chuyển hóa lẫn nhau của isopentenyl diphosphat (IPP) và dimethylallyl diphosphat (DMAPP). Đây là một bước thiết yếu để chuyển hóa mevalonate trong con đường sinh tổng hợp isoprenoid [50]. Trong những năm gần đây, một lượng lớn gen *IDI* đã được nhân bản và xác định. Thực hiện tìm kiếm và phân tích phát sinh loài trên các chuỗi DNA cho thấy có hai họ *IDI* là *IDI1* và *IDI2*, cả hai loại có thể được tìm thấy trong các sinh vật sống theo con đường MEP/MVA [51]. Trong dữ liệu transcriptome sâm Ngọc Linh, 2 unigene được chú giải là mã cho enzyme *IDI* đã được trích xuất. Tuy có chiều dài tương tự nhau nhưng kết quả so sánh với cơ sở dữ liệu cho thấy unigene 800937_g5_i1 chứa đầy đủ vùng ORF và có độ tương đồng cao với trình tự tương ứng có nguồn gốc từ *P. notoginseng* (95,83%). Trong khi đó, unigene c800937_g4_i1 không chứa đầy đủ vùng mã hóa, độ tương đồng với gen *IDI* từ *P. notoginseng* khá cao tuy nhiên độ bao phủ tương đối thấp.

Dữ liệu unigene của gen *SS*: Squalene synthase (*SS*) xúc tác bước đầu tiên từ con đường isoprenoid trung tâm tới sinh tổng hợp sterol và triterpenoid. Enzyme này chịu trách nhiệm cho quá trình sinh tổng hợp squalene từ hai phân tử FPP, một tiền chất quan trọng trong sinh tổng hợp phytosterol và triterpene. *P. ginseng* có 3 gen *SS* chức năng. Lee và đtg (2004) đã phân lập một cDNA đầy đủ của *SS*, được đặt tên là PGSS1, từ thư viện cDNA lá của *P. ginseng*. Trình tự amino acid của PgSS1 có độ tương đồng 84,1%, 75,78%, 81,45% và 71,33% lần lượt với *Glycine max*, *A. thaliana*, *Nicotiana tabacum* và *Oryza sativa* [52]. Bốn unigene được chú giải là gen *SS* được tìm thấy trong dữ liệu transcriptome sâm Ngọc Linh. Hai unigene chỉ chứa một phần bên trong vùng ORF (c669272_g1_i1 và c800646_g6_i2) và 2 unigene (c802035_g3_i1 và c802035_g5_i1) chứa trình

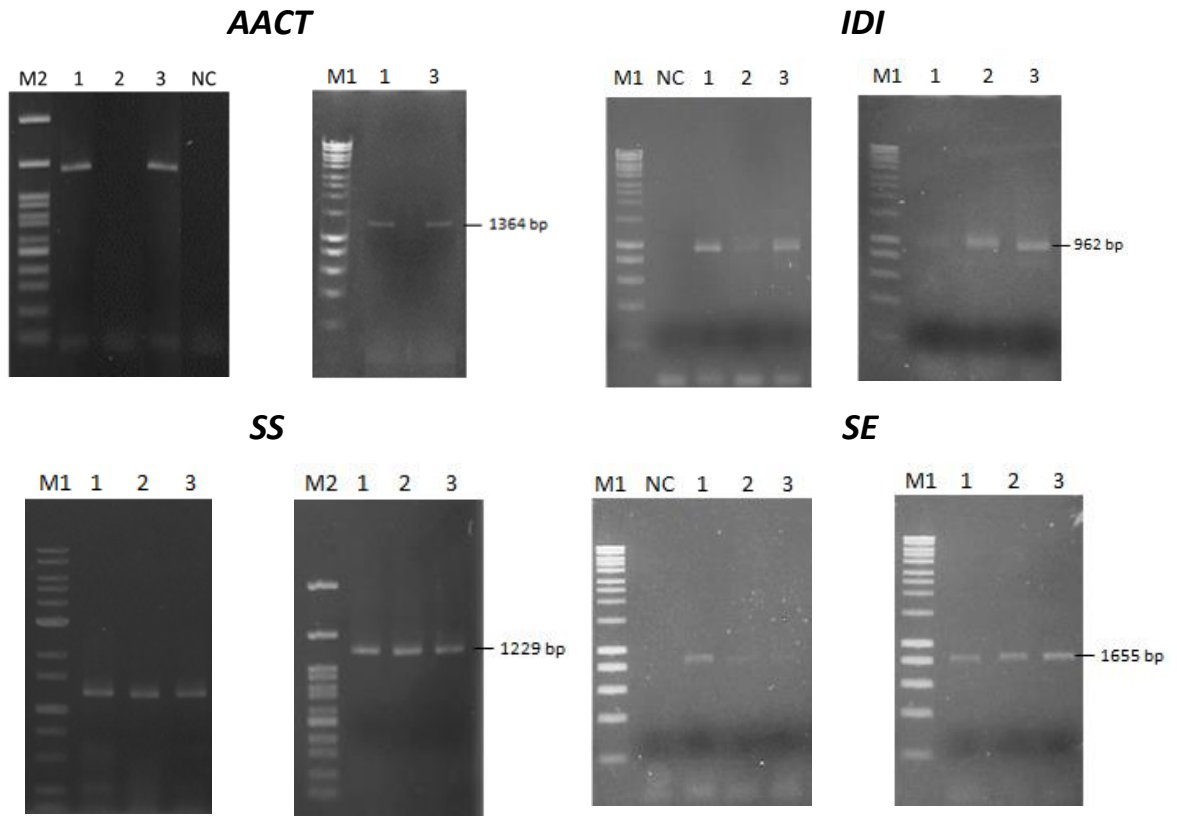
tự đầy đủ của ORF khi so sánh với các trình tự tham chiếu trong các cơ sở dữ liệu. Hai trình tự c802035_g3_i1 và c802035_g5_i1 thuộc cùng cluster và cùng có độ tương đồng cao với các gen *SS* có nguồn gốc từ *P. quinquefolius*. Trong khi unigene c802035_g3_i1 tương đồng 98,77% với trình tự mang mã số AM182456, unigene c802035_g5_i1 tương đồng 96,94% với GU997681.

Dữ liệu unigene của gen *SE*: Squalene epoxidase (*SE*) xúc tác phản ứng oxy hóa liên kết đôi của squalene để tạo ra 2,3-oxidosqualene, bước oxy hóa đầu tiên trong sinh tổng hợp phytosterol và triterpenoid. *P. ginseng* được tìm thấy có 2 bản sao của gen *SE* và trình tự của chúng có độ tương đồng thấp tại đầu N [53]. PgSE1 phiên mã ở tất cả các bộ phận, trong khi PgSE2 có mức độ biểu hiện thấp ngoại trừ ở cuống lá và nụ hoa. Ngoài ra, gen *SE* có nguồn gốc từ củ Tam thất *P. notoginseng* tương đồng 98% với *P. ginseng* [4]. Từ dữ liệu transcriptome sâm Ngọc Linh, có 6 unigene được chú giải liên quan tới gen *SE*. Trong số các unigene được tìm thấy, 2 unigene chứa đầy đủ vùng ORF (c780573_g1_i4 và c782799_g1_i1), 2 unigene ngắn dưới 700 bp chỉ thuộc một phần trong vùng ORF (c739748_g1_i1 và c763721_g3_i1) và 2 unigene chứa một phần vùng ORF mang mã kết thúc. Hai unigene c780573_g1_i4 và c782799_g1_i1 có độ tương đồng cao trên 98% với các gen *SE* có nguồn gốc từ loài *P. ginseng* và *P. vietnamensis* var. *fuscidiscus*.

3.3. KẾT QUẢ NHÂN BẢN VÀ XÁC ĐỊNH TRÌNH TỰ MỘT SỐ GEN LIÊN QUAN ĐẾN CON ĐƯỜNG SINH TỔNG HỢP GINSENOSE Ở SÂM NGỌC LINH SỬ DỤNG PHƯƠNG PHÁP SANGER

Dựa vào các trình tự tham chiếu có trên cơ sở dữ liệu GenBank cũng như trình tự các gen thu được trong dữ liệu hệ gen phiên mã đã xác định bằng phương pháp giải trình tự gen thế hệ mới, các cặp mồi đã thiết kế (Bảng 2.2) được sử dụng để khuếch đại các gen quan tâm, bao gồm *AACT*, *IDI*, *SS*, *SE*. Kết quả PCR khuếch đại các gen thành công được thể hiện ở Hình 3.3. Các sản phẩm PCR sau đó được nhân lượng lớn, tinh sạch, điện di kiểm tra trên gel agarose đảm bảo có một băng duy nhất, sáng, rõ nét và sử dụng để giải trình tự bằng phương pháp Sanger. Sau các bước xử lý, chúng tôi đã thu được trình tự cDNA của các gen nghiên cứu ở sâm Ngọc Linh. Sử dụng phần mềm Bioedit, trình tự cDNA của các gen được so sánh với trình tự tham chiếu có

trên cơ sở dữ liệu và các trình tự được chú giải tương ứng từ dữ liệu hệ gen phiên mã sâm Ngọc Linh, từ đó, chỉ ra những điểm sai khác của trình tự gen thu được bằng hai phương pháp giải trình tự và trình tự gen tương ứng của các loài thuộc chi *Panax* (Phụ lục 3.1).



Hình 3. 3 Kết quả khuếch đại một số gen liên quan đến sinh tổng hợp ginsenoside ở sâm Ngọc Linh

M1: 1 kb (Bioline), M2: 100 bp (Nippon); NC: Đối chứng âm với khuôn là nước; 1-3: Sản phẩm PCR với khuôn là cDNA sâm Ngọc Linh mẫu 1-3. Đối với mỗi gen: Hình trái: Sản phẩm PCR nhân gen; Hình phải: Sản phẩm PCR tinh sạch.

