

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC SƯ PHẠM

Hoàng Việt Cường

NGHIÊN CỨU ĐẶC ĐIỂM HỆ GENE LỤC LẠP VÀ PHÂN TÍCH
HOẠT TÍNH CỦA SAPONIN PHÂN LẬP TỪ CÂY LAN TAI CÁO
(*Hoya verticillata* var. *verticillata*)

LUẬN ÁN TIẾN SĨ SINH HỌC

THÁI NGUYÊN - 2026

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC SƯ PHẠM

Hoàng Việt Cường

NGHIÊN CỨU ĐẶC ĐIỂM HỆ GENE LỤC LẠP VÀ PHÂN TÍCH
HOẠT TÍNH CỦA SAPONIN PHÂN LẬP TỪ CÂY LAN TAI CÁO
(*Hoya verticillata* var. *verticillata*)

Ngành: Di truyền học
Mã số: 9420121

LUẬN ÁN TIẾN SĨ SINH HỌC

CÁN BỘ HƯỚNG DẪN KHOA HỌC:

1. GS.TS. Chu Hoàng Mậu
2. PGS.TS. Từ Quang Tân

THÁI NGUYÊN - 2026

LỜI CAM ĐOAN

Tôi xin cam đoan đây là công trình nghiên cứu của tôi dưới sự hướng dẫn của GS.TS. Chu Hoàng Mậu và PGS.TS. Từ Quang Tân. Các kết quả trình bày trong luận án là trung thực, một phần đã được công bố trong các Tạp chí khoa học - công nghệ, phần còn lại chưa được ai công bố trong bất kỳ công trình nào khác. Mọi trích dẫn đều ghi rõ nguồn gốc.

Thái Nguyên, tháng 5 năm 2026

TÁC GIẢ

Hoàng Việt Cường

LỜI CẢM ƠN

Tôi xin bày tỏ lòng biết ơn sâu sắc đến GS.TS. Chu Hoàng Mậu, PGS.TS. Từ Quang Tân đã tận tình hướng dẫn và chỉ bảo trong suốt quá trình tôi nghiên cứu và thực hiện đề tài luận án.

Tôi xin chân thành cảm ơn PGS.TS. Nguyễn Thị Ngọc Lan, PGS.TS. Nguyễn Đức Hùng, PGS.TS. Nguyễn Hữu Quân, PGS.TS. Vũ Thị Thu Thủy, PGS.TS. Sỹ Danh Thường, PGS.TS. Nguyễn Thị Thu Nga, PGS.TS. Phạm Thị Thanh Nhân – Khoa Sinh học Trường Đại học Sư phạm Thái Nguyên đã tận tình giúp đỡ và tạo điều kiện thuận lợi trong quá trình thực hiện đề tài luận án.

Tôi xin chân thành cảm ơn sự giúp đỡ và sự đóng góp những ý kiến quý báu của các thầy cô giáo và cán bộ Bộ môn Di truyền học và Công nghệ Sinh học trong quá trình học tập, nghiên cứu, seminar và hoàn thành luận án. Xin cảm ơn các thầy cô giáo và cán bộ Khoa Sinh học, Bộ phận đào tạo Sau đại học, Trường Đại học Sư phạm - Đại học Thái Nguyên đã giúp đỡ và tạo điều kiện thuận lợi cho tôi hoàn thành khoá học này.

Tôi xin cảm ơn sự hỗ trợ kinh phí của Bộ Giáo dục và Đào tạo trong đề tài mang mã số B2022-TNA-22-CT562 do PGS.TS. Từ Quang Tân làm chủ nhiệm, và Chương trình sáng kiến KRIBB [KGM4582423] của Hàn Quốc.

Tôi xin trân trọng cảm ơn Ban Giám đốc Sở Giáo dục và Đào tạo tỉnh Thái Nguyên, chân thành cảm ơn các đồng nghiệp Phòng Giáo dục trung học đã tạo điều kiện thuận lợi, giúp đỡ tôi trong quá trình học tập và công tác.

Tôi rất biết ơn những người thân trong gia đình, các bạn bè đã luôn động viên, quan tâm, tạo mọi điều kiện cho tôi học tập và nghiên cứu.

Thái Nguyên, tháng 5 năm 2026

TÁC GIẢ

Hoàng Việt Cường

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
MỤC LỤC	iii
Danh mục bảng	vii
Danh mục hình	viii
Danh mục từ và chữ viết tắt	x
MỞ ĐẦU	1
1. Đặt vấn đề	1
2. Mục tiêu nghiên cứu	2
3. Nội dung nghiên cứu	2
4. Những đóng góp mới của luận án	3
5. Ý nghĩa khoa học và thực tiễn của đề tài luận án	3
Chương 1. TỔNG QUAN TÀI LIỆU	5
1.1. ĐẶC ĐIỂM HỆ GENE LỤC LẠP CỦA THỰC VẬT VÀ CHI <i>HOYA</i>	5
1.1.1. Hệ gene lục lập ở thực vật bậc cao	5
1.1.2. Hệ gene lục lập của các loài thuộc chi <i>Hoya</i>	7
1.1.2.1. Chi <i>Hoya</i> và loài <i>Hoya verticillata</i> var. <i>verticillata</i>	7
1.1.2.2. Đặc điểm hệ gene lục lập của các loài thuộc chi <i>Hoya</i>	9
1.1.3. Nhận xét chung	12
1.2. PHÂN TÍCH PHÁT SINH CHỦNG LOẠI DỰA TRÊN TRÌNH TỰ DNA LỤC LẠP	13
1.2.1. Phát sinh chủng loại học phân tử	13
1.2.2. Phân tích sự phát sinh chủng loại ở thực vật	14
1.2.3. Phân tích sự phát sinh chủng loại chi <i>Hoya</i>	18
1.2.4. Nhận xét chung	21
1.3. THÀNH PHẦN HÓA HỌC, HOẠT TÍNH SINH HỌC CỦA MỘT SỐ LOÀI THUỘC CHI <i>HOYA</i>	22
1.3.1. Các hợp chất hóa học ở thực vật hạt kín và chi <i>Hoya</i>	22

1.3.2. Hoạt tính sinh học của thực vật và một số loài <i>Hoya</i>	25
1.3.3. Saponin	27
1.3.3.1. <i>Saponin thực vật</i>	28
1.3.3.2. <i>Hoạt tính sinh học của saponin thực vật</i>	30
1.3.3.3. <i>Phân lập saponin từ một số loài Hoya</i>	34
1.3.4. Mô phỏng tương tác phân tử và động học phân tử	37
1.3.5. Nhận xét chung	40
Chương 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU	42
2.1. VẬT LIỆU NGHIÊN CỨU	42
2.1.1. Vật liệu thực vật	42
2.1.2. Các chủng vi khuẩn kiểm định	43
2.1.3. Các dòng tế bào ung thư	43
2.1.4. Dữ liệu	43
2.2. HÓA CHẤT, THIẾT BỊ	43
2.2.1. Hóa chất	43
2.2.2. Thiết bị	44
2.3. PHƯƠNG PHÁP NGHIÊN CỨU	44
2.3.1. Nhóm phương pháp nghiên cứu hệ gene lục lạp	44
2.3.1.1. <i>Tách chiết DNA và giải trình tự hệ gene</i>	44
2.3.1.2. <i>Lắp ráp và chú thích hệ gene</i>	45
2.3.1.3. <i>Phân tích hệ gene</i>	45
2.3.2. Nhóm phương pháp phân tích phát sinh chủng loại	46
2.3.2.1. <i>Phương pháp phân tích phát sinh chủng loại trình tự hệ gene lục lạp</i>	46
2.3.2.2. <i>Phương pháp phân tích phát sinh chủng loại dựa trên trình tự một số vùng DNA lục lạp</i>	46
2.3.3. Nhóm phương pháp phân tích hoạt tính sinh học, phân lập, xác định chất	46
2.3.3.1. <i>Định tính các hợp chất hữu cơ có trong Lan tai cáo</i>	47
2.3.3.2. <i>Phương pháp tạo cao chiết</i>	48
2.3.3.3. <i>Phương pháp xác định hoạt tính sinh học</i>	48

2.3.3.4. Phân lập saponin từ lá Lan tai cáo	51
2.3.3.5. Phân tích tương tác phân tử và động học phân tử <i>in silico</i>	55
2.3.4. Phương pháp phân tích số liệu	57
2.4. ĐỊA ĐIỂM NGHIÊN CỨU VÀ HOÀN THÀNH LUẬN ÁN	57
Chương 3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN	58
3.1. ĐẶC ĐIỂM HỆ GENE LỤC LẠP CỦA LOÀI LAN TAI CÁO	58
3.1.1. Cấu trúc hệ gene	58
3.1.2. So sánh hệ gene lục lập hoàn chỉnh của <i>Hoya verticillata</i> var. <i>verticillata</i> thu thập ở Việt Nam với các loài thuộc chi <i>Hoya</i> khác trên GenBank	65
3.2. PHÂN TÍCH PHÁT SINH CHỦNG LOẠI DỰA TRÊN TRÌNH TỰ HỆ GENE LỤC LẠP VÀ MỘT SỐ VÙNG DNA THUỘC HỆ GENE LỤC LẠP	70
3.2.1. Phân tích phát sinh chủng loại dựa trên hệ gene lục lập hoàn chỉnh	70
3.2.2. Phân tích phát sinh chủng loại dựa trên một số vùng DNA của hệ gene lục lập	73
3.2.2.1. Phát sinh chủng loại dựa trên một số vùng gene	74
3.2.2.2. Phát sinh chủng loại dựa trên các vùng đệm	74
3.3. PHÂN TÍCH HOẠT TÍNH SINH HỌC CỦA SAPONIN PHÂN LẬP TỪ LAN TAI CÁO	83
3.3.1. Thành phần hóa học của Lan tai cáo	83
3.3.2. Hoạt tính kháng khuẩn	83
3.3.3. Hoạt tính sinh học của các phân đoạn chứa saponin từ cao chiết Lan tai cáo	84
3.3.3.1. Hoạt tính chống oxy hóa	84
3.3.3.2. Hoạt tính gây độc tế bào ung thư	85
3.3.3.3. Hoạt tính ức chế α -glucosidase	87
3.3.4. Phân lập saponin	88
3.3.4.1. Chiết xuất và phân lập saponin	88
3.3.4.2. Thủy phân acid và sắc ký khí	88
3.3.4.3. Xác định cấu trúc saponin	90
3.3.5. Nghiên cứu tương tác phân tử và động học phân tử <i>in silico</i> của saponin	93

<i>3.3.5.1. Nghiên cứu tương tác phân tử in silico</i>	93
<i>3.3.5.2. Nghiên cứu động học phân tử in silico</i>	100
KẾT LUẬN VÀ ĐỀ NGHỊ	106
1. Kết luận	106
2. Đề nghị	106
CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ LIÊN QUAN ĐẾN LUẬN ÁN	108
TÀI LIỆU THAM KHẢO	109

DANH MỤC BẢNG

	<i>Trang</i>
Bảng 1.1. Thành phần hóa học của các loài trong chi <i>Hoya</i>	24
Bảng 3.1. Tóm tắt hệ gene lục lạp của loài <i>H. verticillata</i> var. <i>verticillata</i>	60
Bảng 3.2. Thành phần gene của hệ gene lục lạp <i>H. verticillata</i> var. <i>verticillata</i>	60
Bảng 3.3. Tần số sử dụng codon đồng nghĩa tương đối (RSCU) của các gene mã hóa protein ở <i>Hoya verticillata</i> var. <i>verticillata</i> .	63
Bảng 3.4. Tỷ lệ phân bố các vùng trong hệ gene lục lạp của 6 loài <i>Hoya</i>	67
Bảng 3.5. Kết quả định tính các nhóm hợp chất hữu cơ có trong Lan tai cáo	83
Bảng 3.6. Hoạt tính kháng khuẩn của cao chiết tổng số Lan tai cáo	84
Bảng 3.7. Hoạt tính trung hòa gốc tự do (DPPH) của các phân đoạn chứa saponin	85
Bảng 3.8. Khả năng gây độc tế bào ung thư của các phân đoạn chứa saponin	86
Bảng 3.9. Tác động ức chế α -glucosidase của các phân đoạn chứa saponin	87
Bảng 3.10. Dữ liệu phổ ^{13}C và ^1H NMR của hợp chất 1 và 2 (dung môi pyridine-d ₅ , δ tính bằng ppm, J tính bằng Hz)	89
Bảng 3.11. Tương tác phân tử giữa saponin với protein 6GL8.	94

DANH MỤC HÌNH

	<i>Trang</i>
Hình 1.1. Hình ảnh loài Lan tai cáo (<i>H. verticillata</i> var. <i>verticillata</i>)	8
Hình 1.2. Sơ đồ phát sinh chủng loại có độ chính xác tối đa (ML) được suy ra từ 27 hệ gene lục lạp hoàn chỉnh của loài <i>Aster</i>	18
Hình 1.3. Cấu trúc của saponin	28
Hình 1.4. Sinh tổng hợp saponin thực vật	29
Hình 2.1. Mẫu cây Lan tai cáo (<i>Hoya verticillata</i> var. <i>verticillata</i>)	42
Hình 2.2. Sơ đồ thí nghiệm phân lập và xác định saponin từ Lan tai cáo	47
Hình 2.3. Sơ đồ phân lập saponin từ Lan tai cáo	52
Hình 3.1. Bản đồ hệ gene lục lạp của loài Lan tai cáo thu hái tại Việt Nam	59
Hình 3.2. Phân tích các đoạn lặp lại trình tự đơn giản (SSR) của hệ gene lục lạp hoàn chỉnh của <i>H. verticillata</i> var. <i>verticillata</i>	64
Hình 3.3. Phân tích trình tự lặp lại lớn (LSR) trên quy mô hệ gene ở <i>H. verticillata</i> var. <i>verticillata</i> .	65
Hình 3.4. So sánh cấu trúc hệ gene lục lạp hoàn chỉnh của 6 loài <i>Hoya</i>	66
Hình 3.5. Đồ thị so sánh các giá trị đa dạng nucleotide (Pi) thông qua phân tích cửa sổ trượt (chiều dài cửa sổ: 600 bp; kích thước bước: 200 bp) trên toàn bộ hệ gene lục lạp của 30 loài <i>Hoya</i> .	68
Hình 3.6. So sánh điểm nối các vùng LSC, SSC và hai vùng IR trong số 6 hệ gene lục lạp hoàn chỉnh của <i>Hoya</i> .	68
Hình 3.7. Sơ đồ phát sinh chủng loại của 50 loài trong họ Apocynaceae dựa trên 79 trình tự gene mã hóa protein.	72
Hình 3.8. Mối quan hệ phát sinh loài được suy ra bằng phương pháp Maximum Likelihood dựa trên trình tự vùng gene <i>ndhF</i> .	75
Hình 3.9. Cây phát sinh chủng loại sử dụng vùng đệm <i>trnK-rps16</i> của hệ gene lục lạp <i>H. verticillata</i> var. <i>verticillata</i> và của các loài khác thuộc họ Apocynaceae	76