AACT: Cặp mồi AACT F/R được sử dụng để khuếch đại vùng CDS của gen AACT. Cặp mồi được thiết kế để khuếch đại từ phần đầu trình tự CDS chứa mã mở đầu đến phần CDS chứa mã kết thúc. Điều kiện khuếch đại gen AACT: 94°C, 3 phút; 40 chu kỳ (94°C-1 phút; 58°C-30 giây; 72°C-45 giây); 72°C-2 phút, sau đó giữ ở 4°C. Kết quả PCR trên Hình 3.4 cho thấy băng sản phẩm có kích thước tương ứng kích thước lý thuyết (1.364 bp). Trình tự

cDNA hoàn chỉnh gen *AACT* phân lập từ sâm Ngọc Linh, *PvAACT*, có độ dài 1.315 bp, vùng CDS ở vị trí 9-1235. Trình tự này được công bố trên GenBank với mã số MZ272018. So sánh với trình tự tham chiếu *AACT* của *P. notoginseng* (KJ804173) là trình tự đầy đủ duy nhất trên cơ sở dữ liệu, trình tự *PvAACT* có độ tương đồng 98,86% với 15 vị trí sai khác và 3 vị trí trống. Đối với các trình tự *AACT* trích xuất từ dữ liệu transcriptome, 2 unigene c790622_g5_i1 và c790622_g1_i3 thuộc cùng một cluster có sự tương đồng với trình tự *AACT* từ *T. ammi*. Có thể thấy được khi so sánh sắp hàng, 2 unigene này tương đồng rất thấp với *PvAACT*. Unigene c801217_g7_i1 có độ tương đồng cao nhất (99,85%) với trình tự *PvAACT* thu được từ phương pháp Sanger. Qua so sánh sắp hàng trình tự, hai trình tự này chỉ có 2 điểm sai khác và 1 vị trí trống.

IDI: Cặp mồi IDI F/R được thiết kế để khuếch đại đoạn gen *IDI* có chứa vùng CDS có kích thước dự kiến là 965 bp, trong đó mồi xuôi IDI F ở vùng 5' _UTR và vùng CDS có chứa bộ ba mở đầu và mồi ngược IDI R ở vùng 3' _UTR. Điều kiện khuếch đại sử dụng cho cặp mồi này như sau: 94°C-3 phút; 40 chu kì (94°C-45 giây; 55°C-30 giây; 72°C-1 phút 30 giây); 72°C-2 phút, sau đó giữ ở 4°C. Kết quả PCR với cặp mồi thiết kế cho thấy sản phẩm PCR thu được có kích thước phù hợp với lý thuyết. Trình tự cDNA gen *IDI*, *PvIDI* thu được có độ dài 965 bp, mang đầy đủ vùng CDS từ vị trí 6-932. Trình tự này được công bố trên GenBank với mã số OK513075. Kết quả so sánh trình tự cho thấy *PvIDI* khá gần gũi với trình tự của gen này ở *P. notoginseng* (độ tương đồng 97,51%). Sự sai khác giữa hai trình tự này xảy ra do thay thế nucleotide đơn. Đối với các unigene mã cho enzyme IDI chọn lọc từ dữ liệu transcriptome, unigene c800937_g5_i1 có độ tương đồng với *PvIDI* cao hơn (96,89%). Trong khi đó, unigene c800937_g4_i1 tương đồng chủ yếu với *PvIDI* ở vùng đầu 5'. Do đó, độ tương đồng của unigene này với *PvIDI* đạt 93,77% nhưng với độ bao phủ chỉ 35%. Sự sai khác giữa trình tự *IDI* từ sâm Ngọc Linh thu được từ hai phương pháp bắt nguồn từ tính chất của giải trình tự transcriptome là giải từng đoạn ngắn và sử dụng phần mềm lắp ráp dựa trên trình tự tham chiếu nhưng nguồn tham chiếu từ *P. vietnamensis* còn hạn chế.

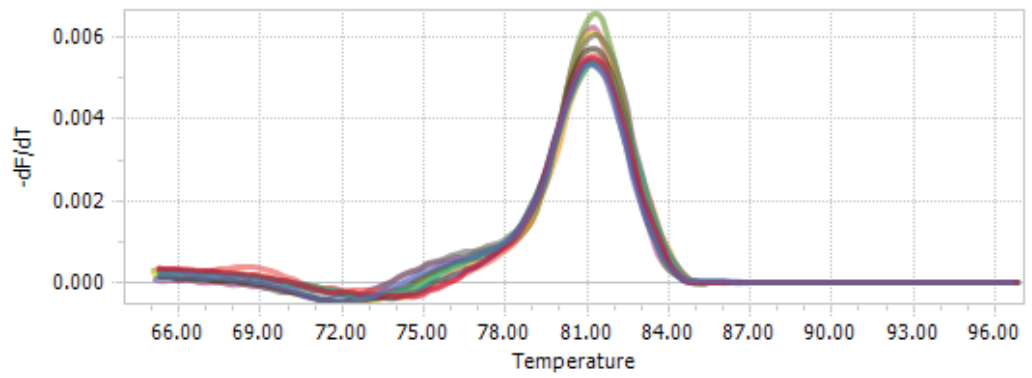
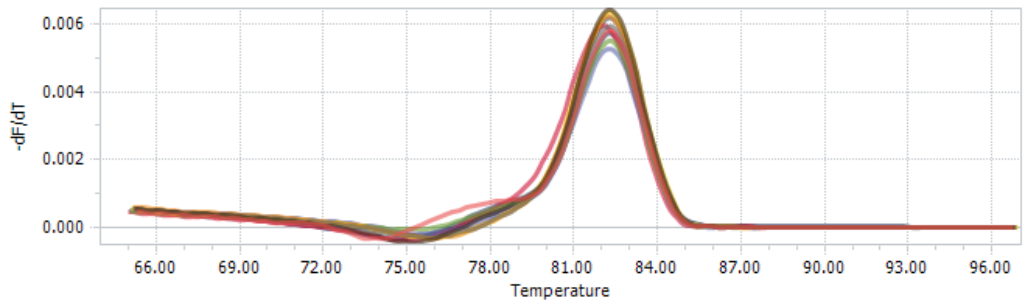
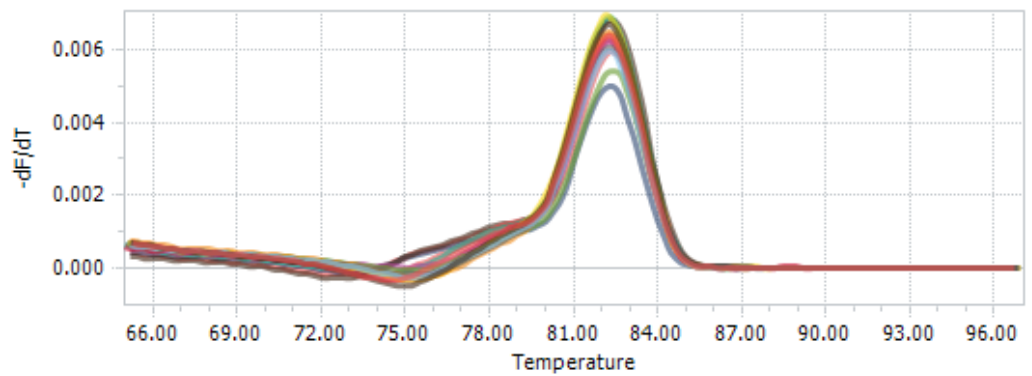
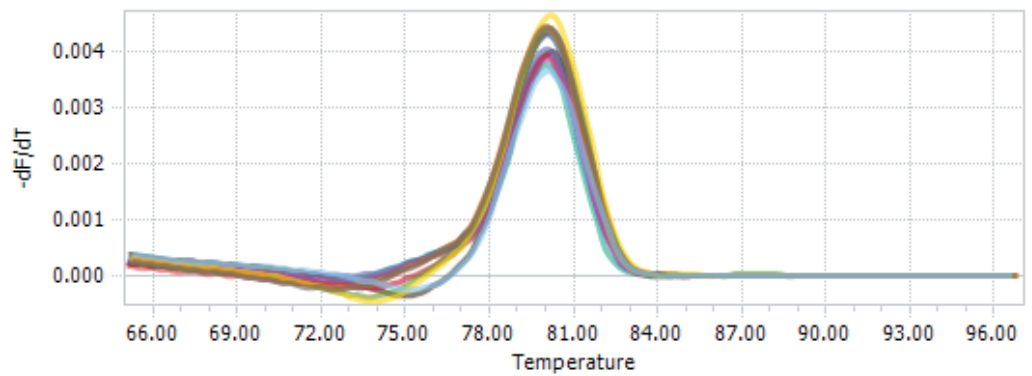
SS: Cặp mồi SS F/R cho phản ứng khuếch đại gen SS, trong đó mồi SS F được thiết kế từ mã mở đầu của gen trong khi mồi SS R thuộc vùng cuối của khung đọc mở. Kích thước sản phẩm theo lý thuyết là 1229 bp. Điều kiện khuếch đại sử dụng cho cặp mồi này như sau: 94°C-3 phút; 40 chu kì (94°C-45 giây; 55°C-45 giây; 72°C-70 giây); 72°C-2 phút, sau đó giữ ở 4°C. Trình tự gen thu được, *PvSS*, có độ dài 1195bp, chứa mã mở đầu và phần lớn vùng CDS của gen SS. Trình tự này được công bố trên GenBank với mã số OK513077. Tra cứu trên GenBank, chúng tôi thu thập được nhiều trình tự gen SS đến từ một số loài thuộc chi *Panax* như *P. ginseng*, *P. quinquefolius* hay *P. japonicus*. Các trình tự này khá đa dạng. Trong đó, trình tự *PvSS* chúng tôi thu được từ sâm Ngọc Linh có độ tương đồng cao nhất với trình tự gen SS2 có nguồn gốc từ *P. ginseng* (GQ468527). Từ dữ liệu transcriptome, chúng tôi thu được 4 unigene được chú giải là gen SS, có độ dài từ 320 bp đến 1665 bp. Unigene c669272_g1_i1 và c800646_g6_i2 không tương đồng với trình tự *PvSS*. Mặc dù, unigene c669272_g1_i1 cho thấy có tương đồng với trình tự SS của *P. japonicus* (KP890782), tuy nhiên sự tương đồng chỉ xảy ra ở 100 nucleotide cuối và KP890782 cũng có khác biệt với *PvSS*. Vì vậy, xét trên toàn trình tự c669272_g1_i1 không tương đồng với *PvSS*. Trong khi đó, c800646_g6_i2 có độ tương đồng cao với trình tự SS GU354313 của *Aralia elata* không phải thành viên chi *Panax*. Hai unigene còn lại đều chứa trình tự ORF đầy đủ. Unigene c802035_g3_i1 tương đồng 100% với trình tự thu được từ phương pháp Sanger. Unigene c802035_g5_i1 chỉ tương đồng với *PvSS* ở 300 nucleotide đầu.

SE: Để khuếch đại gen *SE*, chúng tôi thiết kế 2 cặp mồi, SE F1R1 được sử dụng cho phản ứng khuếch đại đoạn trình tự dài 802 bp chứa mã mở đầu (kí hiệu là SE1) và SE F2R2 được sử dụng cho phản ứng khuếch đại đoạn trình tự dài 874 bp chứa mã kết thúc (kí hiệu là SE2). Hai đoạn DNA này có phần trình tự chồng gối lên nhau thuận lợi cho quá trình ghép nối trình tự. Điều kiện khuếch đại sử dụng cho hai cặp mồi này như sau: 94°C-3 phút; 40 chu kì (94°C-45 giây; 64°C-30 giây; 72°C-30 giây); 72°C-2 phút, sau đó giữ ở 4°C. Kết quả PCR với hai cặp mồi SE F1R1 và SE F2R2, cho thấy chúng tôi không thu được sản phẩm PCR với cặp mồi SE F1R1. Đối với cặp mồi SE F2R2, chúng tôi thu được sản phẩm có kích thước đúng với lý thuyết. Cuối cùng, chúng tôi đã thu được trình tự cDNA một phần gen *SE*, *PvSE2*, có độ

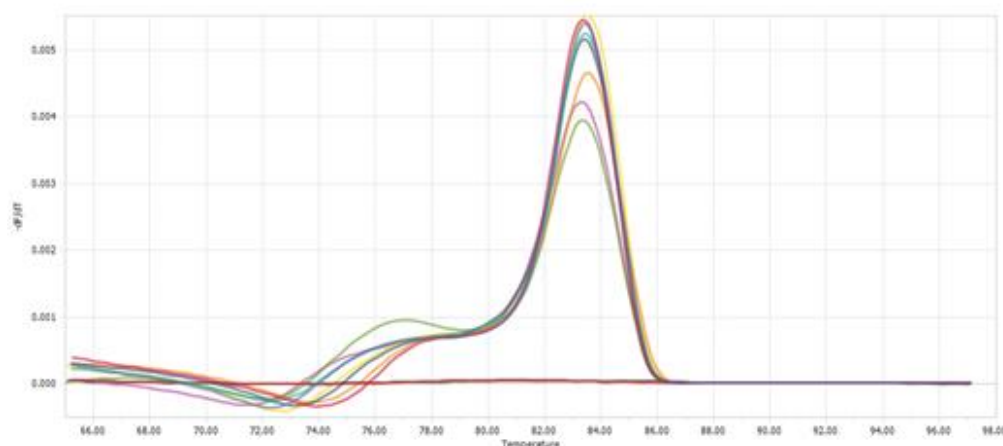
dài 824 bp chứa mã kết thúc. Trình tự này được công bố trên GenBank với mã số OK513079. Đối với các trình tự tham chiếu của các loài thuộc chi *Panax* có trên NCBI, trình tự *PvSE2* có độ tương đồng cao nhất với *SE* của *P. vietnamensis* var. *fuscidiscus* (99,15%). Điều này có thể lý giải vì đây là hai thứ của *P. vietnamensis*. Độ tương đồng thấp nhất thuộc về trình tự *SE* của *P. sokpayensis* (86,41%). Các trình tự tham chiếu còn lại có độ tương đồng từ 87% đến 98%. Các sai khác chủ yếu xảy ra do thay đổi nucleotide đơn. Đối với các unigene tương ứng chọn lọc từ dữ liệu transcriptome, unigene *c780573_g1_i4* có độ tương đồng cao nhất với trình tự thu được bằng phương pháp Sanger (99,51%). Các unigene còn lại có độ sai khác lớn do kích thước nhỏ, tương đồng với các trình tự *SE* thuộc loài xa với chi *Panax*.

3.3. KẾT QUẢ PHÂN TÍCH BIỂU HIỆN CÁC GEN LIÊN QUAN ĐẾN CON ĐƯỜNG SINH TỔNG HỢP GINSENOSE Ở MẪU LÁ SÂM NGỌC LINH 1 VÀ 4 NĂM TUỔI

Sau khi thiết kế trình tự các cặp mồi, bước tiếp theo là kiểm tra độ đặc hiệu cho phản ứng khuếch đại các gen nghiên cứu. Độ đặc hiệu này được kiểm tra bằng cách thiết kế phản ứng PCR định lượng cho từng cặp mồi, trong đó thay đổi nhiệt độ gắn mồi cho mỗi phản ứng. Sản phẩm sau phản ứng được biến tính bằng cách nâng nhiệt độ để xác định nhiệt độ biến tính của sản phẩm. Mồi đặc hiệu sẽ cho một sản phẩm khuếch đại duy nhất, do đó nhiệt độ biến tính sẽ thu được 1 giá trị duy nhất. Bất kỳ phản ứng nào có nhiều hơn 1 nhiệt độ biến tính của sản phẩm đều chứng tỏ phản ứng khuếch đại gen không đặc hiệu, do đó, mồi sẽ không được sử dụng để kiểm tra biểu hiện gen. Kết quả kiểm tra cho thấy, các mồi *AACT*, *IDI*, *SS* và *SE* đều đặc hiệu cho các gen mục tiêu tương ứng. Kết quả tối ưu hóa các mồi được thể hiện trong Hình 3.4.

EF**AACT****IDI****SS**

SE



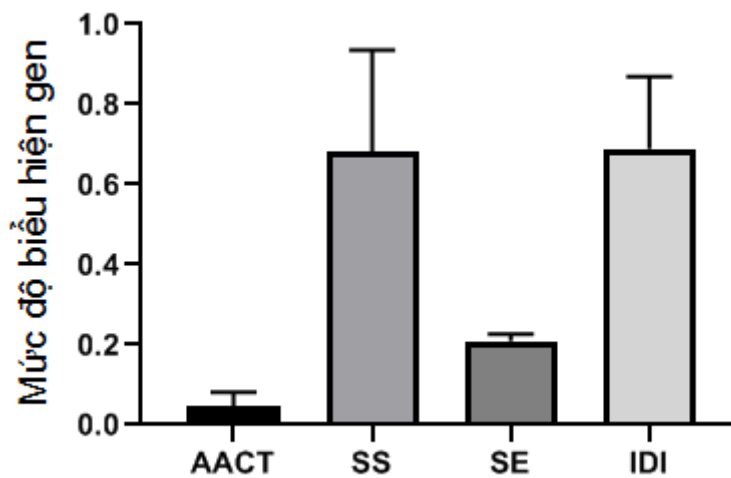
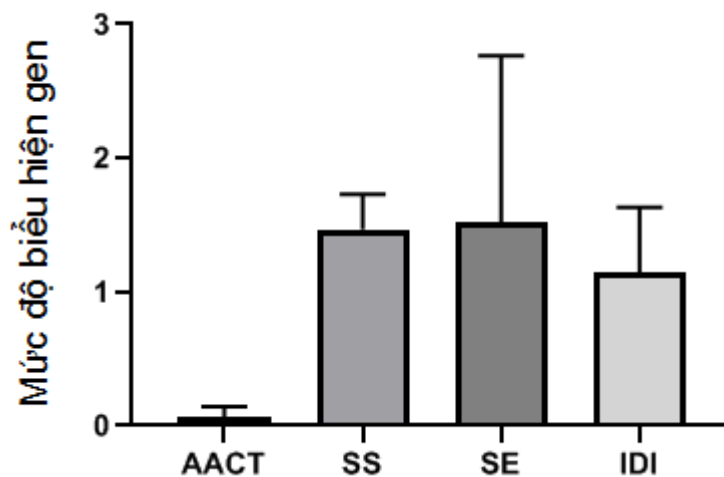
Hình 3. 4 Tối ưu hóa môi khuếch đại các gen *EF*, *AACT*, *IDI*, *SS*, *SE*

Kết quả biểu hiện của các gen liên quan đến quá trình tổng hợp ginsenoside ở lá sâm Ngọc Linh 1 tuổi được thể hiện trên Hình 3.5A và Bảng 3.4. Kết quả phân tích cho thấy, trong lá sâm Ngọc Linh 1 năm tuổi, các gen nghiên cứu liên quan đến tổng hợp ginsenoside có mức độ biểu hiện khác nhau. Trong đó, gen *IDI* có mức biểu hiện cao nhất, sau đó đến các gen *SS*, *SE*, và *AACT* biểu hiện thấp nhất. Phân tích tương tự cho các gen liên quan ở lá sâm Ngọc Linh 4 năm tuổi (Hình 3.5B) cho thấy, các gen *SE*, *SS*, *IDI* biểu hiện cao, trong khi gen *AACT* biểu hiện rất thấp.