Hình 3.10. Cây phát sinh chủng loại sử dụng vùng đệm <i>rps16-trnQ</i> của hệ gene lục lạp <i>H. verticillata</i> var. <i>verticillata</i> và của các loài khác thuộc họ Apocynaceae	78
Hình 3.11. Cây phát sinh chủng loại sử dụng vùng đệm <i>ndhC-trnV</i> của hệ gene lục lạp <i>H. verticillata</i> var. <i>verticillata</i> và của các loài khác thuộc họ Apocynaceae	80
Hình 3.12. Cây phát sinh chủng loại sử dụng vùng đệm <i>psbI-atpA</i> của hệ gene lục lạp <i>H. verticillata</i> var. <i>verticillata</i> và của các loài khác thuộc họ Apocynaceae	81
Hình 3.13. Cây phát sinh chủng loại sử dụng vùng đệm <i>rbcL-accD</i> của hệ gene lục lạp <i>H. verticillata</i> var. <i>verticillata</i> và của các loài khác thuộc họ Apocynaceae	82
Hình 3.14. Cấu trúc liên kết (HMBC) và cấu hình không gian (ROESY) của hợp chất 1	92
Hình 3.15. Cấu trúc liên kết (HMBC) và cấu hình không gian (ROESY) của hợp chất 2	93
Hình 3.16. Cấu trúc 3D của các phân tử thử nghiệm.	95
Hình 3.17. Mô hình tương tác phân tử 2D 3D của Hợp chất 1 (A-1, A-2), Hợp chất 2 (B-1, B-2), Hợp chất 3 (C-1, C-2), Hợp chất 4 (D-1, D-2), Hợp chất 5 (E-1, E-2) và navitoclax (F-1, F-2) với thụ thể 6GL8.	99
Hình 3.18. Biểu đồ so sánh độ ổn định cấu trúc (RMSD) của phức hợp Hợp chất 2 (Parasiticoside B) và Navitoclax trong 10 ns mô phỏng động học phân tử.	101
Hình 3.19. Biểu đồ so sánh dao động cấu trúc (RMSF) của phức hợp Hợp chất 2 (Parasiticoside B) và Navitoclax trong mô phỏng động học phân tử.	102
Hình 3.20. Biểu đồ so sánh sự thay đổi số liên kết hydrogen theo thời gian giữa Hợp chất 2 (Parasiticoside B) và Navitoclax trong mô phỏng động học phân tử.	103

DANH MỤC TỪ VÀ CHỮ VIẾT TẮT

Viết tắt	Tiếng Anh	Nghĩa tiếng Việt
ABTS ⁺	2,2'-azinobis (3-ethylbenzothiazoline-6-sulfonate)	
ACTs	Artemisinin-based combination therapies	Liệu pháp kết hợp dựa trên artemisinin
ADHD	Attention deficit hyperactivity disorder	Rối loạn tăng động giảm chú ý
ATR	Attenuated total reflectance	Tổng phản xạ suy yếu
BEAST2	Bayesian evolutionary analysis 2	Phân tích tiến hóa Bayesian 2
BLAST	Basic Local Alignment Search Tool	Công cụ tìm kiếm trình tự cơ bản
bp	base pairs	Cặp bazơ
CAS	Chemical Abstracts Service	Dịch vụ tóm tắt hóa học
COSY	Correlation Spectroscopy	Phổ tương tác đồng hạt nhân H-H theo mạch liên kết
cp genome	Chloroplast genome	Hệ gene lục lạp
cs		Cộng sự
CTAB	Cetyl trimethylammonium bromide	
CVD	Cardiovascular Disease	Bệnh tim mạch
DMSO	Dimethyl sulfoxide	
DNA	Deoxyribonucleic acid	
DPPH	1,1-diphenyl-2-picrylhydrazyl	
EC ₅₀	50% effective concentration	Nồng độ hiệu quả 50% đối tượng thí nghiệm
ESI	Electrospray ionization	Ion hóa phun điện tử
FBS	Fetal Bovine Serum	Huyết thanh thai bò
GABA	Gamma-aminobutyric Acid	
GATA-3	GATA Binding Protein 3	
GC	Gas chromatography	Sắc ký khí

HMBC	Heteronuclear Multiple Bond Correlation	Tương tác dị hạt nhân qua nhiều liên kết
HPLC	High-performance liquid chromatography	Sắc ký lỏng hiệu năng cao
HPTLC	High-performance thin-layer chromatography	Sắc ký lớp mỏng hiệu suất cao
HR	High Resolution	Độ phân giải cao
HS-SPME	Headspace solid phase micro extraction	Vi chiết pha rắn
HSQC	Heteronuclear single quantum coherence	Tương tác dị hạt nhân qua 1 liên kết
IC ₅₀	50% inhibitory concentration	Nồng độ gây ức chế 50% đối tượng thí nghiệm
IR	Inverted repeat	Vùng lặp lại đảo ngược
ITS	Internal transcribed spacer	
LC	Liquid-chromatography	Sắc ký lỏng
LDL	Low - Density Lipoprotein	Mật độ lipoprotein thấp
LDR	Long-distributed repeat	Lặp lại phân bố dài
LSC	Large single-copy	Sao chép đơn lớn
LSR	Long sequence repeat	Lặp lại chuỗi dài
MALDI	Matrix-assisted laser desorption ionization	Ma trận hấp thụ năng lượng
MD	Molecular dynamics	Động học phân tử
MEME	Minimum Essential Medium Eagle	Môi trường nuôi cấy cơ bản của Eagle
MEP	Methylerythritol Phosphate	
ML	Maximum likelihood	Khả năng tối đa
MPLC	Medium-pressure liquid chromatography	Sắc ký lỏng áp suất trung bình
MS	Mass spectrometry	Khối phổ
MVA	Mevalonic Acid	
Mya	Million years ago	Triệu năm trước
NCBI	National Center for Biotechnology Information	Trung tâm Thông tin Công nghệ sinh học Quốc gia

NCI	National Cancer Institute	Viện Ung thư Quốc gia
NMDA	N-metyl-d-aspartate	
NMR	Nuclear magnetic resonance	Cộng hưởng từ hạt nhân
OD	Optical Density	Mật độ quang học
ov	Overlapped proton signals	Tín hiệu proton chồng chéo
PCR	Polymerase chain reaction	Phản ứng chuỗi polymerase
RAxML	Randomized Axelerated Maximum Likelihood	Phân tích ngẫu nhiên, xác suất tối đa
ROESY	Rotating-Frame Overhauser Enhancement Spectroscopy	Phổ tăng cường Overhauser khung quay
RMSD	Root mean square deviation	Độ lệch bình phương trung bình căn bậc hai
RMSF	Root Mean Square Fluctuations	Biến động bình phương trung bình căn bậc hai
RSCU	Relative synonymous codon usage	Sử dụng codon đồng nghĩa tương đối
RNA	Ribonucleic acid	
SRB	Sulforhodamine B	
SPE	Solid-phase extraction	Trích xuất phá rắn
SSC	Small single copy	Sao chép đơn nhỏ
SSR	Simple sequence repeats	Trình tự lặp lại đơn giản
TCA	Trichloroacetic	
TMS	Tetramethylsilane	
TOCSY	Total Correlation Spectroscopy	Phổ tương quan tổng thể
UV	Ultraviolet	Tia cực tím
VLC	Vacuum Liquid Chromatography	Sắc ký lỏng chân không
VOC	Volatile organic compound	Hợp chất hữu cơ dễ bay hơi

MỞ ĐẦU

1. Đặt vấn đề

Lan tai cáo (*Hoya verticillata* var. *verticillata* hoặc *Hoya parasitica* (Rox) Wall ex Traill) là loài thực vật sống phụ sinh lâu năm ở rừng ẩm nhiệt đới thuộc chi *Hoya*. *Hoya* là một chi thực vật có hoa trong họ Apocynaceae, hầu hết là những cây dây leo có lá mỏng nước và cụm hoa hấp dẫn. Lan tai cáo được tìm thấy ở khu vực từ Đông Ấn Độ đến Nam Trung Quốc và một số nước Đông Nam Á như Philippines, Indonesia, Lào, Myanmar, Campuchia, Thái Lan và Việt Nam.

Trong thực tế, Lan tai cáo rất đa dạng và có nhiều đặc điểm hình thái tương tự một số loài trong chi *Hoya*, đặc biệt khi cây ở trạng thái không còn nguyên vẹn hoặc ở dạng bột, việc xác định loài gặp nhiều khó khăn hoặc dễ nhầm lẫn Lan tai cáo với các loài khác cùng chi nếu chỉ dựa vào hình thái. Chính vì vậy, việc giải trình tự DNA lục lạp và tìm kiếm vùng DNA trong hệ gene lục lạp làm chỉ thị hỗ trợ phân biệt loài Lan tai cáo (*H. verticillata* var. *verticillata*) với các loài khác trong chi *Hoya* là cần thiết và mang tính thời sự.

Nghiên cứu hệ gene lục lạp của thực vật đã đạt được những thành tựu đáng kể và có ý nghĩa rất quan trọng, góp phần nâng cao nhận thức của con người về quá trình quang hợp, phân loại thực vật và tiến hóa. Những kết quả phân tích DNA lục lạp (chloroplast DNA-cpDNA) đã được chứng minh là có giá trị trong việc tìm hiểu sự tiến hóa, tính đa dạng và mối quan hệ phát sinh chủng loại của thực vật. Những tiến bộ trong công nghệ giải trình tự thế hệ mới, đã thúc đẩy mạnh mẽ quá trình hình thành cơ sở dữ liệu cpDNA. Cơ sở dữ liệu chuyên dụng, các công cụ Tin sinh học phát triển đã góp phần nâng cao hiệu quả quản lý và phân tích khối lượng ngày càng tăng của trình tự DNA lục lạp.

Một số nghiên cứu cho thấy Lan tai cáo được sử dụng làm thuốc chống suy thận cấp tính, chữa trị và làm lành các vết thương. Một số nghiên cứu bước đầu đã ghi nhận các loài thuộc chi *Hoya*, trong đó có Lan tai cáo chứa các hợp chất flavonoid, phenolic, saponin. Tuy nhiên, nghiên cứu thành phần hóa học, tìm kiếm các chất mới và khảo sát hoạt tính sinh học của chúng làm cơ sở đề xuất sử dụng trong y học còn ít được quan

tâm nghiên cứu. Vấn đề đặt ra là, Lan tai cáo (*H. verticillata* var. *verticillata*) chứa các chất có hoạt tính sinh học nào và có thể ứng dụng các chất này trong y học hay không? là những câu hỏi tiếp theo cần được làm sáng tỏ.

Từ những lý do trên chúng tôi thực hiện đề tài luận án: **“Nghiên cứu đặc điểm hệ gene lục lạp và phân tích hoạt tính của saponin phân lập từ cây Lan tai cáo (*Hoya verticillata* var. *verticillata*)”**.

2. Mục tiêu nghiên cứu

Phân tích đặc điểm hệ gene lục lạp, thiết lập được sơ đồ cây phát sinh chủng loại của loài Lan tai cáo và các loài khác thuộc chi *Hoya*.

Phân lập và xác định được cấu trúc hợp chất saponin tách chiết từ Lan tai cáo và phân tích hoạt tính sinh học *in silico* của các saponin đã phân lập.

3. Nội dung nghiên cứu

3.1. Nghiên cứu giải trình tự hệ gene lục lạp của loài Lan tai cáo.

- i) Thu thập mẫu cây, phân loại tạo nguyên liệu cho phân tích sinh học phân tử và phân tích thành phần hóa học Lan tai cáo.
- ii) Định danh loài Lan tai cáo (*H. verticillata* var. *verticillata*) bằng phương pháp hình thái so sánh.
- iii) Giải trình tự nucleotide của DNA lục lạp và thiết lập sơ đồ cấu trúc hệ gene lục lạp cây Lan tai cáo. Mô tả đặc điểm hệ gene lục lạp của Lan tai cáo.

3.2. Phân tích sự phát sinh chủng loại, khảo sát một số chỉ thị DNA có độ đa dạng nucleotide cao từ hệ gene lục lạp của loài Lan tai cáo và đề xuất mã vạch tiềm năng cho loài/chi.

- i) Xác định và thu thập dữ liệu trình tự hệ gene lục lạp của các loài thuộc chi *Hoya* trên GenBank. Phân tích tiến hóa phân tử dựa trên trình tự DNA lục lạp của các loài này.
- ii) Xác định và thu thập một số trình tự DNA có độ đa dạng nucleotide cao, xây dựng sơ đồ phát sinh chủng loại dựa trên dữ liệu về các trình tự khai thác được từ GenBank, đề xuất mã vạch tiềm năng sử dụng trong định danh loài và chi *Hoya*.

3.3. Nghiên cứu phân lập, xác định cấu trúc và phân tích hoạt tính sinh học của một số saponin từ cây Lan tai cáo.

i) Đánh giá hoạt tính sinh học của cao chiết phân đoạn chứa saponin từ lá Lan tai cáo.

ii) Phân lập saponin và xác định cấu trúc của các saponin đã phân lập.

iii) Phân tích hoạt tính gây độc tế bào ung thư của các saponin phân lập từ Lan tai cáo thông qua mô hình tương tác phân tử và phân tích động học phân tử *in silico*.

4. Những đóng góp mới của luận án

4.1. Luận án là công trình nghiên cứu mới về đặc điểm của hệ gene lục lập hoàn chỉnh của loài Lan tai cáo (*H. verticillata* var. *verticillata*). Đề xuất được mã vạch DNA tiềm năng hỗ trợ phân biệt loài và nhận diện chi *Hoya*.

4.2. Phân lập và xác định được cấu trúc của hai saponin mới từ lá Lan tai cáo là 3 β ,20-dihydroxy-pregn-5-ene-3-O- β -D-fucopyranoside 20-O- β -D-glucopyranoside và 3 β ,20-dihydroxy-pregn-5-ene-3-O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-fucopyranoside-20-O- β -D-glucopyranoside

4.3. Đánh giá được khả năng tương tác và mô phỏng động học phân tử *in silico* của các saponin phân lập từ Lan tai cáo và xác định được parasiticoside B (3 β ,20-dihydroxy-pregn-5-ene-3-O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-fucopyranoside-20-O- β -D-glucopyranoside) có hoạt tính gây độc cao nhất chống lại protein ung thư Bcl-2. Parasiticoside B là ứng viên tiềm năng cho nghiên cứu thực nghiệm hoạt tính chống lại protein Bcl-2 trong ung thư.

5. Ý nghĩa khoa học và thực tiễn của đề tài luận án

Kết quả đạt được của luận án có giá trị khoa học và thực tiễn trong tiếp cận nghiên cứu tiến hóa phân tử dựa trên trình tự hệ gene lục lập và phân lập, xác định hoạt tính của các saponin phục vụ trong nghiên cứu và ứng dụng dược phẩm từ cây Lan tai cáo.

Về mặt khoa học

Kết quả nghiên cứu góp phần làm sáng tỏ cấu trúc hệ gene lục lập hoàn chỉnh của loài Lan tai cáo và cấu trúc hệ gene lục lập đặc trưng của các loài thuộc chi *Hoya*.

Kết quả bước đầu xây dựng được sơ đồ phát sinh chủng loại thực vật dựa trên trình tự hệ gene hoặc các vùng gene lục lạp, là cơ sở để đánh giá mối quan hệ gần gũi và mức độ tiến hóa phân tử của các loài trong chi *Hoya*.