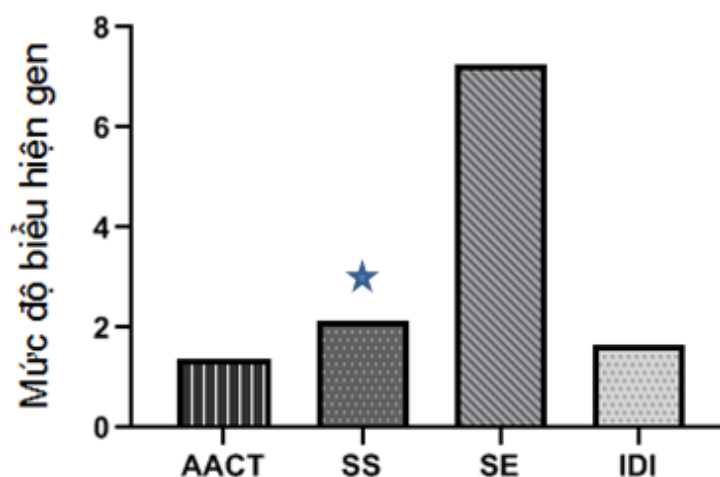
Bảng 3.4 So sánh mức độ biểu hiện của một số gen liên quan đến con đường sinh tổng hợp ginsenoside ở mô lá và thân rễ sâm Ngọc Linh 1 năm tuổi và 4 năm tuổi

Tên gen		<i>AACT</i>	<i>IDI</i>	<i>SS</i>	<i>SE</i>
Lá 1 tuổi	Trung bình	0,047	0,689	0,683	0,209
	SD	0,034	0,181	0,252	0,018
Lá 4 tuổi	Trung bình	0,065	1,146	1,468	1,515
	SD	0,079	0,486	0,262	1,252
Thân rễ 1 tuổi	Trung bình	0,242	0,824	0,715	0,152
	SD	0,040	0,076	0,063	0,030
Thân rễ 4 tuổi	Trung bình	0,274	2,182	0,718	1,035
	SD	0,222	1,881	0,568	0,628

Hình 3.5C trình bày kết quả so sánh mức độ biểu hiện của các gen liên quan đến quá trình tổng hợp ginsenoside ở lá sâm Ngọc Linh 4 năm tuổi so với 1 năm tuổi. Kết quả phân tích cho thấy có sự khác biệt trong sự biểu hiện gen của các gen nghiên cứu. Gen *SE* có mức biểu hiện ở lá 4 năm tuổi cao hơn nhiều so với lá 1 năm tuổi. Gen *SS* biểu hiện cao hơn ở lá sâm 4 tuổi so với lá sâm 1 tuổi khoảng 2 lần. Một số gen khác như *AACT*, *IDI* biểu hiện tăng không đáng kể trong lá ở 2 độ tuổi khác nhau. Có thể nhận định rằng hầu hết các gen liên quan đến tổng hợp ginsenoside ở lá sâm 4 năm tuổi đều biểu hiện tăng theo các mức độ khác nhau so với lá của cây 1 năm tuổi.

A**B**

C



Hình 3. 5 Biểu hiện của các gen liên quan đến sinh tổng hợp ginsenoside ở mô lá sâm Ngọc Linh

A: Biểu hiện của các gen ở mô lá sâm 1 năm tuổi; B: Biểu hiện của các gen ở mô lá sâm 4 năm tuổi; C: Tỷ lệ mức độ biểu hiện của các gen liên quan đến tổng hợp ginsenoside ở mô lá sâm 4 năm tuổi và 1 năm tuổi. Dấu sao chỉ ra sự khác biệt có ý nghĩa thống kê ($p < 0,05$) theo kiểm định t -test

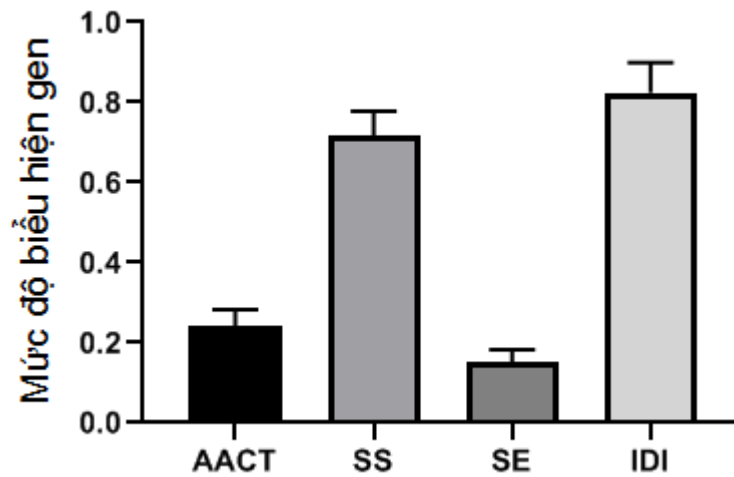
3.4. KẾT QUẢ PHÂN TÍCH BIỂU HIỆN CÁC GEN LIÊN QUAN ĐẾN CON ĐƯỜNG SINH TỔNG HỢP GINSENOSE Ở MÃU THÂN RỄ SÂM NGỌC LINH 1 VÀ 4 NĂM TUỔI

Kết quả phân tích (Hình 3.6) cũng cho thấy, các gen liên quan đến tổng hợp ginsenoside có mức độ biểu hiện khác nhau trong mẫu thân rễ sâm Ngọc Linh 1 năm tuổi và 4 năm tuổi. Ở sâm 1 năm tuổi, gen *IDI* thể hiện mức độ biểu hiện cao nhất, sau đó đến gen *SS*, trong khi các gen còn lại *AACT* và *SE* có mức biểu hiện thấp (Hình 3.6A). Ở sâm 4 năm tuổi, gen *IDI* biểu hiện cao. Gen *SS*, *SE* biểu hiện ở mức độ trung bình, trong khi gen *AACT* biểu hiện thấp hơn (Hình 3.6B).

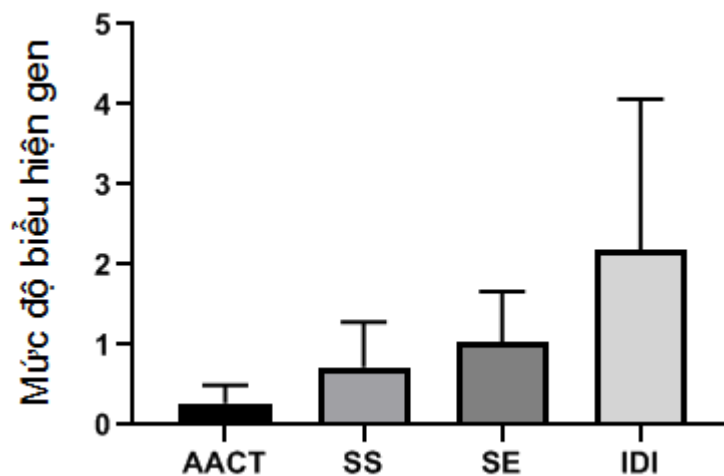
So sánh mức độ biểu hiện của các gen ở mô thân rễ sâm 4 năm tuổi so với 1 năm tuổi cho thấy có sự khác biệt trong sự biểu hiện gen của các gen liên quan đến tổng hợp ginsenoside (Hình 3.6C). Trong đó, gen *SE* biểu hiện cao hơn 6 lần ở thân rễ sâm 4 năm tuổi so với sâm 1 năm tuổi. Gen *IDI* có mức độ biểu hiện tăng khoảng 2 lần. Kết quả này cũng tương đồng với kết quả của Li và đtg (2019) thể hiện kết quả biểu hiện cao của gen *SE* ở thân rễ sâm 3 năm tuổi so với sâm 1 năm tuổi [49]. Một số gen khác như *AACT* và *SS*

biểu hiện tăng không đáng kể trong mô thân rễ ở 2 độ tuổi. Tương tự như ở mô lá, hầu hết các gen liên quan đến tổng hợp ginsenoside ở thân rễ sâm 4 năm tuổi đều biểu hiện tăng theo các mức độ khác nhau so với ở thân rễ của cây 1 năm tuổi. Năm 2008, He và đtg nghiên cứu biểu hiện của gen mã hóa squalene epoxidase đóng vai trò quan trọng trong con đường tổng hợp ginsenoside ở *P. notoginseng* cũng đã chỉ ra độ tuổi của cây càng cao thì mức độ phiên mã của gen này càng tăng [39]. Mức độ biểu hiện gen khác nhau ở 2 loại mô, thuộc 2 độ tuổi khác nhau có thể quyết định đến sự khác biệt trong hàm lượng các hoạt chất ginsenoside tương ứng ở sâm 1 năm tuổi và 4 năm tuổi.

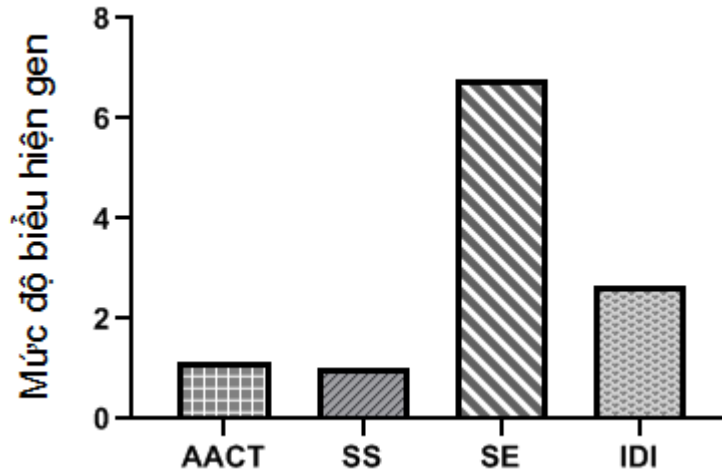
A



B



C



Hình 3. 6 Biểu hiện của các gen liên quan đến sinh tổng hợp ginsenoside ở mô thân rễ sâm Ngọc Linh

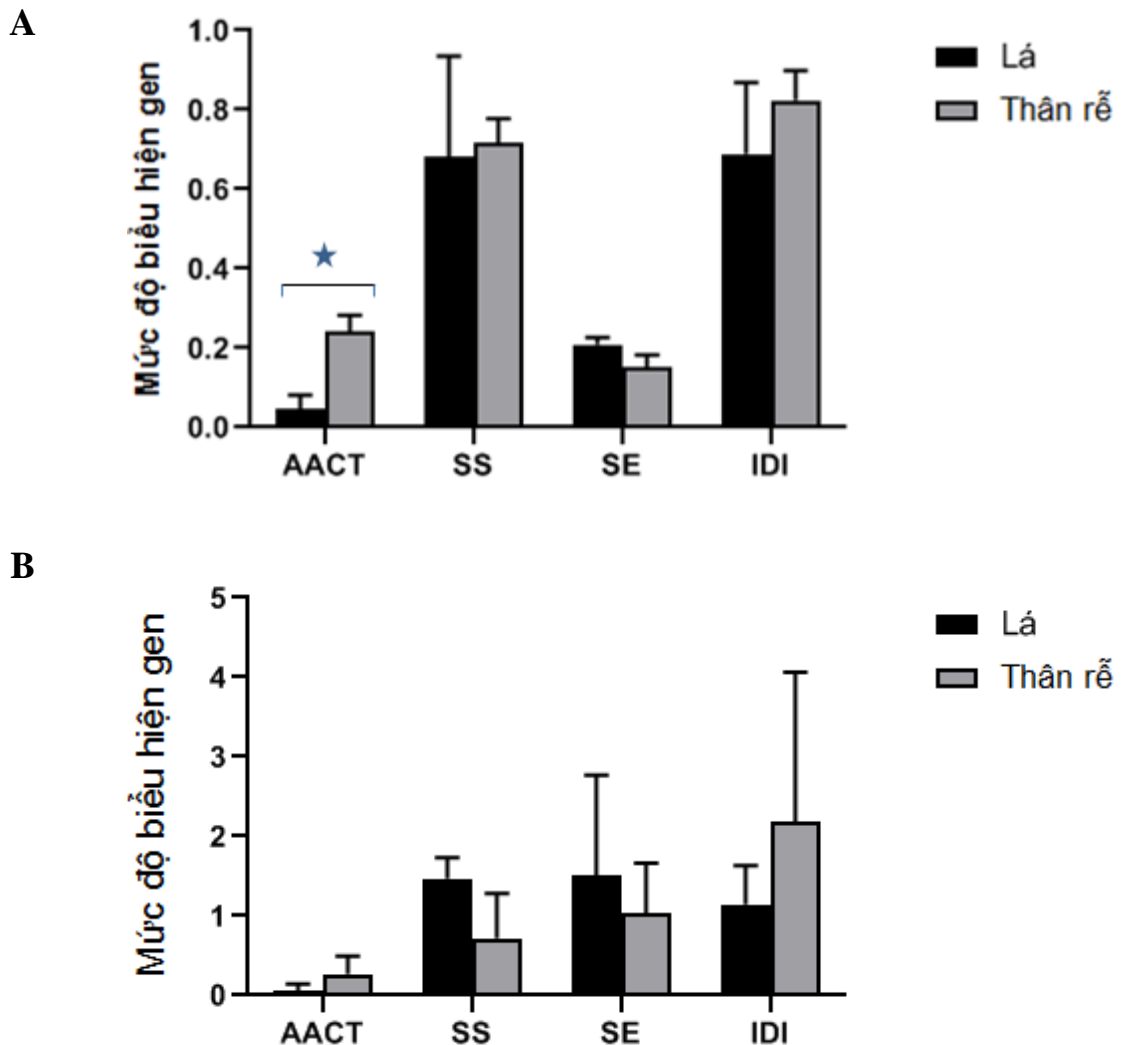
A: Biểu hiện của các gen ở thân rễ sâm 1 năm tuổi; B: Biểu hiện của các gen ở thân rễ sâm 4 năm tuổi (B); C: Tỷ lệ mức độ biểu hiện của các gen liên quan đến tổng hợp ginsenoside ở mô thân rễ sâm 4 tuổi và 1 năm tuổi

3.5. KẾT QUẢ SO SÁNH MỨC BIỂU HIỆN CỦA CÁC GEN LIÊN QUAN ĐẾN TỔNG HỢP GINSENOSE GIỮA MÔ LÁ VÀ THÂN RỄ SÂM NGỌC LINH Ở HAI ĐỘ TUỔI

So sánh kết quả biểu hiện của các gen liên quan đến quá trình tổng hợp ginsenoside ở mô lá và thân rễ sâm Ngọc Linh 1 năm tuổi được thể hiện trên Hình 3.7 và Bảng 3.4. Kết quả phân tích cho thấy, các gen liên quan đến tổng hợp ginsenoside có mức độ biểu hiện khác nhau giữa mô lá và thân rễ sâm 1 năm tuổi (Hình 3.7A). Hầu hết các gen nghiên cứu đều có mức biểu hiện ở thân rễ cao hơn ở lá, ngoại trừ gen *SE* có mức biểu hiện ngược lại, ở lá cao hơn so với thân rễ. Trong đó, gen *AACT* có mức biểu hiện ở thân rễ cao hơn ở lá khoảng 5 lần. Kết quả này phù hợp với một số nghiên cứu đã công bố. Dựa vào hệ gen phiên mã của *P. notoginseng*, Liu và đtg (2015) đã phân tích và so sánh sự biểu hiện các gen liên quan đến quá trình sinh tổng hợp ginsenoside ở mô lá, thân rễ và hoa. Kết quả cũng cho thấy gen mã hóa protein *AACT* có mức độ biểu hiện ở thân rễ cao hơn so với lá và hoa [31].

Đối với sâm 4 năm tuổi, kết quả phân tích và so sánh cho thấy gen *AACT*, *IDI* có mức biểu hiện ở thân rễ cao hơn ở lá. Các gen *SS* và *SE* có mức

biểu hiện ngược lại, ở lá cao hơn so với thân rễ. Các gen được phát hiện này sẽ là nguồn tài nguyên quý giá để nghiên cứu về cơ chế phân tử của quá trình sinh tổng hợp ginsenoside. Như vậy, có thể thấy mức độ biểu hiện gen giữa mô lá và mô thân rễ cũng có những điểm khác biệt ở 2 độ tuổi của cây (1 và 4 năm tuổi), có thể tương ứng với thay đổi thành phần ginsenoside giữa các loại mô ở các độ tuổi.



Hình 3. 7 So sánh mức biểu hiện của các gen liên quan đến sinh tổng hợp ginsenoside ở mô lá và thân rễ sâm Ngọc Linh 1 và 4 năm tuổi

A: So sánh mức biểu hiện của các gen liên quan đến sinh tổng hợp ginsenoside ở mô lá và thân rễ sâm Ngọc Linh 1 năm tuổi; B: So sánh mức biểu hiện của các gen liên quan đến sinh tổng hợp ginsenoside ở mô lá và thân rễ sâm Ngọc Linh 4 năm tuổi. Dấu sao chỉ ra sự khác biệt có ý nghĩa thống kê ($p < 0,05$) theo kiểm định t -test

KẾT LUẬN VÀ KIẾN NGHỊ

KẾT LUẬN

1. Số lượng và mức độ biểu hiện tương đối của các unigene tham gia vào con đường sinh tổng hợp ginsenoside từ kết quả giải trình tự hệ gen phiên mã sâm Ngọc Linh đã được xác định. Mô hình con đường giả định sinh tổng hợp ginsenoside ở sâm Ngọc Linh đã được xây dựng.
2. Bốn gen tham gia vào con đường sinh tổng hợp ginsenoside (*AACT*, *IDI*, *SS*, *SE*) ở sâm Ngọc Linh đã được nhân bản và xác định trình tự cDNA.
3. Bốn gen (*AACT*, *IDI*, *SS*, *SE*) đã được đánh giá mức độ biểu hiện ở hai loại mô và hai độ tuổi (1 năm tuổi và 4 năm tuổi) của sâm Ngọc Linh. Các gen nghiên cứu có mức độ biểu hiện khác nhau giữa mô lá và thân rễ ở hai độ tuổi. Đặc biệt, ở sâm 1 năm tuổi và 4 năm tuổi, các gen *AACT* và *IDI* có mức biểu hiện ở thân rễ cao hơn ở lá. Các gen được phát hiện là nguồn thông tin hữu ích hỗ trợ nghiên cứu cơ chế phân tử của quá trình sinh tổng hợp ginsenoside ở sâm Ngọc Linh.

KIẾN NGHỊ

Kết quả đạt được của đề tài là cơ sở để tiếp tục mở rộng các nghiên cứu xây dựng dữ liệu toàn diện về gen và hệ gen của sâm Ngọc Linh, góp phần cung cấp thông tin đầy đủ và hữu ích về di truyền, các gen chức năng liên quan đến các tính trạng quan trọng, phục vụ công tác giám định, đánh giá nguồn gen, quản lý, bảo tồn, chọn tạo giống sâm Ngọc Linh chất lượng bằng công nghệ sinh học.