Hai saponin mới phân lập được từ Lan tai cáo là dữ liệu bổ sung cho nghiên cứu sàng lọc chất có hoạt tính sinh học ứng dụng trong nghiên cứu sinh dược học.

Về mặt thực tiễn

Trình tự DNA lục lạp hoàn chỉnh của Lan tai cáo (*H. verticillata* var. *verticillata*) đã bổ sung nguồn dữ liệu phục vụ phân tích tiến hóa phân tử của chi *Hoya*. Mã vạch DNA tiềm năng có thể sử dụng hỗ trợ định danh chi *Hoya* và loài Lan tai cáo.

Các bài báo đăng tải trên các tạp chí khoa học công nghệ quốc tế và trong nước cùng với cấu trúc hệ gene lục lạp hoàn chỉnh của loài Lan tai cáo công bố trên GenBank là những tư liệu có giá trị tham khảo trong thực tiễn nghiên cứu và đào tạo.

Chương 1. TỔNG QUAN TÀI LIỆU

1.1. ĐẶC ĐIỂM HỆ GENE LỤC LẠP CỦA THỰC VẬT VÀ CHI HOYA

1.1.1. Hệ gene lục lập ở thực vật bậc cao

Những nghiên cứu khoa học đầu tiên đã chứng minh rằng lục lập cũng có khả năng lưu trữ thông tin di truyền dưới dạng DNA. Công bố của Tewari và Wildman (1970) là nghiên cứu tiên phong về hàm lượng DNA và tổ chức của lục lập, đánh dấu một cột mốc quan trọng trong lĩnh vực sinh học phân tử thực vật. Tewari và Wildman đã định lượng DNA có trong lục lập và so sánh với hàm lượng DNA trong nhân, chỉ ra rằng DNA lục lập có cấu trúc hình vòng và mô tả một số gene của DNA lục lập có chức năng chủ yếu là mã hóa các protein liên quan đến quá trình quang hợp (thành phần của chuỗi vận chuyển điện tử quang hợp và các enzyme cần thiết cho quá trình cố định carbon). Nghiên cứu này đặt nền móng cho các nghiên cứu tiếp theo về hệ gene lục lập, bao gồm nỗ lực giải trình tự và so sánh hệ gene giữa các loài thực vật khác nhau. Nghiên cứu đã nhấn mạnh những đặc điểm độc đáo của DNA lục lập và tầm quan trọng trong sinh học và tiến hóa thực vật [152].

Năm 1972, Sager và Lane đã nghiên cứu các đặc điểm và chức năng của lục lập ở tảo xanh *Chlamydomonas reinhardtii*, trong đó tiến hành nghiên cứu rộng về các đột biến của *Chlamydomonas* có ảnh hưởng đến cấu trúc và chức năng của lục lập [124]. Kung và cs (1977) cũng cho rằng DNA lục lập có dạng vòng và bản đồ hệ gene lục lập đã được ông mô tả. Nghiên cứu đã làm sáng tỏ thêm về tổ chức của hệ gene lục lập bằng cách xác định các vùng cụ thể và vị trí tương đối của chúng trong phân tử DNA lục lập. Đồng thời, cung cấp phương pháp nghiên cứu và những thông tin rất quan trọng để làm rõ cấu trúc di truyền và ý nghĩa chức năng của DNA lục lập. Trên cơ sở lập bản đồ gene lục lập của các loài thực vật khác nhau, Kung và cs (1977) đã phân tích và so sánh để khám phá mối quan hệ tiến hóa của các loài thực vật [76].

Nghiên cứu năm 1978 của Kirk và Tilney-Bassett tập trung vào phân tích về quá trình sao chép và cơ chế di truyền của hệ gene lục lập ở thực vật, bao gồm thời gian, các enzyme liên quan và các yếu tố điều hòa kiểm soát quá trình sao chép DNA lục lập. Một phát hiện quan trọng của nghiên cứu này là đã xác nhận, lục lập chủ yếu được di truyền

từ mẹ ở hầu hết các loài thực vật, phát hiện này rất quan trọng để có những hiểu biết sâu sắc về cơ sở của sự di truyền các bào quan ở thực vật cũng như sinh lý thực vật [73].

Gillham và cs (1978) đã tiến hành các nghiên cứu lập bản đồ di truyền để mô tả hệ gene lục lạp của *Chlamydomonas reinhardtii*. Các tác giả đã xác định và lập bản đồ các gene khác nhau trong DNA lục lạp, đặc biệt là những gene liên quan đến quá trình quang hợp và các quá trình trao đổi chất thiết yếu khác. Nghiên cứu này đã xác định được các cơ chế biểu hiện và điều hòa gene trong lục lạp. Nghiên cứu của Gillham cung cấp những hiểu biết sâu sắc về cách các gene lục lạp được phiên mã, hoàn chỉnh và dịch mã để tạo ra các protein chức năng cần thiết cho quá trình quang hợp và chuyển hóa tế bào. Nhóm của Gillham đã sử dụng các chủng *Chlamydomonas* đột biến có khiếm khuyết cụ thể trong hệ gene lục lạp để nghiên cứu chức năng của các gene này cũng như vai trò của chúng trong các quá trình sinh học và chức năng của lục lạp. Nghiên cứu phát triển kỹ thuật di truyền để đưa các gene ngoại lai vào hệ gene lục lạp nhằm nghiên cứu chức năng gene và xử lý lục lạp cho các ứng dụng công nghệ sinh học [42], [49].

Mô hình cấu trúc hệ gene hoàn chỉnh đầu tiên được xây dựng là trình tự nucleotide hoàn chỉnh của cây thuốc lá được Shinozaki và cs công bố năm 1986. Công trình nghiên cứu của Shinozaki và cs đã tìm ra được cách thức thực vật điều chỉnh biểu hiện gene trong lục lạp dưới áp lực của điều kiện môi trường khác nhau như hạn hán, nhiệt độ khắc nghiệt [131].

Năm 2017, Zhang và cs đã phát triển Thư viện tiêu chuẩn tham khảo về trình tự hệ gene lục lạp (GenomeTrakrCP) trên NCBI nhằm xây dựng cơ sở dữ liệu toàn diện về trình tự hệ gene lục lạp hoàn chỉnh nhằm giúp tăng cường khả năng định danh, hỗ trợ các nghiên cứu về gene, về sự phát sinh chủng loại và tiến hóa của thực vật [175]. Đến tháng 12 năm 2018, đã có 3654 hệ gene lục lạp hoàn chỉnh được công bố, kết quả này đã nâng cao đáng kể sự hiểu biết của chúng ta về quá trình tiến hóa thực vật, cung cấp một bức tranh khá chi tiết về hệ gene lục lạp ở các loài thực vật từ đó đánh giá được mức độ đa dạng và ý nghĩa của nó đối với thích nghi theo thời gian của thực vật. Những hiểu biết sâu sắc này có ý nghĩa quan trọng đối với các lĩnh vực như phân loại học, tiến hóa phân tử, sinh học thực vật, nông nghiệp, bảo tồn thực vật để ứng dụng trong y học [172]. Đến năm 2020, cơ sở dữ liệu về các vùng gene và trình tự đầy đủ về hệ gene lục

lập của 3823 loài thực vật thuộc 1527 chi và 256 họ đã được công bố [134]. Trên cơ sở những tiến bộ về công nghệ giải trình tự hiệu suất cao và sự quan tâm ngày càng tăng đối với hệ gene của thực vật, đến năm 2024, đã có hơn 6.000 hệ gene lục lập hoàn chỉnh được công bố trên cơ sở dữ liệu NCBI [160].

Đặc điểm đáng nổi bật của hệ gene lục lập thực vật hạt kín là đều có cấu trúc 4 vùng, vùng sao chép đơn lớn (Large single-copy-LSC), vùng sao chép đơn nhỏ (Small single-copy-SSC), hai vùng lặp lại đảo ngược (Inverted repeats-IR). Những nỗ lực giải trình tự hệ gene lục lập đã tạo điều kiện thuận lợi cho sự phát triển của nhiều lĩnh vực nghiên cứu khác nhau, bao gồm nghiên cứu chức năng hệ gene thực vật, nghiên cứu tiến hóa phân tử và phân tích phát sinh chủng loại. Sự sẵn có ngày càng tăng của trình tự DNA lục lập trong các cơ sở dữ liệu đã cho phép phân tích so sánh toàn diện, tiết lộ những hiểu biết sâu sắc về quá trình tiến hóa của hệ gene lục lập và của thực vật.

1.1.2. Hệ gene lục lập của các loài thuộc chi *Hoya*

1.1.2.1. Chi *Hoya* và loài *Hoya verticillata* var. *verticillata*

Chi *Hoya* có 562 loài (Plants of the World Online: POWO) và là chi lớn nhất trong họ Apocynaceae, các loài trong chi này là loại cây bụi hoặc dây leo, sống biểu sinh, phân bố ở vùng nhiệt đới và cận nhiệt đới từ dãy Himalaya đến Okinawa, Nam và Đông Nam Á, Úc và quần đảo Fiji [43], [146]. *Hoya* là chi trong tông Marsdenieae (gồm 17 chi), cụ thể là *Hoya*, *Dischidia*, *Lygisma*, *Sarcolobus*, *pseudosarcolobus*, *Spirella*, *Gymnemopsis*, *Harmandiella*, *Gymnema*, *Dischidanthus*, *Marsdenia*, *Dregea*, *Cosmostigma*, *Telosma*, *Campestigma*, *Pentasacme*, *Heterostemma* [166].

Theo POWO, trong 562 loài của chi *Hoya*, *H. verticillata* có 32 tên đồng nghĩa dị loại (heterotypic synonyms), trong đó có nhiều đơn vị dưới loài (variety) ví dụ như *H. verticillata* var. *verticillata*, *H. parasitica* (Roxb.) Wall. ex Wight), *H. pottsii* var. *angustifolia*, *H. parasitica* var. *geoffrayi* Costantin, ... Dựa vào hệ thống phân loại của S. Liede và F. Albert (1994), chi *Hoya* thuộc họ Apocynaceae ở Việt Nam gồm 31 loài, trong đó có loài Lan tai cáo [2]. Về phân loại ở Việt Nam, Lan tai cáo (Dây lười lợn, Cầm cù) có tên khoa học là (*H. verticillata* var. *verticillata* là một trong 31 loài thuộc chi *Hoya* được tìm thấy ở Việt Nam, chúng là những loài thực vật phụ sinh. Lan

tai cáo (*H. verticillata* var. *verticillata*) là loài thảo dược dùng làm thuốc uống có tác dụng lợi sữa, dùng để đắp rịt vết thương chống mất máu và nhiễm khuẩn vết thương, đồng thời còn được dùng để làm cây cảnh (Hình 1.1) [2].



Hình 1.1. Hình ảnh loài Lan tai cáo (*H. verticillata* var. *verticillata*). A, B, C: Mẫu ép của *H. verticillata* var. *verticillata*.

(<https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:77173117-1>); D, E, G: Cây và cành mang hoa, quả của loài Lan tai cáo (Ảnh tác giả chụp)

Lan tai cáo được tìm thấy ở nhiều địa phương của Việt Nam như: Tuyên Quang, Thái Nguyên, Quảng Ninh (Tiên Yên), Thành phố Hà Nội, Thành phố Hải Phòng (Vườn quốc gia Cát Bà), Hà Nam (Võ Xá, Ninh Bình cũ), Thành phố Hồ Chí Minh (Bà Rịa: Núi Đỉnh) ... Loại cây dạng dây leo phụ sinh. Lá mọc đối, nạc, dày, nhiều hình thái khác nhau như: hình trứng, bầu dục hay hình mác; cỡ $5-18 \times 3-7$ cm; đầu nhọn, có mũi; gốc tròn hay nêm; nhẵn 2 mặt; gân gốc 3-5, gân bên 2 cặp, rõ; cuống mập, dài 6-20 mm. Cụm hoa hình tán, nhiều hoa; cuống cụm hoa dài 5-10 cm, đỉnh hình trụ và dày, cỡ $3-5 \times 0,5$ cm, nhiều sọc. Cuống hoa dài 1,5-2 cm, nhẵn hay có lông. Nụ hoa hình con quay, 5 cạnh. Hoa màu trắng hay vàng, đôi khi ở đỉnh có màu da cam, hồng hay tím, thơm. Thủy đài hình trứng. Tròng hình bánh xe, nạc, tiền khai hoa van; thủy tròng hình trứng-tam giác, ngắn hơn 4 mm. Vây tròng phụ ngắn hơn 3 mm, cạnh ngoài nhọn. Chỉ nhị dính nhau; bao phấn 2 ô, có phần phụ chất màng mỏng ở đỉnh; hạt phấn dính thành khối phấn và có sáp bao bên ngoài vách khối phấn, khối phấn không có mòm ở đỉnh; cơ quan truyền phấn có gót dính và 2 chuỗi; khối phấn hướng lên, dài khoảng 500 μ m; chỉ có một khối phấn trong mỗi ô phấn. Quả đại, cỡ $14 \times 0,4$ cm. Hạt cỡ $3-4 \times 1$ mm, có mào lông ở đỉnh. Mùa hoa tháng 7-8 [2].

1.1.2.2. Đặc điểm hệ gene lục lập của các loài thuộc chi *Hoya*

Những nghiên cứu về phân loại trước đây đối với các loài trong chi *Hoya* chủ yếu dựa trên đặc điểm hình thái như thân, lá, hoa, nhựa mủ [2], [17], [117]. Tuy nhiên, việc phân loại như vậy chưa đảm bảo độ tin cậy trong trường hợp các loài trong chi *Hoya* có nhiều các đặc điểm hình thái tương tự nhau hoặc các cơ quan sinh dưỡng bị biến dạng. Dữ liệu của nhiều hệ gene lục lập hoàn chỉnh được giải trình tự và những phần mềm chuyên dụng đã cung cấp nguồn vật liệu có giá trị cũng như các phương tiện hiện đại trong nghiên cứu tiến hóa, phân loại.

Trình tự hệ gene lục lập hoàn chỉnh đầu tiên của các loài trong chi *Hoya* được nghiên cứu, giải trình tự, công bố năm 2018 là của hai loài *H. pottsii* Traill và *H. liangii* Tsiang (lấy mẫu từ Hải Nam, Trung Quốc (110,111°Đ, 19,897°N)) [146]. Hệ gene lục lập của *H. pottsii* và *H. liangii* được phát hiện có tổng chiều dài lần lượt là 161.565 bp và 162.989 bp. Cả 2 hệ gene lục lập trên đều có cấu trúc 4 vùng điển hình của hệ gene lục lập thực vật hạt kín, chứa 2 vùng lặp lại đảo ngược (IR) là 24.657 bp (*H. pottsii*) và 24.841 bp (*H. liangii*), 1 vùng sao chép đơn lớn (LSC) là 92.532 bp (*H. pottsii*) và 93.292 bp (*H. liangii*), 1 vùng sao chép đơn nhỏ (SSC) gồm 19.719 bp (*H. pottsii*) và 20.015 bp (*H. liangii*). Hệ gene lục lập của loài *H. pottsii* và *H. liangii* chứa 115 gene, bao gồm 81 gene mã hóa protein (7 gene trong vùng IR), 30 gene tRNA (7 trong vùng IR) và 4 gene rRNA; 5 gene giả (*rpoC1*, *accD*, *ycf2*, *ndhH*, *ycf1*) ở *H. pottsii* và 6 gene giả (*rpoC2*, *accD*, *ycf2*, *ycf1*, *ndhH*, *ycf1*) ở *H. liangii*, 14 gene chứa một intron và 3 gene (*ycf3*, *clpP*, *rps12*) chứa 2 intron. Gene *rps12* được phát hiện là có tính chất vận động, đặc trưng của thực vật hạt kín. Hàm lượng A/T tổng thể trong hệ gene lục lập của *H. pottsii* là 62,40%, trong đó giá trị tương ứng của vùng LSC, SSC và IR lần lượt là 64,10; 68,20 và 56,80%. Hàm lượng A/T tổng trong hệ gene lục lập của *H. liangii* là 62,30%, trong đó giá trị tương ứng của vùng LSC, SSC và IR lần lượt là 64,10; 68,20 và 56,80%.

Rodda và Niissalo (2021) đã phân tích tương đối đầy đủ về sự tiến hóa trong tổ chức hệ gene lục lập của các loài trong chi *Hoya*, ngoài ra có thêm những phân tích so sánh giữa hệ gene lục lập và hệ gene ty thể để có những thông tin chính xác trong những kết luận của mình. Nghiên cứu chỉ ra rằng, trong mỗi tế bào có nhiều bản sao hệ gene lục lập. Tổng chiều dài của các hệ gene lục lập hoàn chỉnh thay đổi từ 175.405 bp

(*Dischidia milnei* Hemsl) đến 178.525 bp (*H. omlorii*) và từ 161.660–161.700 bp ở các nhóm ngoài. Độ dài của các vùng IR hoàn chỉnh thay đổi từ 41.272 (*D. milnei*) đến 42.069 bp (*H. exilis* Schltr.), từ 26.117–28.287 bp ở các nhóm ngoài. Vùng sao chép đơn lớn (LSC) hoàn chỉnh của các nhóm có độ dài khác nhau từ 90.248 bp (*H. ignorata*) đến 92.364 bp (*H. megagalaster* Warb. ex K.Schum. & Lauterb.), vùng sao chép đơn nhỏ (SSC) hoàn chỉnh thay đổi từ 2.285 bp (*H. diversifolia*) đến 2.304 bp (*H. omlorii*), nghiên cứu cũng đã chỉ ra vùng SSC ở nhóm *Hoya* trở thành vùng ngắn nhất được biết đến trong thực vật hạt kín [17], [118]. Cấu trúc hệ gene lục lạp ở tất cả các loài trong nhóm *Hoya* được đặc trưng bởi sự gia tăng lớn về kích thước của các trình tự lặp lại đảo ngược so với các nhóm bên ngoài; ranh giới IR/SSC đã chuyển từ *ycf1* sang *ndhF*. Loài ở nhóm ngoài có quan hệ gần gũi nhất với nhóm *Hoya* là *Marsdenia flavesces* A. Cunn [118].

Nghiên cứu năm 2022 của Odago và cs về 31 hệ gene mới được giải trình tự (*Dischidia australis*, *D. griffithii*, *D. nummularia*, *D. ruscifolia*, *H. angustifolia* (pottsii), *H. ariadna*, *H. caudata*, *H. chinghungensis*, *H. commutata*, *H. dimorpha*, *H. griffithii*, *H. kerrii*, *H. lacunosa*, *H. lanceolata* subs. *Bella*, *H. liangii*, *H. longifolia*, *H. meliflua* subs. *fraterna*, *H. ovalifolia*, *H. pandurata*, *H. pottsii*, *H. pubicalyx*, *H. radicalis*, *H. rigida*, *H. sylvatica*, *H. sp.* 11 ZCF6107, *H. sp.* 2 ZCF6006, *H. sp.* 3 ZCF6076, *H. sp.* 4 ZCF6004, *H. sp.* 8 ZCF6076, *H. thomsonii* và *H. volubilis*) và so sánh với những hệ gene lục lạp của chi *Hoya* đã được công bố nhằm làm rõ mối quan hệ tiến hóa giữa các loài *Hoya* và các loài liên quan khác [104]. Nghiên cứu đã xác định các vùng có độ biến đổi cao (*ndhF*, *ycf1*, *rpl22*, *matK*, *trnT-trnL* và *trnL-trnF*) và các vùng lặp lại (-mono, -tri, -penta và -hexanucleotide) được xác định là có tiềm năng với vai trò là chỉ thị phân tử cho phát sinh chủng loại. Những nghiên cứu phân tích phát sinh chủng loại đối với các loài thuộc nhóm *Hoya* đều hỗ trợ mạnh mẽ, góp phần làm sáng tỏ mối quan hệ giữa các loài trong chi. Các kết quả và dữ liệu được trình bày trong nghiên cứu này cung cấp cái nhìn sâu sắc về mối quan hệ tiến hóa và lịch sử địa sinh học của các loài thuộc nhóm *Hoya*.

Nghiên cứu của Wang và cs năm 2023 đã tập hợp được 101 hệ gene lục lạp hoàn chỉnh bao gồm 35 hệ gene lục lạp mới được giải trình tự, 24 hệ gene lục lạp từ dữ liệu

tho đã được công bố và phần còn lại là từ cơ sở dữ liệu NCBI (GenBank), đại diện cho 26 trong số 27 tông trong họ Apocynaceae, được sử dụng để phân tích, so sánh. Các phân tích phát sinh chủng loại được tiến hành bằng cách sử dụng ma trận dữ liệu gene lục lạp kết hợp của 77 gene mã hóa protein từ 162 đơn vị phân loại, bao gồm tất cả 41 trong số 49 phân tông của họ Apocynaceae. Chiều dài hệ gene lục lạp dao động từ 150.897 bp ở loài *Apocynum venetum* đến 178.616 bp ở *H. exilis*. Các nghiên cứu đã xác định được 6 loại ranh giới giữa vùng lặp lại đảo ngược (IR) và vùng LSC. Kích thước vùng lặp lại đảo ngược (IR) mở rộng trên các vùng SSC với mức độ khác nhau đã được tìm thấy ở 3 tông bao gồm Alyxieae, Ceropegieae và Marsdenieae trong họ Apocynaceae. Các vùng IR của tông Marsdenieae đã phát triển theo hai cách: Mở rộng sang vùng sao chép đơn lớn (LSC) ở Lygisma, Stephanotis, Ruehssia, Gymnema (Nhánh Cosmopolitan) và mở rộng sang cả vùng LSC và SSC ở Dischidia-Hoya và Marsdenia (nhánh châu Á-Thái Bình Dương). Có 6 gene mã hóa và 5 vùng không mã hóa được xác định là có tính biến đổi cao bao gồm *accD*, *ccsA-ndhD*, *clpP*, *matK*, *ndhF*, *ndhG-ndhI*, *trnG (GCC)-trnfM (CAU)*, *trnH (GUG)-psbA*, *trnY (GUA)-trnE (UUC)*, *ycf1* và *ycf2*. Khả năng tối đa và các phân tích phát sinh chủng loại Bayes đã dẫn đến cấu trúc các cây gần như giống hệt nhau và tạo ra một cây phát sinh tốt nhất chia Thiên lý thành 15 nhánh. Phân họ Periplocoideae được xếp vào lớp Apocynoid và là chị em với nhánh Echiteae-Odontadenieae-Mesechiteae với giá trị tin cậy cao. Ba tông (Melodineae, Vinceae và Willughbeieae), phân tông Amphineuriinae và 4 chi (*Beaumontia*, *Ceropegia*, *Hoya* và *Stephanotis*) không được coi là đơn ngành. Công trình nghiên cứu đã làm sáng tỏ các mối quan hệ phát sinh chủng loại cơ bản trong họ Apocynaceae và cung cấp những hiểu biết sâu sắc về sự tiến hóa của các hệ gene lục lạp của họ Apocynaceae thông qua việc sử dụng dữ liệu hệ gene lục lạp được lấy mẫu nhiều nhất cho đến nay [162].

Đến thời điểm hiện tại đã có khoảng gần 50 hệ gene lục lạp của các loài thuộc chi *Hoya* được công bố trên NCBI. Các nghiên cứu về hệ gene lục lạp của các loài trong chi *Hoya* đã góp phần làm sáng tỏ những đặc điểm cấu trúc độc đáo và những hiểu biết sâu sắc về tiến hóa của chúng. Hệ gene lục lạp của loài thuộc chi *Hoya* và các chi liên quan cho thấy sự mở rộng của các vùng lặp lại đảo ngược, dẫn tới vùng sao chép đơn

nhỏ bị thu hẹp đáng kể, khiến nó trở thành một trong những chi có vùng sao chép đơn nhỏ ngắn nhất được biết đến trong thực vật hạt kín [104], [118], [167]. Sự mở rộng này đã dẫn đến hầu hết các gene của vùng SSC tồn tại dạng cặp đôi [167] hoặc chỉ còn lại 1 gene *ndhF*. Hàm lượng và tổ chức gene tổng thể ở các loài *Hoya* phần lớn được bảo tồn, tuy nhiên cũng có những thay đổi nhỏ (ví dụ như mất gene hoặc tạo gene giả) [79]. Phân tích dữ liệu hệ gene lục lập hoàn chỉnh đã được công bố là cơ sở cho việc giải quyết các mối quan hệ giữa các loài trong nhóm *Hoya* [104]. Thông qua một số gene và vùng gene có khả năng biến đổi cao sẽ đề xuất được các chỉ thị phân tử DNA tiềm năng để xác định loài và nghiên cứu phát sinh chủng loại [79], [104]. Những phát hiện này bổ sung vào sự hiểu biết của chúng ta về quá trình tiến hóa hệ gene lục lập của *Hoya* và cung cấp nguồn thông tin cần thiết cho những nghiên cứu trong tương lai.

1.1.3. Nhận xét chung

Hệ gene lục lập của thực vật hạt kín nói chung và của chi *Hoya* nói riêng có cấu trúc điển hình gồm bốn vùng: Một vùng sao chép đơn lớn (LSC), một vùng sao chép đơn nhỏ (SSC) và hai vùng lặp đảo ngược (IR). Ở các loài thuộc chi *Hoya*, vùng IR thường có xu hướng mở rộng, dẫn đến sự thu hẹp đáng kể của vùng SSC, khiến *Hoya* trở thành một trong những chi thực vật hạt kín có vùng SSC ngắn nhất được ghi nhận cho đến nay.

Tính đến nay, đã có hơn 50 hệ gene lục lập của các loài *Hoya* được giải trình tự và công bố trên các cơ sở dữ liệu quốc tế, góp phần quan trọng trong việc làm sáng tỏ đặc điểm cấu trúc hệ gene, cũng như hỗ trợ các nghiên cứu về phân loại học và tiến hóa của chi này. Tuy nhiên, so với mức độ đa dạng loài rất cao của *Hoya*, số lượng hệ gene lục lập hoàn chỉnh đã được công bố vẫn còn tương đối hạn chế, cho thấy nhu cầu cấp thiết phải tiếp tục bổ sung dữ liệu hệ gene cho các loài và đơn vị phân loại còn thiếu.

Đáng chú ý, trong số các đơn vị dưới loài của *H. verticillata*, hiện mới chỉ có hệ gene lục lập của *H. pottsii* var. *angustifolia* được công bố, trong khi các biến thể khác, bao gồm *H. verticillata* var. *verticillata* (Lan tai cáo), vẫn chưa có bất kỳ dữ liệu hệ gene lục lập hoàn chỉnh nào được báo cáo. Vì vậy, việc giải trình tự và phân tích hệ gene lục lập của *H. verticillata* var. *verticillata* là hết sức cần thiết, không chỉ nhằm bổ sung dữ

liệu di truyền cho chi *Hoya*, mà còn góp phần làm rõ mối quan hệ phát sinh chủng loại và vị trí phân loại của nhóm này trong họ Apocynaceae.

1.2. PHÂN TÍCH SỰ PHÁT SINH CHỦNG LOẠI DỰA TRÊN TRÌNH TỰ DNA LỤC LẠP

1.2.1. Phát sinh chủng loại học phân tử

Phát sinh chủng loại phân tử liên quan đến việc suy luận mối quan hệ tiến hóa giữa các loài, các bậc phân loại trên loài bằng cách sử dụng dữ liệu phân tử. Thông qua mô hình hóa các thay đổi ở cấp độ phân tử trong chuỗi protein và DNA theo thời gian, thông qua đó tái tạo lịch sử tiến hóa của các loài, đồng thời đánh giá mức độ tin cậy của các suy luận. Các suy luận phát sinh chủng loại phân tử không chỉ hỗ trợ cho các phát sinh chủng loài truyền thống mà còn đóng vai trò quan trọng trong việc giải quyết một số vấn đề khó liên quan đến các sự phân nhánh trong hướng tiến hóa. Do cơ sở dữ liệu thông tin trình tự phân tử ngày càng tăng, lĩnh vực này có nhiều tiềm năng và sẽ là chìa khóa quan trọng để hiểu lịch sử và mối quan hệ của các dạng sống [19].

Thuyết trung tính cho rằng, phần lớn những thay đổi và tiến hóa ở cấp độ phân tử là do sự cố định ngẫu nhiên (do sự trôi dạt lấy mẫu ngẫu nhiên trong các quần thể hữu hạn) của các đột biến trung tính có chọn lọc. Thuyết này cũng khẳng định rằng hầu hết các biến dị di truyền của các loài ở cấp độ phân tử (như đa hình protein và DNA) là trung tính và chúng được duy trì trong quá trình phát triển của loài. Thuyết trung tính dựa trên các giả định đơn giản, cho phép phát triển các lý thuyết toán học dựa trên di truyền quần thể để phân tích sự tiến hóa và biến dị phân tử theo các thuật ngữ định lượng [72]. Sự tiến bộ không ngừng trong những nghiên cứu trong lĩnh vực sinh học phân tử tạo điều kiện thuận lợi cho việc áp dụng lý thuyết di truyền quần thể vào phân tích dữ liệu thực tế. Thông tin di truyền được lưu trữ trong các chuỗi DNA được di truyền ổn định từ thế hệ này sang thế hệ khác, tuy nhiên vẫn xảy ra quá trình biến đổi và tích lũy biến đổi đối với vật chất di truyền. Người ta cũng có thể so sánh các cấu trúc bậc hai, bậc ba của protein và nucleic acid từ nhiều nguồn khác nhau cho thấy hầu hết các thay đổi tiến hóa ở cấp độ phân tử là do trôi dạt di truyền ngẫu nhiên của các đột biến trung tính có chọn lọc hoặc gần trung tính chứ không phải do chọn lọc tự nhiên [105]. Kimura

đưa ra giả thuyết đồng hồ phân tử cho rằng các đột biến (đặc biệt là đột biến trung tính) tích tụ trong chuỗi DNA hoặc protein với tốc độ gần như không đổi theo thời gian, cho phép các nhà khoa học ước tính thời điểm phân kỳ tiến hóa giữa các loài [72].