DANH MỤC CÔNG TRÌNH CỦA TÁC GIẢ

Công bố: 01 bài báo được đăng trên tạp chí quốc tế (đồng tác giả), 01 bài báo đã được công bố trên tạp chí quốc gia (tác giả chính):

1. Luu LH, Pham HL, Huynh HT, Nguyen LN, **Vu TT**, Le HT (2022). Differential expression of ginsenoside biosynthesis-related genes at an early developmental stage of *Panax vietnamensis*. *Plant Biotechnology Reports*, 16(2), 215-228.
2. **Vũ Thị Trinh**, Luu Hàn Ly, Huỳnh Thị Thu Huệ, Lê Thị Thu Hiền (2021) Phân lập, xác định trình tự và đánh giá biểu hiện của gen mã hóa Acetoacetyl-coa thiolase (AACT) ở sâm Ngọc Linh (*Panax vietnamensis* Ha et Grushv.). *Tạp chí Công nghệ sinh học*, 19(1): 107-117, ISSN: 1811-4989.

TÀI LIỆU THAM KHẢO

1. Jung J., Lee NK., Paik HD., 2017, Bioconversion, health benefits, and application of ginseng and red ginseng in dairy products, *Food Sci Biotechnol*, 26, pp. 1155-1168.
2. Patel S., Rauf A., 2017, Adaptogenic herb ginseng (*Panax*) as medical food: status quo and future prospects, *Biomed Pharmacother*, 85, pp. 120-127.
3. Ha TD., Grushvitzky IV., 1985, A new species of the genus *Panax* (Araliaceae) from Vietnam, *Bot Zhurn*, 70, pp. 519-522.
4. Nguyễn Tập, 2005, Các loài thuộc chi *Panax* ở Việt Nam, *Tạp chí Dược liệu*, 10(3), tr. 71-76.
5. Sách Đỏ Việt Nam, Phần II, Thực vật, 2007, *Nhà xuất bản Khoa học tự nhiên và Công nghệ, Hà Nội*.
6. Lã Đình Mỡi, Châu Văn Minh, Trần Văn Sung, Phạm Quốc Long, Phan Văn Kiệm, Trần Huy Thái, Trần Minh Hợi, Ninh Khắc Bản, Lê Mai Hương, 2013, Họ Nhân sâm (Araliaceae Juss.) - Nguồn hoạt chất sinh học đa dạng và đầy triển vọng ở Việt Nam, *Hội nghị khoa học Toàn quốc về Sinh thái và Tài nguyên sinh vật lần thứ 5*, tr. 1152-1158.
7. Phan Kế Long, 2013, Báo cáo tổng kết: “Nghiên cứu phân loại, phân bố và thành phần hóa học của cây Sâm mọc tự nhiên ở Lai Châu”, Đề tài thuộc 7 hướng KHCN ưu tiên cấp Viện Hàn lâm Khoa học và Công nghệ Việt Nam 2012-2013.
8. Nguyễn Ngọc Dung, 1995, Nhân giống sâm Ngọc Linh (*Panax vietnamensis* Ha et Grushv.) bằng con đường sinh học, *Nhà xuất bản Nông nghiệp*, tr. 43-100.
9. Xiang QB., Lowry IIPP., 2007, *Panax*: 489-491, *Flora of China*, tr. 13.
10. Roskov Y., Kunze T., Orrell T., Abucay L., Paglinawan L., Culham A., Bailly N., Kirk P., Bourgoin T., Baillargeon G, 2014, Species 2000 & ITIS catalogue of life. 2014 Annual Checklist. Digital resource at, www.catalogueoflife.org/annual-checklist/2014. Species 2000: Naturalis, Leiden, the Netherlands.
11. Shin BK., Kwon SW., Park JH., 2015, Chemical diversity of ginseng saponins from *Panax ginseng*, *J Ginseng Res*, 39(4), pp. 287-298.

12. Case MA., Flinn KM., Jancaitis J., Alley A., Paxton A., 2007. Declining abundance of American ginseng (*Panax quinquefolius* L.) documented by herbarium specimens, *Biol Conserv*, 134, pp. 22-30.
13. McGraw JB., Lubbers AE., Van der Voort M., Mooney EH., Furedi MA., Souther S., Turner JB., Chandler J., 2013, Ecology and conservation of ginseng (*Panax quinquefolius*) in a changing world, *Ann NY Acad Sci* 1286, pp. 62-91
14. Trần Ngọc Lê, Nguyễn Tiến Dũng, Nguyễn Thị Thu, Lê Thị Thu Hiền, Ngô Hoàng Linh, Nguyễn Đức Nam, Trần Quốc Thành, Hoàng Nghĩa Nhạc, Phùng Văn Hào, 2016, Kết quả nghiên cứu về loài sâm Puxailaileng ở vùng núi cao tỉnh Nghệ An, *Tạp chí Khoa học-Công nghệ Nghệ An*, 12, tr. 7-11.
15. Pham TN., Pham TH., Nguyen QN., Phan VT., Nguyen MK., Dinh DL., 2020, A molecular phylogeny of *Panax* L. genus (Araliaceae) based on ITS-rRNA and *matK* support for identification of *Panax* species in Vietnam, *VNU Journal of Science: Medical and Pharmaceutical Sciences*, 36(2), pp. 91-99.
16. Lương Đức Toàn, 2018, Tính chất, chất lượng đặc thù của sâm củ và điều kiện tự nhiên vùng trồng sâm Ngọc Linh tỉnh Kon Tum, Kỷ yếu Hội nghị Đầu tư và phát triển sâm Ngọc Linh Kon Tum và các dược liệu khác, Bộ Y tế, UBND Tỉnh Kon Tum, Bộ Nông nghiệp và Phát triển Nông thôn, Kon Tum.
17. Zhang GH., Ma CH., Zhang JJ., Chen JW., He MH., Xu XZ., Jiang NH., Yang SC., 2015, Transcriptome analysis of *Panax vietnamensis* var. *fuscidicus* discovers putative ocotillol-type ginsenosides biosynthesis genes and genetic markers, *BMC Genomics*, 16, tr. 159.
18. Zhang T., Han M., Yang L., Han Z., Cheng L., Sun Z., Yang L., 2018, The effects of environmental factors on ginsenoside biosynthetic enzyme gene expression and saponin abundance, *Molecules*, 24(1), pp. 14.
19. Gachon CM., Langlois-Meurinne M., Saindrenan P., 2005, Plant secondary metabolism glycosyltransferases: the emerging functional analysis, *Trends Plant Sci*, 10(11), pp. 542-549.
20. Meesapyodsuk D., Balsevich J., Reed DW., Covello PS., 2007, Saponin biosynthesis in *Saponaria vaccaria*. cDNAs encoding beta-amyrin

- synthase and a triterpene carboxylic acid glucosyltransferase, *Plant Physiol*, 143(2), pp. 959-969.
21. Shibuya M., Hoshino M., Katsube Y., Hayashi H., Kushiro T., Ebizuka Y., 2006, Identification of beta-amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay, *FEBS J*, 273(5), pp. 948-959.
 22. Nguyen MD., Kasai R., Ohtani K., Ito A., Nguyen TN., Yamasaki K., Tanaka O., 1993, Saponins from Vietnamese ginseng, *Panax vietnamensis* Ha et Grushv. collected in central Vietnam. I. *Chem Pharm Bull*, 41, pp. 2010-2014.
 23. Nguyen MD., Kasai R., Ohtani K., Ito A., Nguyen TN., Yamasaki K., Tanaka O., 1994b, Saponins from Vietnamese ginseng, *Panax vietnamensis* Ha et Grushv. collected in central Vietnam. III, *Chem Pharm Bull*, 42, pp. 634-640.
 24. Trần Công Luận, 2003, Kết quả nghiên cứu về hóa học sâm Việt Nam 1978 – 2002, *Hội thảo Bảo tồn và Phát triển cây sâm Việt Nam, Tam Kỳ*, tr. 62-75.
 25. Nguyễn Thượng Dong, Trần Công Luận, Nguyễn Thị Thu Hương, 2007, Sâm Việt Nam và một số cây thuốc họ nhân sâm, Nhà xuất bản Khoa học và Kỹ thuật.
 26. Konno C., Sugiyama K., Kano M., Takahashi M., Hikino H., 1984, Isolation and hypoglycaemic activity of panaxans A, B, C, D and E, glycans of *Panax ginseng* Roots, *Planta Med*, 50, pp. 434-436.
 27. Nguyễn Thị Thu Hương, 2003, Kết quả nghiên cứu về tác dụng dược lý sâm Việt Nam, *Hội thảo bảo tồn và phát triển cây sâm Việt Nam, Tam Kỳ*, tr. 76-90.
 28. Nguyễn Thị Thu Hương, Phạm Thị Mỹ Loan, 2007, Tác dụng của sâm Việt Nam và dinh dưỡng trên thực nghiệm gây suy giảm miễn dịch, *Dược Liệu*, 3(4), tr. 121-126.
 29. Sun C., Li Y., Wu Q., Luo H., Sun Y., Song J., Lui EM., Chen S., 2010, *De novo* sequencing and analysis of the American ginseng root transcriptome using a GS FLX Titanium platform to discover putative genes involved in ginsenoside biosynthesis, *BMC Genomics*, 11, pp. 262.