Xây dựng cây phát sinh chủng loại là một công cụ quan trọng trong các nghiên cứu tiến hóa phân tử, cho phép các nhà nghiên cứu hình dung và phân tích các mối quan hệ tiến hóa giữa các sinh vật. Có nhiều phương pháp và phần mềm khác nhau đã được ứng dụng để xây dựng sơ đồ phát sinh chủng loại, bao gồm các phương pháp dựa trên khoảng cách, dựa trên ký tự và khả năng tối đa [107]. Các phần mềm thường được sử dụng để xây dựng sơ đồ phát sinh chủng loại gồm: Phylip, PAUP, PAML, TreePuzzle, PhyML, MyBayes, MEGA, DNASTAR...

Mối quan hệ di truyền giữa các sinh vật có thể được ước lượng từ mô hình thay đổi của các nucleotide trong trình tự DNA. Vấn đề đầu tiên đặt ra là cần phải chọn một gene, vùng gene hoặc hệ gene làm vật liệu di truyền/chỉ thị phân tử phù hợp cho nghiên cứu, phân tích. Việc lựa chọn, sử dụng đoạn DNA đã được xác định làm chỉ thị phân tử để định danh loài sinh vật gọi là mã vạch DNA (DNA barcoding, chỉ thị DNA). Các lựa chọn phổ biến ở thực vật bao gồm hệ gene lục lạp hoặc các thành phần của hệ gene lục lạp, gene RNA ribosome nhân hoặc gene nhân. Vấn đề thứ hai là cần phải lựa chọn phương pháp phù hợp sử dụng để phân tích, so sánh, phân kỳ trình tự DNA từ đó xây dựng bản đồ mô tả đặc điểm của phân tử DNA đang được nghiên cứu. Các nghiên cứu thực nghiệm chỉ ra rằng việc phân tích những vị trí nhất định trên phân tử DNA lục lạp cung cấp độ phân giải tốt cho những nghiên cứu hệ thống phân loại ở cấp độ loài, chi, họ. Do đó việc giải trình tự DNA, trình tự gene hoặc vùng DNA lục lạp quan tâm có vai trò rất quan trọng, cung cấp dữ liệu lớn giúp phân tích mối quan hệ di truyền và tiến hóa sâu sắc hơn [29].

1.2.2. Phân tích sự phát sinh chủng loại ở thực vật

Phân tích sự phát sinh chủng loại dựa trên hệ gene lục lạp đã nâng cao đáng kể sự hiểu biết về quá trình tiến hóa của thực vật, cung cấp những thông tin về nguồn gốc, sự đa dạng và chiều hướng tiến hóa thích nghi của các loài thực vật. Phát sinh chủng loại dựa trên trình tự DNA lục lạp liên quan đến việc sử dụng thông tin di truyền có trong lục lạp của thực vật để suy ra mối quan hệ tiến hóa giữa các loài khác nhau. Sử

dụng trình tự gene lục lạp để phân tích phát sinh chủng loại là một công cụ mạnh mẽ và hiệu quả. Hệ gene lục lạp có tính bảo tồn cao khiến chúng đặc biệt hữu ích cho mục đích này [157]. Các vùng khác nhau của hệ gene lục lạp, bao gồm cả các gene mã hóa (ví dụ: *rbcL*, *matK*) và các vùng không mã hóa (ví dụ: intron, spacers), có thể được sử dụng cho các nghiên cứu phát sinh chủng loại ở các cấp độ phân loại khác nhau [153]. Trong khi các gene mã hóa thường có nhiều thông tin hơn để giải quyết các mối quan hệ ở cấp độ phân loại cao hơn, thì những vùng không mã hóa thường hữu ích hơn ở các cấp độ phân loại thấp hơn do tốc độ tiến hóa của chúng nhanh hơn [153].

Những tiến bộ gần đây trong công nghệ giải trình tự gene thế hệ mới đã cho phép phân tích trình tự hệ gene lục lạp hoàn chỉnh từ DNA tổng số, cung cấp một phương pháp hiệu quả để xác định thực vật và thiết lập các mối quan hệ di truyền. Phương pháp này đã được áp dụng thành công cho các loài cây cỏ và xác định chính xác mối quan hệ phát sinh chủng loại của chúng. Những nghiên cứu phân tích quy mô lớn sử dụng hệ gene lục lạp đã cung cấp rất nhiều cây phát sinh có độ phân giải tốt cho hầu hết các họ thực vật có hoa [81]. Việc sử dụng phân tích dữ liệu gene nhân đã bổ sung thêm những hiểu biết về các mối quan hệ nhánh chính và thời gian phân kỳ của các sơ đồ phát sinh chủng loại [174]. Việc tích hợp các loại dữ liệu đa dạng, phương pháp phân tích tiếp tục hoàn thiện góp phần nâng cao sự hiểu biết về phát sinh chủng loại thực vật hạt kín, cung cấp nền tảng vững chắc cho các nghiên cứu tiến hóa. Tuy nhiên, một số mối quan hệ giữa các nhánh phân loại chính vẫn chưa được làm sáng tỏ, cần tiếp tục có thêm những nghiên cứu giải trình tự gene và phát triển, hoàn thiện các phương pháp nghiên cứu tiến hóa hiện có [81].

Các nghiên cứu gần đây sử dụng hệ gene lục lạp hoàn chỉnh để nghiên cứu phát sinh chủng loại của chi *Sinodielsia* và các chi liên quan của họ Apiaceae (Họ Hoa tán), qua phân tích 39 hệ gene lục lạp đã xác định được 6 vùng gene là những vùng có khả năng đột biến cao trong chi *Sinodielsia* [168]. Gou và cs (2020) đã nghiên cứu hệ gene lục lạp của 8 loài trong các chi *Hansenia*, *Haplosphaera* và *Sinodielsia*, là cơ sở để cung cấp những hiểu biết về phát sinh chủng loại và cung cấp mã vạch DNA tiềm năng phục vụ định danh loài trong các chi này [50]. Cả hai nghiên cứu đều nhận thấy các vùng *ndhF-rpl32* và *ycf1* rất đa dạng và hữu ích cho các phân tích phát sinh chủng loại. Các

nghiên cứu về phát sinh thực vật dựa trên chỉ thị phân tử là vùng gene, gene, hệ gene lục lập chỉ ra rằng một số gene, đặc biệt là những gene liên quan đến quá trình quang hợp, có dấu hiệu tiến hóa thích nghi đáp ứng với môi trường sống đa dạng [50] do đó dễ biến đổi hơn và có ý nghĩa quan trọng trong nghiên cứu phát sinh chủng loại [54].

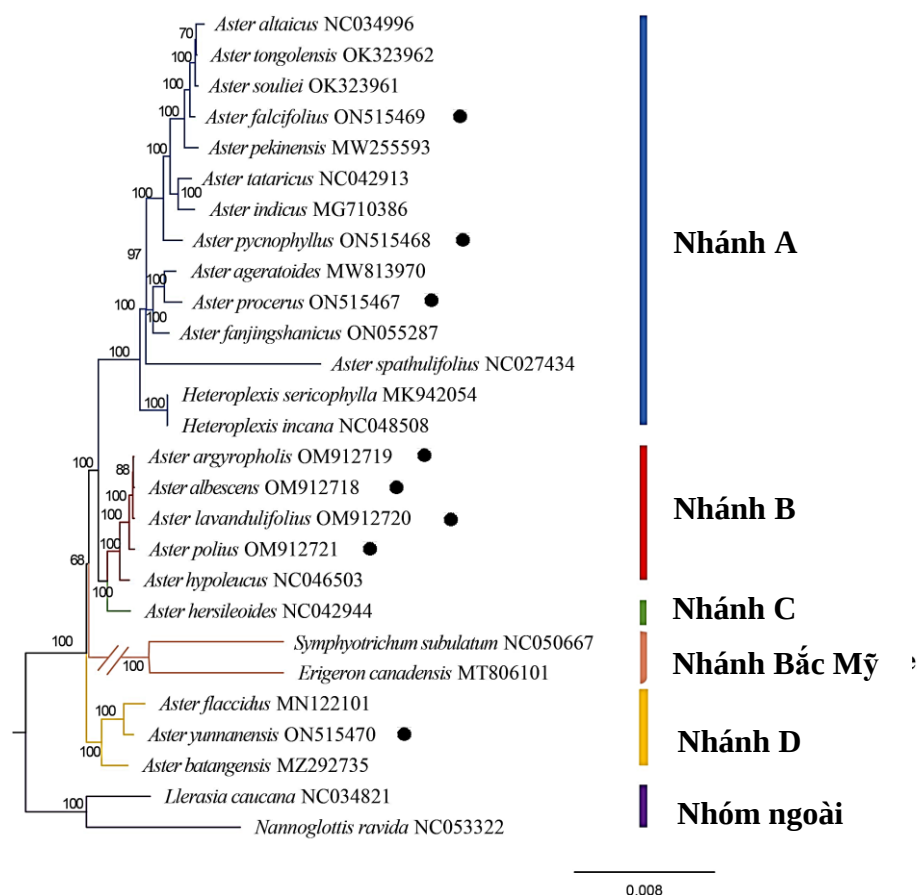
Li và cs (2021) đã đề xuất cây phân tích con đường phát sinh chủng loại (PPA II), là một sơ đồ phát sinh chủng loại mở rộng có nguồn gốc từ việc phân tích một tập dữ liệu lớn về hệ gene lục lập. Sơ đồ này nhằm mục đích giải quyết các mối quan hệ tiến hóa của thực vật hạt kín (thực vật có hoa) ở cấp độ họ trở lên. Cây PPA II trong nghiên cứu này được xây dựng dựa trên một tập dữ liệu mở rộng bao gồm 4792 trình tự, trong đó 4627 mẫu đại diện cho 4498 loài thực vật hạt kín trong tất cả các họ và bộ thực vật hạt kín đã được công bố. Phạm vi nghiên cứu có độ bao phủ đáng kể đối với các chi và loài, cung cấp một cái nhìn tổng quan toàn diện về sự đa dạng của thực vật hạt kín [81]. Orton và cs (2021) đã phân tích 313 hệ gene lục lập của họ Poaceae (Họ Hòa thảo) nhằm có thêm những thông tin về mối quan hệ giữa phân họ Pooideae (Phân họ Hòa thảo), gồm 164 đại diện với các loài khác trong họ (164 Pooideae, 40 Bambusoideae, 15 Oryzoideae, 44 Panicoideae, 20 Chloridoideae, 6 Danthonioideae, 6 Arundinoideae, 5 Micrairoideae, 6 Aristidoideae, 1 Pueloideae, 4 Pharoideae và 2 Anomochlooideae). Tổng cộng, 47 hệ gene lục lập đã được giải trình tự và lắp ráp hoàn chỉnh cho nghiên cứu này. Đây là nghiên cứu lớn nhất thuộc loại này bao gồm dữ liệu ở cấp độ hệ gene, không chỉ tăng cường bổ sung dữ liệu phân tử mà còn là cơ sở cho phân loại và xây dựng sơ đồ phát sinh tiến hóa BEAST2 (Cây lấy mẫu phân tích tiến hóa Bayes) phục vụ cho các nghiên cứu tiến hóa và phát sinh chủng loại [106].

Nghiên cứu của Gao và cs đối với chi *Rosa* (Hoa hồng) dựa trên trình tự hệ gene lục lập hoàn chỉnh đã cho thấy các vùng lặp lại đơn lớn (có số lượng từ 100-116 bp hoặc từ 78-90 bp) cùng một số vùng đệm giữa các gene (*ycf3-trnS*, *trnT-trnL*, *psbE-petL* và *ycf1*) được xác định là những ứng cử viên chỉ thị phân tử để phân biệt các loài trong chi *Rosa*. Nghiên cứu chỉ ra mối quan hệ giữa các hệ gene lục lập được chọn lọc, phân tích phát sinh chủng loại trong chi *Rosa* góp phần nhận biết và bảo tồn những loài hoa hồng hoang dã, quý hiếm [47].

Năm 2023, Su và cs đã giải trình tự, chú thích đầy đủ hệ gene lục lạp của 13 loài thuộc tông Datureae (Cà độc dược) và 2 hệ gene từ loài *Solanaceae* được lấy từ cơ sở dữ liệu NCBI. Nghiên cứu cho thấy hệ gene lục lạp có kích thước dao động từ 154.686 đến 155.979 bp và từ 155.497 đến 155.919 bp tương ứng đối với các loài *Datura* và *Brugmansia*. Đối với *Datura* và *Brugmansia*, tổng cộng đã xác định được lần lượt là 128 và 132 gene, trong đó 83 và 87 gene mã hóa protein, 37 gene tRNA và 8 gene rRNA. Phân tích các trình tự lặp lại chỉ ra rằng số lượng và loại khác nhau giữa các loài đối với trình tự lặp lại trình tự đơn giản (SSR), lặp lại trình tự dài (LSR) và lặp lại song song lần lượt có số lượng từ 53 đến 59, 98 đến 99 và 22 đến 30. Phân tích phát sinh chủng loại dựa trên hệ gene lục lạp đã hỗ trợ mối quan hệ đơn ngành giữa *Datura*, *Brugmansia* và *Trompettia*, đồng thời giải quyết mối quan hệ phát sinh chủng loại giữa các loài [139].

Chen và cs (2024) đã giải trình tự hệ gene lục lạp hoàn chỉnh của 8 loài cúc trong chi *Aster* (Cúc tây) bằng cách sử dụng bộ dữ liệu giải trình tự thế hệ tiếp theo. Kết quả cho thấy các hệ gene lục lạp hoàn chỉnh của *Aster* có cấu trúc 4 phần. Những hệ gene này có chiều dài 152.045–152.729 bp và chứa 132–133 gene, bao gồm 87 gene mã hóa protein, 37–38 gene tRNA và 8 gene rRNA. Sự mở rộng hoặc thu hẹp của các vùng lặp lại đảo ngược, sao chép đơn lớn, sao chép đơn nhỏ cũng được phát hiện ở 8 loài chi *Aster*. Ngoài ra, các phân tích trong nghiên cứu cho thấy trình tự lặp lại đơn giản phong phú nhất là mononucleotide A/T và 14 vùng có tính biến đổi cao đã được phát hiện bằng cách phân tích các vùng giáp ranh, phân kỳ trình tự và các điểm nóng.

Các phân tích phát sinh chủng loại chỉ ra rằng 27 loài trong Astereae được tập hợp thành 6 nhánh, gồm các nhánh từ A đến D và các loài ngoài nhóm, giả thuyết ủng hộ rằng các chi *Heteropappus*, *Kalimeris* và *Heteroplexis* có quan hệ gần gũi với chi *Aster*. Kết quả các nhánh B đến D cho thấy có thể các nhánh này được coi là chi. Ước tính thời gian phân kỳ cho thấy các nhánh A, B, C và D phân kỳ lần lượt ở mức 23,15 Mya, 15,13 Mya, 24,29 Mya và 21,66 Mya (Mya-Triệu năm trước, là đơn vị thời gian thường được sử dụng trong phân tích phân tử để ước tính thời gian phân tách của các nhánh trong sơ đồ phát sinh chủng loại). Những kết quả nghiên cứu này đã làm sáng tỏ mối quan hệ phát sinh chủng loại của chi *Aster* và cung cấp thông tin hữu ích về việc xác định loài trong chi *Aster* và các chi khác có quan hệ gần gũi [26].



Hình 1.2. Sơ đồ phát sinh chủng loại có độ chính xác tối đa (ML) được suy ra từ 27 hệ gene lục lạp hoàn chỉnh của loài *Aster* (giá trị bootstrap được thể hiện trên các nhánh) [24].

Hệ gene lục lạp đã trở thành dữ liệu quan trọng và phân tích hệ gene lục lạp là công cụ mạnh mẽ trong việc làm sáng tỏ lịch sử tiến hóa của thực vật có hoa, cung cấp những hiểu biết sâu sắc về phát sinh chủng loại, tiến hóa thích nghi và định dạng loài. Những nghiên cứu này nhấn mạnh tầm quan trọng của việc tích hợp dữ liệu hệ gene lục lạp với dữ liệu hình thái và phân tử khác để có kết quả phân tích tiến hóa đầy đủ và toàn diện hơn.

1.2.3. Phân tích sự phát sinh chủng loại ở chi *Hoya*

Năm 2006, Wanntorp và cs là nhóm nghiên cứu tiên phong trong phân tích phát sinh chủng loại trong chi *Hoya* dựa trên hệ gene, tác giả xác định rằng các nhóm loài chính trong chi *Hoya* dựa trên phân tích, so sánh các vùng đệm *atpB-rbcL* lục lạp, gene *trnL* lục lạp kết hợp với vùng ITS (DNA ribosome nhân) [163], [165].

Nghiên cứu cho thấy, sơ đồ phát sinh chủng loại có độ phân giải tốt dựa trên trình tự của các vùng đệm hệ gene lục lạp là *trnH-psbA*, *trnT-trnL*, các vùng ITS và 5'-ETS

gene mã hóa vùng nhân của 154 loài. Nghiên cứu cũng đã phân chia các loài thuộc chi *Hoya* thành 6 dòng chính (Nhánh I–VI) có sự phân bố địa lý rộng khắp, các nhóm biểu sinh phân kỳ sớm nhất của Nhánh I được giới hạn ở lục địa châu Á cận nhiệt đới hoặc vùng nhiệt đới đến cận nhiệt đới châu Úc (Nhánh IV), trong khi họ hàng sống trên đất chỉ giới hạn ở các vùng nhiệt đới Indo-Malaysia (Nhánh II/III) và châu Úc (Nhánh II) của Quần đảo Ấn-Úc. Nhánh V (vùng nhiệt đới Indo-Malaysia và châu Úc) và Nhánh VI với các đại diện đầu tiên ở dãy Himalaya, Nhật Bản và khắp châu Úc, đại diện cho giai đoạn đa dạng hóa thứ hai ở *Hoya*. Tác giả đã đưa ra giả thuyết rằng *Hoya* có nguồn gốc từ vùng nhiệt đới đến cận nhiệt đới Ấn Độ-Miến Điện/Himalaya và đặc điểm chung là phụ sinh có liên quan đến sự khởi đầu của khí hậu gió mùa trong quá trình nâng cao của dãy Himalaya [164].

Năm 2018, Tan và cs đã giải trình tự và chú thích đầy đủ đối với hệ gene lục lạp của *H. pottsii* Traill (MH678667), *H. liangii* Tsiang (MH678666), cùng với 17 trình tự hệ gene lục lạp hoàn chỉnh đã được công bố trên NCBI (*Cynanchum wilfordii* NC_029459.1, *Asclepias nivea* NC_022431.1, *Oncinotis tenuiloba* NC_025657.1, *Beaumontia murtonii* NC_037934.1, *Streptoechites chinensis* NC_037931.1, *Stripecoma peltigera* NC_037933.1, *Pentalinon luteum* NC_025658.1, *Echites umbellatus* NC_025655.1, *Pachypodium baronii* NC_037929.1, *Alafia barteri* NC_037930.1, *Nerium oleander* NC_025656.1, *Carissa macrocarpa* NC_033354.1, *Rhazya stricta* NC_024292.1, *Tonduzia stenophylla* NC_037935.1, *Catharanthus roseus* NC_021423.1, *Vinca Major* NC_037928.1, *Genetiana veitchiorum* NC_037985.1), các tác giả đã sử dụng phần mềm RAxML để xây dựng sơ đồ phát sinh chủng loại của 19 loài, trong đó 2 loài thuộc chi *Hoya* được xếp ở cùng một nhóm và cho kết quả giá trị bootstrap 100 [146]. Wei và cs (2020) trên cơ sở công bố giải trình tự hệ gene lục lạp hoàn chỉnh của *H. carnosa*, kết hợp với 18 hệ gene lục lạp hoàn chỉnh đã được công bố trên NCBI của các chi trong họ Apocynaceae và 2 loài được sử dụng ở nhóm ngoài gồm *Genetiana straminea* (Genetianaceae), *Coffea arabica* (Rubiaceae) đã xây dựng được cây phát sinh chủng loại tối đa, cây phát sinh chủng loại cho thấy *H. Carnosa*, *H. liangii* và *H. pottsii* là 3 loài trong chi *Hoya* tạo thành một nhánh duy nhất [167].

Nghiên cứu của Rodda và cs công bố 20 hệ gene lục lạp được giải trình tự của các loài trong nhóm *Hoya*, chỉ ra mối liên hệ giữa chi *Hoya* với các chi khác trong tông Mardenieae. Cơ sở của nghiên cứu phát sinh chủng loại này dựa trên tập hợp hai bộ dữ liệu locus lục lạp *trnT-trnL-trnF* từ 110 loài; 3 locus gene lục lạp (*trnT-trnL-trnF*, *atpB-psbA* spacer và *matK*) và 2 locus gene nhân từ 54 loài. Kết quả nghiên cứu đã xây dựng được sơ đồ phát sinh chủng loại tối đa và cho thấy *Hoya* có quan hệ gần gũi với ngành Anatropanthus, Dischidia và Oreosparte; có 3 loài mới có hình thái giống Oreosparte và giống với các loài thuộc chi *Hoya* được đề xuất là chi mới *PapuaHoya* [118], [119].

Odago và cs (2022) đã giải trình tự 31 hệ gene lục lạp của các loài thuộc chi *Hoya* của Trung Quốc bằng cách sử dụng dữ liệu quét hệ gene và thực hiện nhiều phân tích để tìm hiểu sự biến đổi trong hệ gene nhằm phát hiện những vị trí tiềm năng sử dụng trong phân tích phát sinh chủng loại. Nghiên cứu đã sàng lọc các điểm nóng tiềm năng (vùng DNA có độ đa dạng nucleotide cao, Pi) *trnT-trnL-trnF*, *psba-trnH*, *trnG-UCC*, *ndhF*, *ycf1*, *matK*, *rps16* và *accD* có thể được sử dụng làm chỉ thị phân tử DNA để hỗ trợ xác định loài. Sử dụng khả năng tối đa (ML) và suy luận Bayes (BI) xây dựng được sơ đồ phát sinh chủng loại. Hệ gene lục lạp mới được lắp ráp cho thấy đặc điểm cấu trúc của *Hoya* và *Dischidia* điển hình gồm 3 vùng, trong đó vùng SSC ngày càng được thu hẹp và 2 vùng IR được mở rộng. Kích thước hệ gene dao động từ 175.404 bp ở *H. lacunosa* đến 179.069 bp ở *H. ariadna*. Vùng LSC dao động từ 80.795 bp (*H. liangii*) đến 92.072 bp (*H._sp2_ZCF6006*). Các vùng IR được mở rộng mạnh, vùng SSC được thu hẹp lại còn một gene duy nhất là *ndhF*. Nghiên cứu đã xác định được 235 điểm lặp lại phân bố dài (LDR) và 10 điểm có độ phân kỳ cao trong 31 hệ gene lục lạp của chi *Hoya*, có thể được sử dụng làm mã vạch DNA để nhận dạng loài. Phát sinh chủng loại cho thấy *Clemensiella* tồn tại như một chi riêng biệt. *H. ignorata* được coi là họ hàng với các loài trong Nhánh VI (Nhánh số VI chỉ bao gồm các loài trong chi *Hoya*). Nghiên cứu này đã đánh giá được những lợi ích của việc sử dụng dữ liệu hệ gene lục lạp đã được giải trình tự để nghiên cứu tiến hóa và phân tích mối quan hệ phát sinh chủng loại [104].

1.2.4. Nhận xét chung

Trong bối cảnh khoa học và công nghệ sinh học phát triển mạnh mẽ, các chỉ thị phân tử đã trở thành công cụ then chốt trong nghiên cứu phát sinh chủng loại và phân loại sinh vật. Việc phân tích dựa trên các chỉ thị DNA (gene, vùng gene, vùng đệm, intron) hoặc toàn bộ hệ gene, đặc biệt là hệ gene lục lạp, cho phép tiếp cận một cách khách quan và có độ tin cậy cao hơn trong việc làm sáng tỏ lịch sử tiến hóa của sinh vật nói chung, cũng như các loài thuộc chi *Hoya* và các chi trong họ Apocynaceae nói riêng. Đồng thời, việc xác định các vùng DNA có tiềm năng sử dụng làm chỉ thị mã vạch phân tử (DNA barcoding) phục vụ định danh loài đang trở thành một hướng nghiên cứu quan trọng, có ý nghĩa khoa học và thực tiễn sâu sắc.

Tuy nhiên, cho đến nay, số lượng loài thuộc chi *Hoya* đã được giải trình tự hệ gene lục lạp vẫn còn hạn chế so với mức độ đa dạng rất cao của chi này, dẫn đến cơ sở dữ liệu phục vụ phân tích phát sinh chủng loại còn chưa đầy đủ và tính đại diện chưa cao. Điều này phần nào làm giảm độ tin cậy và độ phân giải của các cây phát sinh chủng loại đã công bố. Mặc dù các công cụ và phương pháp tin sinh học hiện đại ngày càng hoàn thiện, cho phép phân tích dữ liệu ở quy mô lớn và với độ chính xác cao, nhưng hiệu quả của các phân tích này vẫn phụ thuộc chặt chẽ vào sự đầy đủ và chất lượng của dữ liệu hệ gene nền.

Đáng chú ý, cho đến thời điểm hiện tại, chưa có công trình nào công bố kết quả phân tích phát sinh chủng loại dựa trên hệ gene lục lạp của các loài *Hoya* phân bố tại Việt Nam, đồng thời cũng chưa có dữ liệu giải trình tự hệ gene lục lạp hoàn chỉnh của *H. verticillata* var. *verticillata*. Đây là một khoảng trống dữ liệu quan trọng, hạn chế khả năng đánh giá đầy đủ vị trí phân loại và mối quan hệ tiến hóa của nhóm này trong họ Apocynaceae.

Vì vậy, việc tiến hành giải trình tự hệ gene lục lạp của *H. verticillata* var. *verticillata*, kết hợp phân tích phát sinh chủng loại và sàng lọc các vùng DNA có tiềm năng làm chỉ thị phân tử phục vụ định danh loài, là hết sức cần thiết và mang tính cấp thiết trong bối cảnh hiện nay. Nghiên cứu này không chỉ góp phần bổ sung dữ liệu di

truyền cho chi *Hoya* mà còn cung cấp cơ sở khoa học vững chắc cho công tác phân loại, bảo tồn và khai thác bền vững nguồn tài nguyên thực vật ở Việt Nam.

1.3. THÀNH PHẦN HÓA HỌC, HOẠT TÍNH SINH HỌC CỦA MỘT SỐ LOÀI THUỘC CHI HOYA

1.3.1. Các hợp chất hóa học ở thực vật hạt kín và chi *Hoya*

Các nghiên cứu về hợp chất thứ cấp trong thực vật đã phát hiện nhiều nhóm chất hữu cơ có hoạt tính sinh học. Alkaloid là nhóm hợp chất chứa nitrogen phổ biến ở khoảng 20% loài thực vật, thường liên quan đến hoạt tính sinh học mạnh. Terpenes và terpenoid là nhóm chất chuyển hóa lớn nhất, có cấu trúc đa dạng và vai trò sinh học quan trọng như tiền chất steroid, truyền tín hiệu và chuyển hóa tế bào. Một số terpen như vitamin A còn có giá trị chữa bệnh [74], [173].

Phenolic là nhóm hợp chất thứ cấp dồi dào trong thực vật, bao gồm phenol đơn giản, phenolic acid, flavonoid, tannin và lignin [69]. Trong đó, flavonoid là nhóm nổi bật với hơn 5.000 loại, chia thành sáu phân nhóm chính: Flavonol, anthocyanidin, isoflavone, flavan-3-ols, flavones và flavanone [41].

Glycoside là các hợp chất có trong nhiều họ thực vật, với hoạt tính sinh học đa dạng. Một số loại phổ biến gồm glucosinolate, glycoside cyanogene, glycoalkaloid, saponin và cardenolide [129].

Saponin là nhóm hợp chất phân cực lưỡng tính, xuất hiện rộng rãi ở hơn 90 họ thực vật [58], phân bố trong rễ, lá và hạt. Về cấu trúc, chúng gồm aglycone steroid hoặc triterpenoid liên kết với chuỗi đường (sapogenin) [126]. Saponin có hoạt tính sinh học mạnh như chống viêm, hạ cholesterol, điều hòa miễn dịch, nhưng cũng có thể ảnh hưởng tiêu cực đến vi sinh vật dạ cỏ và sức khỏe vật nuôi [5], [145]. Hàm lượng saponin biến đổi theo giống, tuổi cây, trạng thái sinh lý và vị trí địa lý [58].

Các nghiên cứu hóa học trên các loài *Hoya* cho thấy sự đa dạng của các hợp chất hữu cơ, đặc biệt là triterpenoid và sterol. Các dẫn xuất phổ biến gồm: β -amyrin và α -amyrin được tìm thấy ở *H. multiflora*, *H. diversifolia*, *H. cumingiana*, *H. paziae*; β -

amyrin cinnamate có ở *H. wayetii*, *H. buotii*, *H. madulidii*; lupeol xuất hiện trong nhiều loài như *H. mindorensis*, *H. multiflora*, *H. cagayanensis*; α -amyrin cinnamate ở *H. buotii*, *H. madulidii*; lupenone trong *H. mindorensis*, *H. pubicalyx*, *H. cagayanensis*; squalene có trong *H. mindorensis*, *H. buotii*, *H. madulidii* và nhóm sterol như β -sitosterol, stigmasterol có mặt ở hầu hết các loài kể trên [95].

Thành phần terpenoid chiếm 39,51% trong hoa *H. cagayanensis*, *H. lacunosa*, *H. coriacea* và là thành phần chính tạo hương thơm [18]. Từ sáp của lá già của *H. lacunosa* phân lập được β -amyrin, taraxerol, α -amyrin và taraxasterol [12]. *H. naumanii* chứa khoảng 50% triterpenol (β -amyrin, lupeol, α -amyrin), 25% n-hydrocarbon (n-C31, n-C33) và 20% acid hữu cơ có nguồn gốc từ dihydronyctanthic acid [14]. Các nghiên cứu chỉ ra ở loài *H. australis* có chứa australinal A–D (triterpene mới), δ -amyrin, β -amyrin, lupeol và α -amyrin [15].