30. Wu D., Austin RS., Zhou S., Brown D., 2013, The root transcriptome for North American ginseng assembled and profiled across seasonal development, *BMC Genomics*, 14, pp. 564.
31. Liu MH., Yang BR., Cheung WF., Yang KY., Zhou HF., Kwok JSL., Liu GC., Li XF., Zhong S., Lee SMY., Tsui SKW., 2015, Transcriptome analysis of leaves, roots and flowers of *Panax notoginseng* identifies genes involved in ginsenoside and alkaloid biosynthesis, *BMC Genomics*, 16, pp. 265.
32. Chen S., Luo H., Li Y., Sun Y., Wu Q., Niu Y., Song J., Lv A., Zhu Y., Sun C., Steinmetz A., Qian Z., 2011, 454 EST analysis detects genes putatively involved in ginsenoside biosynthesis in *Panax ginseng*, *Plant Cell Rep*, 30(9), pp. 1593-1601.
33. Li C., Zhu Y., Guo X., Sun C., Luo H., Song J., Li Y., Wang L., Qian J., Chen S., 2013, Transcriptome analysis reveals ginsenosides biosynthetic genes, microRNAs and simple sequence repeats in *Panax ginseng* C. A. Meyer, *BMC Genomics*, 14(1), pp. 245.
34. Mathiyalagan R., Subramaniyam S., Natarajan S., Kim YY., Sun MS., Kim SY., Kim Y-J., Yang DC., 2013, *In silico* profiling of microRNAs in Korean ginseng (*Panax ginseng* Meyer), *J Ginseng Res*, 37(2), pp. 227-247.
35. Rai A., Yamazaki M., Takahashi H., Nakamura M., Kojoma M., Suzuki H., Saito K., 2016, RNA-seq transcriptome analysis of *Panax japonicus*, and its comparison with other *Panax* species to identify potential genes involved in the saponins biosynthesis, *Front Plant Sci*, 7, pp. 481
36. Nguyen B., Kyunghee K., Young-Chang K., Sang-Choon L., Ji ES., Junki L., Nam-Hoon K., Woojong J., Hong-II C., Tae-Jin Y., 2015, The complete chloroplast genome sequence of *Panax vietnamensis* Ha et Grushv (Araliaceae), *Mitochondrial DNA*.
37. Bustin SA., 2002, Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems, *J Mol Endocrinol*, 29, pp. 23-39.
38. Derveaux S., Vandesompele J., Hellemens J., 2010, How to do successful gene expression analysis using real-time PCR, *Methods*, 50, pp. 227-230.

39. He F., Zhu Y., He M., Zhang Y., 2008, Molecular cloning and characterization of the gene encoding *squalene epoxidase* in *Panax notoginseng*, *DNA Sequence*, 19(3), pp. 270-273.
40. Luo H., Sun C., Sun Y., Wu Q., Li Y., Song J., Niu Y., Cheng X., Xu H., Li C., Liu J., Steinmetz A., Chen S., 2011, Analysis of the transcriptome of *Panax notoginseng* root uncovers putative triterpene saponin-biosynthetic genes and genetic markers, *BMC Genomics*, 5, pp. S5.
41. Park SB., Chun JH., Ban YW., Han JY., Choi YE., 2016, Alteration of *Panax ginseng* saponin composition by overexpression and RNA interference of the protopanaxadiol 6-hydroxylase gene (*CYP716A53v2*), *J Ginseng Res*, 40(1), pp. 47-54.
42. Chomczynski P., Sacchi N., 2006, The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on, *Nat Protoc*, 1(2), pp. 581-585.
43. Livak KJ., Schmittgen TD., 2001, Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method, *Methods*, 25(4), pp. 402-408.
44. Han JY., Hwang HS., Choi SW., Kim HJ., Choi YE., 2012, Cytochrome *P450 CYP716A53v2* catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in *Panax ginseng*, *Plant Cell Physiol*, 53(9), pp. 1535-1545.
45. Jung SC., Kim W., Park SC., Jeong J., Park MK., Lim S., Kim SC., 2014, Two ginseng UDP-glycosyltransferases synthesize ginsenoside Rg3 and Rd, *Plant Cell Physiol*, 55(12), pp. 2177-2188.
46. Lee S., Rhee DK., 2017, Effects of ginseng on stress-related depression, anxiety, and the hypothalamic-pituitary-adrenal axis, *J Ginseng Res*, 41(4), pp. 589-594.
47. Jayakodi M., Lee SC., Lee YS., Park HS., Kim NH., Jang W., Yang TJ., 2015, Comprehensive analysis of *Panax ginseng* root transcriptomes, *BMC Plant Biol*, 15(1), pp. 1-12.
48. Wang J., Li J., Li J., Liu S., Wu X., Li J., Gao W., 2016, Transcriptome profiling shows gene regulation patterns in ginsenoside pathway in response to methyl jasmonate in *Panax Quinquefolium* adventitious root, *Sci Rep-UK*, 6(1), pp. 1-11.

49. Li J., Ma L., Zhang S., Zuo C., Song N., Zhu S., Wu J., 2019, Transcriptome analysis of 1- and 3-year-old *Panax notoginseng* roots and functional characterization of saponin biosynthetic genes *DS* and *CYP716A47-like*, *Planta*, 249(4), pp. 1229-1237.
50. Wu Z., Wouters J., Poulter CD., 2005, Isopentenyl diphosphate isomerase. Mechanism-based inhibition by diene analogues of isopentenyl diphosphate and dimethylallyl diphosphate, *J Am Chem Soc*, 127(49), pp. 17433-17438.
51. Laupitz R., Hecht S., Amslinger S., Zepeck F., Kaiser J., Richter G., Rohdich F., 2004, Biochemical characterization of *Bacillus subtilis* type II isopentenyl diphosphate isomerase, and phylogenetic distribution of isoprenoid biosynthesis pathways, *Eur J Biochem*, 271(13), pp. 2658-2669.
52. Lee MH., Jeong JH., Seo JW., Shin CG., Kim YS., In JG., Choi YE., 2004, Enhanced triterpene and phytosterol biosynthesis in *Panax ginseng* over expressing squalene synthase gene, *Plant Cell Physiol*, 45(8), pp. 976-984.
53. Han JY., In JG., Kwon YS., Choi YE., 2010, Regulation of ginsenoside and phytosterol biosynthesis by RNA interferences of *squalene epoxidase* gene in *Panax ginseng*, *Phytochemistry*, 71(1), pp. 36-46.

PHỤ LỤC

Phụ lục 3.1 Ma trận so sánh cDNA của một số gen liên quan đến con đường sinh tổng hợp ginsenoside phân lập từ sâm Ngọc Linh với trình tự tham chiếu và trình tự trích xuất từ dữ liệu transcriptome

A. *PvAACT*; B. *PvIDI*; C. *PvSS*; D. *PvSE*;

A

```

510      520      530      540      550      560      570      580      590      600
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
P.vietnamensis_AACT
KJ804173.1 Panax notoginseng
c801217_g7_i1
c790622_g5_i1
c790622_g1_i3
          AGATCATCAT---GGCTCCT--GCAGCAGC
TCTCTCTCTCTCCCTTACACTTCTCCAGCTT-CCGGAACCCATAACGTGAGATAAGCCGTTTCTGT.....|
TCTCTCTCTCTCCCTTACACTTCTCCAGCTT-CCGGAGCCCTAAACGTGAGATAAGCCGTTTCTGT.....|
AATTTCCCCCTCTCTGCTCACGCAAGCAAAATAGTTGAAAGTTTGAATCCCTTCTATTAAATCCCA...TG.T.TAGATAAT..ATG..TT.G.A
TTTAGTTGTGATACAAGTCTGAAGATTGAGATATACTCTCTTTGAAACATAATGAAAACCTTAATCAA.AGAT..TCCCCTCAA..TT.-----TGTG

610      620      630      640      650      660      670      680      690      700
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
P.vietnamensis_AACT
KJ804173.1 Panax notoginseng
c801217_g7_i1
c790622_g5_i1
c790622_g1_i3
          AACAGATTGTTTCCAGATCCATCAAGGCTAGAGATGTTTGTATTGTTGGTGCCGCACGTACAC---CAATGGGTGGCT-TTCTTGGAACACTTTCATCATT
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
          .GTG.GA.T..G..T..A..T.G.CT.C.....C..A.....TT.....A.....-G.....AT..C.....TT.T...G..G..
G..TC..ATC.TG.TCA.A.A..GT.....GAG.AC..-C.A..GCATAT..A..T.TATAT..AA...ATGA.C.A.A.T.TGTA.GG...C..

710      720      730      740      750      760      770      780      790      800
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
P.vietnamensis_AACT
KJ804173.1 Panax notoginseng
c801217_g7_i1
c790622_g5_i1
c790622_g1_i3
          ATCAGCTACCAAAC TTGGAT-CCATAGCAATTCAAAGCGCTCT-----TGAAAGGGCAAATGTTGA-TCCAGCACTAGTACAAGAGGTTTTCTTTGG
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
          G..T.....GG.....C..T.....T..G..T.....A.....A.....-GT...T..G.....C..T.....
..A..GC.TAC.C.A..T..A.A..T.A..A.A.G.AG.AA.AACTACGA.T.G.AACTG..AA...CG..TGCT...T..GT..T..ACGGTAAAAA

810      820      830      840      850      860      870      880      890      900
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
P.vietnamensis_AACT
KJ804173.1 Panax notoginseng
c801217_g7_i1
c790622_g5_i1
c790622_g1_i3
          AAA---TGTTCTCAGTGCAAATTTGGGTCAGGCTCCGCTAGACAAGCAGCATTAGGTGCAGGAATACCTAACGCGGTAGTCTGTACCACCATTAAACA
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
          .....T.....|
          .....T.A.....A.A.....A.....C.....T.T.....TA.A..G..T..C.....
...CACA.A.CA..C.CTACT.A.A.AAG...A.A.TAATC...AT.AAG.T..G.C.TG.TC...AA.GTTTA..GTT.A.TA...G.TTT..

910      920      930      940      950      960      970      980      990

```

1000

```

P.vietnamensis_AACT
KJ804173.1 Panax notoginseng
c801217_g7_i1
c790622_g5_i1
c790622_g1_i3
    
```