Từ thân *H. carnososa* đã phân lập các oligosaccharide, ở loài *H. parasitica* phân lập được hoyasterone (androstanoid), acid 15-bulnesolic (sesquiterpene) và hợp chất phenolic đặc trưng [122].

Phân tích dịch chiết rễ, thân, lá loài *H. mindorensis* thu được lupeol, lupenone (trong rễ); squalene (trong lá), botulin (trong thân) [35]. *H. multiflora* chứa nhiều hợp chất như lupeol, α/β -amyrin, bauerenol, lutein, squalene, β -sitosterol và stigmasterol [36]. *H. buotii* chứa taraxerone, taraxerol, hỗn hợp α/β -amyrin cinnamate, β -sitosterol, stigmasterol với thành phần các chất thay đổi tùy theo cơ quan được dùng để phân tích [37]. Từ dịch chiết của loài *H. cagayanensis* phân lập được lupeol, lupenone, dihydrocanaric acid, các hydrocarbon bão hòa và 2-hydroxyethyl benzoate [113]. Chiết xuất lá *H. madulidii* có hỗn hợp lupeol cinnamate, amyrin cinnamate (2:2:1), lupeol, β -sitosterol và stigmasterol (5:1), cùng các hydrocarbon không bão hòa [23].

Dịch chiết *H. meliflua* thu được squalene và hỗn hợp β -amyrin, α -amyrin, lupeol, oleanone, ursenone, lupenone, các dẫn xuất cinnamate và hỗn hợp sterol với tỉ lệ đặc trưng [95].

Thành phần hóa học của các loài trong chi *Hoya* đã được nhiều tác giả nghiên cứu và các hợp chất hữu cơ thu được từ các bộ phận của các loài trong chi *Hoya* được mô tả tại Bảng 1.1 [121].

Bảng 1.1. Thành phần hóa học của một số loài trong chi *Hoya*. [121]

Loài/bộ phận	Thành phần hóa học	Tác giả
<i>H. buotii</i> Kloppenb (Hoa)	Taraxerone, taraxerol, β -sitosterol, stigmasterol, α -amyrin, β -amyrin cinnamate, squalene	(Ebajo và cs 2015b)
<i>H. carnos</i> a (Lá)	Flavonoids, tannins, phenols, and alkaloids	(Rahayu 2019)
<i>H. diversifolia</i> Blume (Lá)	β -amyrin cinnamate, squalene, β -sitosterol, α -amyrin, β - amyrin, lupeol	(Panajon và cs 2016)
<i>H. mindorensis</i> (Lá)	Squalene, β -sitosterol	(Ebajo và cs 2014)
<i>H. multiflora</i> Blume (Lá)	Bauerenol, squalene, lutein, β -sitosterol, stigmasterol	(Ebajo và cs 2015c)
<i>H. parasitica</i> (Lá)	Reducing sugar, tannin, flavonoid, gum, saponin	(Reza và cs 2007)
<i>H. wayetii</i> Kloppenb (Lá)	Triglycerides, chlorophyll a, β -sitosterol, stigmasterol	(Ebajo và cs 2015a)
<i>H. wightii</i> ssp. <i>Palniensis</i> (Lá)	Alkaloids, phenolics, steroids, flavonoids, triterpenoids, saponins, xanthoproteins	(Lakshmi và cs 2014)
<i>H. diversifolia</i> Blume (Thân)	Taraxerol, lupeol cinnamate, stigmasterol	(Panajon và cs 2016)
<i>H. kerrii</i> (Thân)	Lupeol acetate, taraxerol, lupeol, sitosterol, stigmasterol, p- hydroxybenzaldehyde, scopoletin, lariciresinol	(Seeka và cs 2017)
<i>H. mindorensis</i> (Thân)	Betulin	(Ebajo và cs 2014)
<i>H. multiflora</i> Blume (Thân)	Lupeol, α -amyrin, β -amyrin, lupeol acetate, α -amyrin acetate, β -amyrin acetate	(Ebajo và cs 2015c)

<i>H. parasitica</i> (Thân)	Tannin, flavonoid, alkaloid, steroid	(Khatun và cs 2014)
<i>H. wayetii</i> <i>Kloppenb</i> (Thân)	β -amyrin cinnamate, taraxerol	(Ebajo và cs 2015a)
<i>H. wightii</i> ssp. <i>Palniensis</i> (Thân)	Alkaloids, phenolics, steroids, flavonoids, triterpenoids, xanthoproteins	(Lakshmi và cs 2014)
<i>H. mindorensis</i> (Rễ)	Lupenone, lupeol	(Ebajo và cs 2014)

1.3.2. Hoạt tính sinh học của thực vật và một số loài *Hoya*

Các hợp chất có hoạt tính sinh học từ thực vật ngày càng được quan tâm nhờ tiềm năng trong phòng ngừa và điều trị nhiều loại bệnh như ung thư, viêm mạn tính, rối loạn chuyển hóa [135]. Chúng đóng vai trò quan trọng trong việc khám phá và phát triển thuốc thông qua quy trình sàng lọc, đánh giá sinh học – hóa học và thử nghiệm lâm sàng [25], [111].

Trong điều trị ung thư, nhiều hợp chất có nguồn gốc từ thực vật như curcumin, vincristine, vinblastine, paclitaxel và lycopene đã chứng minh hiệu quả thông qua các cơ chế như ức chế tăng sinh, cảm ứng apoptosis và chống oxy hóa [68], [140]. Một số chất có khả năng tăng cường hiệu quả của hóa trị đồng thời giảm tác dụng phụ, bảo vệ tế bào bình thường trong khi ức chế tế bào ung thư.

Với các rối loạn thần kinh, phytochemical thể hiện khả năng bảo vệ thần kinh thông qua chống oxy hóa, chống viêm và ngăn ngừa chết tế bào thần kinh [53]. Nhiều sản phẩm tự nhiên từ *Rhodiola rosea*, *Ginkgo biloba* và *Morinda lucida* đã cho thấy hiệu quả trong cải thiện triệu chứng của các bệnh thoái hóa thần kinh như Alzheimer hoặc Parkinson.

Trong lĩnh vực tim mạch, các hợp chất như flavonoid, phenolic và carotenoid có tác dụng hạ huyết áp, chống viêm, chống oxy hóa và điều hòa lipid máu. Isoflavone, resveratrol, và quercetin đã được ghi nhận giúp giảm nguy cơ mắc bệnh tim mạch [133]. Tuy nhiên, khả năng hòa tan và hấp thụ kém của chúng là hạn chế lớn trong điều trị bệnh

tim mạch. Công nghệ nano như liposome, hạt nano polymer hoặc nano bạc đang được nghiên cứu để tăng cường hiệu quả sinh học và tính ổn định của các hợp chất hóa học thực vật ứng dụng trong y dược [133].

Về khả năng chống viêm và điều hòa miễn dịch, polyphenol từ trà xanh, curcumin, resveratrol và boswellic acid có khả năng ức chế quá trình viêm, điều chỉnh biểu hiện cytokine và gene miễn dịch [161], [149]. Các chất này được nghiên cứu như lựa chọn thay thế cho thuốc điều trị viêm khớp, viêm ruột mạn tính.

Trong bối cảnh kháng kháng sinh gia tăng, các chất tự nhiên như alkaloid, flavonoid và terpenoid có tiềm năng lớn trong điều trị nhiễm khuẩn, nấm và virus [62], [85], [127], [138]. Một số triterpenoid cho thấy khả năng ức chế coronavirus, trong khi flavonoid đã chứng minh hiệu quả kháng SARS-CoV-2. Nhiều hướng tiếp cận như hệ dẫn nano đang được nghiên cứu nhằm cải thiện độ bền và khả năng hấp thụ tế bào của các hợp chất này [138].

Đối với rối loạn chuyển hóa và nội tiết, các hợp chất trong trái cây, rau củ và ngũ cốc như epigallocatechin gallate từ *Camellia sinensis* (trà xanh) có hiệu quả trong điều trị đái tháo đường và béo phì [125]. Hợp chất ở các loài như *Citrullus colocynthis* (Dưa đắng), *Silybum marianum* (Cúc gai) và *Aloe vera* (Lô hội) có tác dụng hạ đường huyết. Ngoài ra, Nhân sâm, Mướp đắng, Ma hoàng cũng được nghiên cứu trong kiểm soát cân nặng và hội chứng chuyển hóa. Tuy nhiên, việc ứng dụng hiệu quả đòi hỏi các nghiên cứu chuẩn hóa và thử nghiệm lâm sàng nghiêm ngặt [51].

Trong điều trị bệnh gan, các liệu pháp truyền thống còn nhiều hạn chế và gây tác dụng phụ. Một số hợp chất thực vật như silymarin, andrographolide, curcumin và glycyrrhizin đã cho thấy tác dụng bảo vệ gan thông qua việc giảm stress oxy hóa và ức chế con đường viêm [93], [128].

Phytochemical cũng có tiềm năng lớn trong lĩnh vực mỹ phẩm và dược phẩm nhờ hoạt tính kháng oxy hóa, kháng viêm và kháng khuẩn. Các chất như aloin, ginsenoside và curcumin giúp trung hòa gốc tự do, giảm mất nước, tăng cường khả năng chống nắng

và làm chậm quá trình lão hóa [86], [92]. Chúng có thể kích hoạt cơ chế tự thực bào, sửa chữa DNA và điều hòa chuyển hóa tế bào.

Nghiên cứu của Pandey và cs (2004) đã chỉ ra rằng các loài *Hoya* chứa lipid, flavanol, sterol, triterpenes, sesquiterpenes và disaccharides, với hoạt tính chống sốt rét, gây độc tế bào, chống dị ứng và điều hòa miễn dịch. Các hợp chất từ *Hoya* được sử dụng điều trị hen suyễn mạn tính, các bệnh viêm nhiễm, đồng thời ứng dụng trong vai trò như thuốc trừ sâu [108].

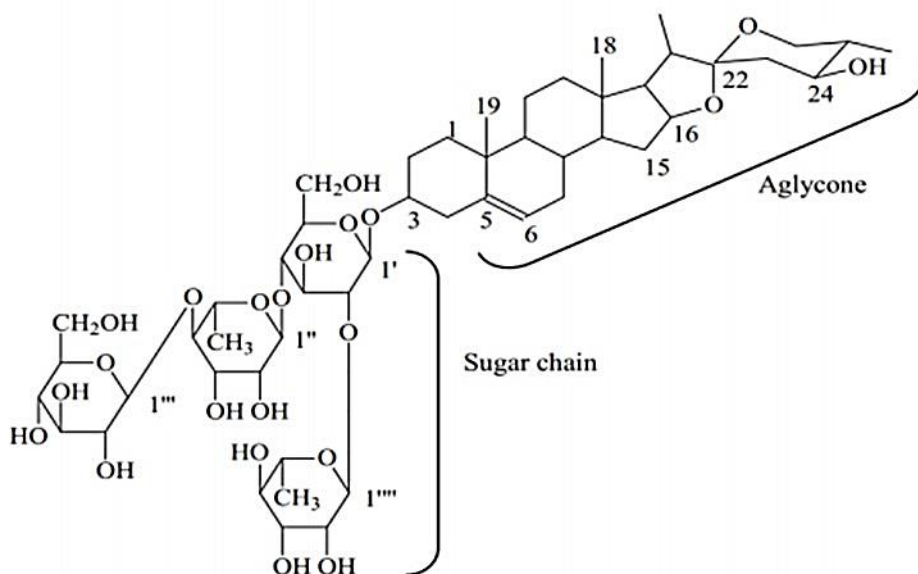
Tania và cs (2016) đã chiết xuất lá Lan tai cáo bằng methanol, ethanol và chloroform và đánh giá hoạt tính sinh học ban đầu. Kết quả cho thấy cả ba dung môi đều tạo ra chiết xuất có giá trị IC_{50} tương đương ascorbic acid chuẩn. Chất chiết chloroform thể hiện khả năng khử mạnh, chống tan máu; trong khi chiết ethanol ổn định màng hồng cầu. Các thử nghiệm *in vivo* cho thấy chiết methanol và ethanol làm tăng nhu động ruột rõ rệt [148].

Nguyen, H. D. (2023), dựa trên nghiên cứu y học cổ truyền về cách sử dụng Lan tai cáo điều trị các bệnh như thấp khớp, vàng da, rối loạn tiết niệu, sốt, táo bón và đau, đã thực hiện chiết xuất lá bằng n-hexane và ethyl acetate, đồng thời đánh giá hoạt tính thông qua thần kinh trung ương thông qua các thí nghiệm trên động vật. Kết quả cho thấy các chiết xuất này có đặc tính an thần, chống lo âu và ức chế thần kinh trung ương. Tuy nhiên, chúng cũng thể hiện độc tính tế bào, đặt ra yêu cầu nghiên cứu sâu hơn để xác định hợp chất hoạt động chính xác, từ đó phát triển các sản phẩm an toàn và hiệu quả [99].

1.3.3. Saponin

Saponin là những glycoside có khối lượng phân tử lớn, thường gặp trong nhiều loài thực vật. Saponin có tính hoạt động bề mặt mạnh, khi hoà tan vào nước có tác dụng làm giảm sức căng bề mặt của dung dịch, tạo nhiều bọt, có tính chất phá huyết, độc đối với động vật máu lạnh nhất là đối với cá, tạo thành phức với cholesterol, có vị hắc và làm hắt hơi mạnh. Phân tử saponin của nó có hai phần: Phần không phải saccharide của

phần tử saponin được gọi là genin hay sapogenin (hay aglycone) và phần đường được liên kết với nhau bằng liên kết glycosidic được thể hiện trong Hình 1.3 [90].