B

```

    410      420      430      440      450      460      470      480      490      500
    |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
PvIDI
KJ804168.1 Panax notoginseng i
c800937_g5_i1
c800937_g4_i1
    
```

```

    510      520      530      540      550      560      570      580      590
    |-----|-----|-----|-----|-----|-----|-----|-----|-----|
PvIDI
KJ804168.1 Panax notoginseng i
c800937_g5_i1
c800937_g4_i1
    
```

```

    610      620      630      640      650      660      670      680      690
    |-----|-----|-----|-----|-----|-----|-----|-----|-----|
PvIDI
KJ804168.1 Panax notoginseng i
c800937_g5_i1
c800937_g4_i1
    
```

C

```

    510      520      530      540      550      560      570      580      590      600
    |-----|-----|-----|-----|-----|-----|-----|-----|-----|
P.vietnamensis_SS
GQ468527.2 Panax ginseng
GU183406.1 Panax ginseng
AM182457.1 Panax quinquefolius
AM182456.1 Panax quinquefolius
KP689321.1 Panax ginseng
KP689320.1 Panax ginseng
    
```

```

KJ804172.1 Panax notoginseng .....
KP689323.1 Panax ginseng .....G.....T.....
KP689322.1 Panax ginseng .....G.....T.....
KP689319.1 Panax ginseng .....G.....T.....
KP689318.1 Panax ginseng .....G.....T.....
GU997681.1 Panax quinquefolius .....T..C.....T..C.....
KT123897.1 Panax notoginseng .....T..C.....T..C.....
KP890782.1 Panax japonicus .....T..C.....T..C.....
KC953032.1 Panax notoginseng .....T..C.....T..C.....
KT936528.1 Panax sokpayensis .....T..C.....T..C.....
DQ186630.1 Panax notoginseng .....T..C.....T..C.....
AB115496.1 Panax ginseng .....T..C.....T..C.....
AB010148.1 Panax ginseng .....T..C.....T..C.....
KP689317.1 Panax ginseng .....T..C.....T..C.....
KP689316.1 Panax ginseng .....T..C.....T..C.....
KP689315.1 Panax ginseng .....T..C.....T..C.....
KP689314.1 Panax ginseng .....T..C.....T..C.....
EU131090.1 Panax notoginseng .....
KJ939264.1 Panax ginseng .....T..C.....T..C.....
KC524471.1 Panax quinquefolius A.CCAGGA.G.T.T..A.A.T..G.TGAAT.AGGC.TT.AGG..TGTGT..AT..AA..G.T.C..A.CG.G.TT..C.GA.A.GCC.TTT...GG.
KC524469.1 Panax quinquefolius A.CCAGGA.G.T.T..A.A.T..G.TGAAT.AGGC.TT.AGG..TGTGT..AT..AA..G.T.C..A.CG.G.TT..C.GA.A.GCC.TTT...GG.
KC422650.1 Panax notoginseng .....T..C.....T..C.....
EU502717.1 Panax ginseng .....T..C.....T..C.....
c802035_g3_i1 .....
c802035_g5_i1 .....
c669272_g1_i1 .....
c800646_g6_i2 .TCAA..AA.ATTT.GT.G.T.A.AA.ACT.ATT.AT.TTG...G...TT..T.AA.T....GT...GA.C.CT..AT.C...TC.TTCTGCTAT

```

610

620

630

640

650

660

670

680

690

700

```

P.vietnamensis_SS
GQ468527.2 Panax ginseng .....A.....T.....
GU183406.1 Panax ginseng .....A.....
AM182457.1 Panax quinquefolius .....A.....T.....G.....
AM182456.1 Panax quinquefolius .....A.....T.....
KP689321.1 Panax ginseng .....A.....T.....
KP689320.1 Panax ginseng .....A.....T.....
KJ804172.1 Panax notoginseng .....T.....
KP689323.1 Panax ginseng .....A.....
KP689322.1 Panax ginseng .....A.....
KP689319.1 Panax ginseng .....A.....
KP689318.1 Panax ginseng .....A.....
GU997681.1 Panax quinquefolius .....G.....C.....T.....T.....T.....G.....

```

```

KT123897.1 Panax notoginseng .....G.....C.....T.....T.....G.....A....
KP890782.1 Panax japonicus .....G.....C.....T.....T.....G.....A....
KC953032.1 Panax notoginseng .....G.....C.....T.....T.....G.....A....
KT936528.1 Panax sokpayensis .....G.....C.....T.....T.....T.G.....A....
DQ186630.1 Panax notoginseng .....G.....C.....T.....T.....G.....A....
AB115496.1 Panax ginseng .....G.....C.....T.....T.....T.G.....A....
AB010148.1 Panax ginseng .....G.....C.....T.....T.....T.G.....A....
KP689317.1 Panax ginseng .....G.....C.....A.....T.....T.....G.....A....
KP689316.1 Panax ginseng .....G.....C.....T.....T.....T.G.....A....
KP689315.1 Panax ginseng .....G.....C.....A.....T.....T.....G.....A....
KP689314.1 Panax ginseng .....G.....C.....T.....T.....T.G.....A....
EU131090.1 Panax notoginseng .....T.....G.....C.....T.....T.....T.G.....A....
KJ939264.1 Panax ginseng .....G.....C.....T.....T.....T.G.....A....
KC524471.1 Panax quinquefolius ..G...AA.TACTAGGC...T.A.CCCCT.G.A..T.C.----TGCA...TAGCAGGAAGAAGC.T.C..AA.GGC.GTTTCA..CA.C.GATG
KC524469.1 Panax quinquefolius ..G...AA.TACTAGGC...T.A.CCCCT.G.A..T.C.----TGCA...TAGCAGGAAGAAGC.T.C..AA.GGC.GTTTCA..CA.C.GATG
KC422650.1 Panax notoginseng .....G.....C.....T.....T.....G.....A....
EU502717.1 Panax ginseng .....G.....C.....T.....T.....T.G.....A....
c802035_g3_i1 .....G.....C.....T.....T.....T.G.....A....
c802035_g5_i1 .....G.....AGTATTTT.C.TTCT.CT.C..G.T.GTTT..TAACTGT...TA...G.C.ATTGA...ATTCC.C.A.TC-TACA
c669272_g1_i1 .....G.....AGTATTTT.C.TTCT.CT.C..G.T.GTTT..TAACTGT...TA...G.C.ATTGA...ATTCC.C.A.TC-TACA
c800646_g6_i2 .....G...TTTG...T.TGG.----.AA.AT.GCTTT.T..TGT.C.TCTG.ACT..T.CCTTT...TT.CA..TTCTTG...TACTT.TCAAA.TAAA..

```

D

```

                1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
PvSE2
AGATCCTTCTCCGATCTTGTTTTAT-AAAATTAGTAGTACCGAGATTCGCTGTTTGGTTGATGTGCCCTGGAC-AAAAGGTGCCTTCTATTTCTAATGGGG
DQ457054.1 Panax notoginseng .....T.....T.....G.....A....
KT123898.1 Panax notoginseng .....T.....T.....G.....A....
KT936529.1 Panax sokpayensis .....A.....C..C.....A.....A.....GA.....G.....T....
JX625132.1 Panax notoginseng .....A.....C.....A.....A.....GA.A.....G.....A....
DQ386734.1 Panax notoginseng .....A.....C..C.....A.....A.....T.....G.....A....
FJ393274.2 Panax ginseng .....A.....C..C.....A.....A.....A.....G.....T....
AB122078.1 Panax ginseng .....A.....G.....G.....G.....G.....G.....G.....G.....
AB003516.1 Panax ginseng .....A.....G.....G.....G.....G.....G.....G.....G.....
KJ939265.1 Panax ginseng .....G.....G.....G.....G.....G.....G.....G.....G.....
KJ804171.1 Panax notoginseng .....T.....T.....T.....T.....T.....T.....T.....T.....
KJ946469.1 Panax vietnamensis .....G.....G.....G.....G.....G.....G.....G.....G.....
KJ946468.1 Panax vietnamensis .....A.....C..C.....A.....A.....A.....GA.....G.....T....
KJ946467.1 Panax vietnamensis ..C..A..A..C.....CCC..C.....C..A.....C.....T.....T.....C.....A.....C....
KC953033.1 Panax notoginseng .....T.....T.....T.....T.....T.....T.....T.....T.....
KC422651.1 Panax notoginseng .....T.....T.....T.....T.....T.....T.....T.....T.....
MK603118.1 Panax japonicus .....A.....G.....G.....G.....G.....G.....G.....G.....
c780573_g1_i4 .....A.....G.....G.....G.....G.....G.....G.....G.....

```

```

c782799_g1_i1      . . . C . A . A . C . . . . . - CCC . C . . . . C . A . . . . C . . . . . T . . . . T . - . . . . C . . . . A . . . . C .
c739748_g1_i1      . CCGG . GA . A . . AGCTCCAACGATGATG . CG . C . GC . . CTGATTGA . TAGCCCG . . . G . TG . CGC . A . . TTC . G . TC . . TT . GGG . C . CAAA . GAA . TA
c746527_g1_i1      . . . . . A . C . . TC . A . . . . - CC . . C . . C . . T . . G . A . T . . . . . A . . . . T . - . . . . C . T . . . . C . G . . . . T .
c763721_g3_i1      ----- . C . C . G . C . - . . A . T . GG . - . . AC . . ATCAA . AACAA . G . . TGA - . . TA . . . . TCC . - . . T . . AAATA
c779029_g1_i1      TACAT . A . . G . . A . . TCCAAGAA . A - . G . . CC . AA . - . . A . . A . GCGA . . C . . TACTTCT . TTGA . AC . CG . . . C . AT . TGAACCA . GGCG . CAACC

```