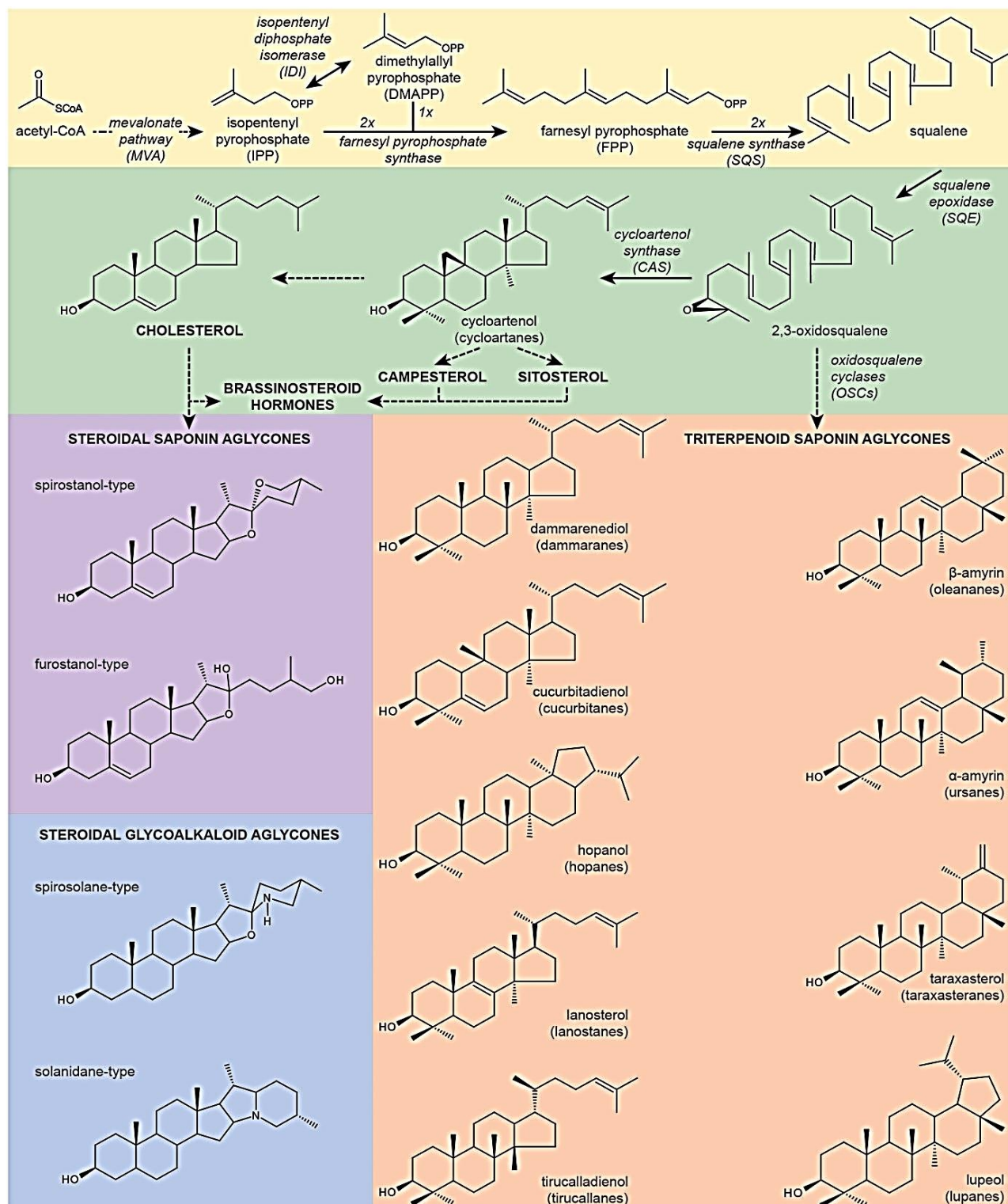


Hình 1.3. Cấu trúc của saponin (Moghimpour và Handali, 2015) [90]

1.3.3.1. Saponin thực vật

Con đường sinh tổng hợp saponin (được thể hiện trong Hình 1.4) bắt đầu với acetyl-CoA, chất này thông qua con đường mevalonate được chuyển hóa thành tiền chất squalene mạch thẳng 30 carbon (màu vàng). Squalene được oxy hóa thành 2,3-oxidosqualene, chất này được vòng hóa thành aglycone cycloartane, tiền chất chuyên biệt cho quá trình sinh tổng hợp sterol ở thực vật (màu xanh lá cây). Cholesterol là tiền chất chuyên biệt cho quá trình tổng hợp saponin steroid (màu tím) và glycoalkaloid steroid (màu xanh lam). Các saponin triterpenoid (màu cam) được tổng hợp từ 2,3-oxidosqualene được vòng hóa thành các aglycone triterpene chuyên biệt. Tên enzyme

và con đường sinh tổng hợp được in nghiêng. Mũi tên nét đứt biểu thị nhiều bước trong con đường sinh tổng hợp [91].



Hình 1.4. Sinh tổng hợp saponin thực vật [91]

Một số nghiên cứu đã chứng minh hệ gene lục lạp tổng hợp các protein tham gia vào quá trình điều hòa sản xuất isoprenoid trong lục lạp là cơ sở quan trọng để tổng hợp các hợp chất và hormone thiết yếu liên quan đến quang hợp [46].

Sapogenin của saponin triterpenoid là một dẫn xuất triterpenoid bao gồm 30 nguyên tử carbon và khung cơ bản gồm 6 đơn vị isopren. Sapogenin - saponin triterpenoid được chia thành triterpenoid tetracycle và pentacycle. Các loại triterpenoid 4 vòng bao gồm lanostane, euphane, loại dadamane [101], ankan cucurbitacin, loại cycloxyane, meliacanes, ...; các loại triterpenoid 5 vòng bao gồm loại oleanane [78], ursane, lupane, friedelanes và các loại khác. Sapogenin của saponin steroid là dẫn xuất steroid thường chứa 27 nguyên tử carbon và có khung cơ bản là spirostane và furostane. Saponin steroid được chia thành các loại spirostanol, isospirostanol, furostanol và spirostanol biến dạng. Saponin được phân bố rộng rãi trong tự nhiên và thường được tìm thấy trong rễ, thân, lá, hoa (quả) và hạt của thực vật có hoa trên cạn [5]. Saponin triterpenoid phân bố chủ yếu ở các họ Leguminosae (Đậu), Araliaceae (Nhân sâm), Umbelliferae (Hoa tán), Polygalaceae (Nghê) và các họ thực vật khác. Saponin steroid chủ yếu được tìm thấy trong các họ thực vật một lá mầm, như các họ thực vật Dioscoreaceae (Họ Củ nâu), Liliaceae (Họ Loa kèn) và Scrophulariaceae (Họ Huyền sâm). Saponin là một trong những thành phần chính của nhiều loại thuốc thảo dược Trung Quốc, chẳng hạn như nhân sâm, xương cựa, bupleurum, *Ophiopogon japonicus*, notoginseng và cam thảo [112].

1.3.3.2. Hoạt tính sinh học của saponin thực vật

Saponin được nghiên cứu rộng rãi vì hoạt tính sinh học đa dạng của chúng, bao gồm tác dụng kháng khuẩn, kháng nấm, kháng virus, gây độc tế bào ung thư, chống viêm, chống oxy hóa, hạ cholesterol và điều hòa miễn dịch. Nghiên cứu gần đây đã nhấn mạnh tiềm năng của chúng như là sự thay thế hữu hiệu cho các chất hoạt động bề mặt tổng hợp và tác nhân dược lý [64]. Saponin đã cho thấy nhiều tiềm năng trong việc nâng cao hiệu quả của thuốc chống ung thư và kháng sinh khi được sử dụng trong các liệu pháp kết hợp. Sự đa dạng về cấu trúc của chúng góp phần tạo ra nhiều ứng dụng công nghiệp khác nhau trong lĩnh vực dược phẩm, mỹ phẩm và thực phẩm. Tuy nhiên, vẫn

cần có những nghiên cứu sâu về cấu trúc, hoạt tính sinh học, tiềm năng trong phòng và trị bệnh, nâng cao sức khỏe, kéo dài tuổi thọ [64].

Hoạt tính kháng khuẩn: Saponin chiết xuất từ loài *Glycyrrhiza glabra* (Cam thảo nhẵn) và *Quillaja saponaria* có tác dụng kháng khuẩn đối với cả vi khuẩn Gram dương và Gram âm. Saponin của loài *Glycyrrhiza glabra* có khả năng ức chế mạnh đối với *Escherichia coli*, *Staphylococcus aureus*, *Pseudomonas aeruginosa* và *Proteus mirabilis* [90]. Glabridin được phân lập từ *Glycyrrhiza glabra* có hoạt tính kháng vi khuẩn lao (chủng *Mycobacterium tuberculosis*). Saponin của loài *Quillaja saponaria* có tác dụng kháng khuẩn, tuy nhiên ở nồng độ thấp lại cho thấy khả năng tăng cường sinh trưởng của *E. coli* [10] và ở nồng độ cao lại gây độc tính đối với tế bào nhân thực. Những phát hiện này chứng minh rằng saponin là hợp chất kháng khuẩn tự nhiên. Cơ chế tác động của saponin dựa trên cấu trúc lưỡng tính, dẫn tới sự tương tác của chúng với các thành phần cấu trúc của màng như cholesterol và phospholipid [84] từ đó có thể phá vỡ màng tế bào và gây nên sự phân hủy tế bào. Các nghiên cứu đã chỉ ra rằng saponin có thể làm nhiễu loạn màng sinh học, làm giảm sức căng bề mặt trong dung dịch nước và tạo thành micelle. Saponin của loài *Quillaja saponaria* có thể gây ra sự phân hủy tế bào ở các chủng nấm men công nghiệp [21]. Saponin oleanane cho thấy có tác dụng thẩm thấu khác nhau trên màng tế bào và lysosome ở nồng độ thấp, trong khi đó chúng làm tan máu ở nồng độ cao [48]. Hoạt tính phá vỡ màng tế bào của saponin có tiềm năng ứng dụng trong điều trị ung thư, tuy nhiên cần có thêm những nghiên cứu để hạn chế tác dụng làm tan máu của chúng [84].

Hoạt tính kháng nấm: Saponin của loài *Medicago sativa* (Cỏ linh lăng) có hoạt tính kháng nhiều loài nấm, trong đó có loài *Candida albicans*. Một số nghiên cứu đã chỉ ra rằng các phân đoạn giàu saponin từ *Medicago sativa* ức chế sự phát triển của *Candida albicans* thông qua ức chế hình thành chồi, sự phát triển của sợi nấm, sự bám dính và hình thành màng [123]. Hoạt tính của saponin chiết xuất từ loài *Medicago sativa* còn được chứng minh trên các loại nấm men và vi khuẩn khác, đặc biệt là các loài Gram dương.

Hoạt tính kháng virus: Nghiên cứu chỉ ra rằng saponin có thể ức chế sự nhân lên của các loại virus như HIV (Human Immunodeficiency Virus) và HSV (Herpes Simplex

Virus). Chúng can thiệp vào cấu trúc lớp vỏ của virus, ngăn chặn virus xâm nhập vào tế bào vật chủ. Saponin triterpenoid cho thấy hoạt tính kháng virus đối với HSV-1 bằng cách ức chế tổng hợp DNA và tổng hợp protein vỏ. Tiềm năng kháng virus của saponin còn được chứng minh trên các loại virus khác như SARS-CoV, rotavirus và poliovirus [130]. Trong thời kỳ hậu COVID-19, saponin cho thấy triển vọng như các tác nhân kháng virus SARS-CoV-2 [89]. Những phát hiện này làm nổi bật vai trò của saponin như các tác nhân kháng virus tự nhiên có nguồn gốc từ thực vật với tác dụng phụ tối thiểu, là cơ sở để thực hiện những nghiên cứu sâu hơn nữa để đưa saponin vào các liệu pháp điều trị kháng virus.

Gây độc tế bào ung thư: Saponin từ *Panax ginseng* (Nhân sâm) gây ra hiện tượng chết theo chương trình ở nhiều dòng tế bào ung thư khác nhau, bao gồm ung thư vú, ruột kết và phổi thông qua nhiều con đường truyền tín hiệu như PI3K/Akt, NF- κ B, EGFR [71], [171]. Chiết xuất nhân sâm đã cho thấy hoạt động chống tăng sinh của các tế bào ung thư phổi A549, điều chỉnh các gene liên quan đến truyền tín hiệu tế bào, chuyển hóa và điều hòa hình thành khung xương tế bào [70]. Saponin, bao gồm cả saponin từ nhân sâm, ức chế sự phát triển của tế bào khối u bằng cách gây ra sự ngừng chu kỳ tế bào và hiện tượng chết theo chương trình, như vậy nghiên cứu cho thấy tiềm năng của saponin nhân sâm trong liệu pháp điều trị ung thư [71]. Nghiên cứu về chiết xuất các hợp chất từ loài *Dioscorea villosa* (khoai mỡ dại) đã cho thấy tác dụng chống lại các tế bào ung thư vú. Một glycoside spirostanol được phân lập từ khoai mỡ dại đã chứng minh hoạt tính gây độc đối với nhiều dòng tế bào ung thư [59]. Chiết xuất saponin từ khoai mỡ dại ở nồng độ dưới mức gây chết có gây độc tế bào, kích thích các gene miễn dịch và ức chế các gene liên quan đến chu kỳ tế bào [40], [88]. Dioscin là một saponin steroid từ khoai mỡ dại có tiềm năng chống các tế bào ung thư vú xâm lấn bằng cách tăng cường biểu hiện GATA3, thúc đẩy các đặc điểm biểu hiện thành tế bào biểu mô và làm giảm xâm lấn tế bào [11]. Các nghiên cứu chỉ ra rằng saponin có hoạt tính kháng tế bào ung thư thông qua nhiều cơ chế khác nhau, mặc dù các nghiên cứu này chứng minh tiềm năng của saponin trong điều trị ung thư, nhưng điều quan trọng là những nghiên cứu này mới chủ yếu tập trung vào các tế bào ung thư vú nên cần phải có thêm nhiều nghiên cứu để đánh giá tác dụng cụ thể của nó đối với các dòng tế bào ung thư khác nhau [39].

Chống viêm: Saponin từ loài *Glycyrrhiza uralensis* (Cam thảo) đã được chứng minh có tác dụng chống viêm bằng cách ức chế sản xuất cytokine tiền viêm (gây viêm). Các nghiên cứu đã chỉ ra rằng chiết xuất từ loài *Glycyrrhiza uralensis* tác động đến đại thực bào, ức chế hoạt động giải phóng các chất trung gian gây viêm như NO, IL-1 β và TNF- α [83]. Saponin tổng số từ loài *Glycyrrhiza uralensis* ức chế các enzyme chính trong con đường chuyển hóa arachidonic acid, làm giảm tổng hợp PGE2 [83] hoặc ức chế hoạt hóa NLRP3 inflammasome, ức chế sản xuất caspase-1 và IL-1 β trong các đại thực bào và mô mỡ. Nhiều thành phần khác nhau trong saponin của loài *Glycyrrhiza uralensis* bao gồm isoliquiritigenin, isoliquiritin và liquiritigenin, đã được phát hiện có tác dụng ức chế sản xuất NO thông qua cơ chế tác động ức chế biểu hiện enzyme xúc tác chuyển hóa sản xuất NO có thể gây ra ở các tế bào gan được xử lý bằng IL-1 β [147]. Trong khi glycyrrhizin là thành phần chính của *Glycyrrhiza uralensis* cho thấy tác dụng chống viêm, các thành phần phụ lại có tác dụng cao hơn trong việc ức chế các chất trung gian gây viêm [147]. Saponin của loài *Hedera helix* (cây Thường xuân) có đặc tính chống viêm trong cả mô hình viêm cấp tính và mãn tính [141]. Lá của cây chứa nhiều hợp chất có hoạt tính sinh học khác nhau, bao gồm: Saponin, flavonoid và terpenoid [132], [8]. Các chất này có hoạt tính đa dạng như chống viêm, giảm đau, long đờm và kháng khuẩn [8]. Saponin, đặc biệt là α -hederin, β -hederin và hederacoside C là những chất chính trong điều trị viêm [150]. Các thử nghiệm của các sản phẩm chiết xuất saponin thô và chiết xuất saponin tinh khiết của loài *Hedera helix* thể hiện tác dụng chống viêm mạnh mẽ, trong đó chiết xuất thô thể hiện hoạt động chống viêm cấp tính mạnh hơn và chiết xuất tinh khiết thể hiện tác dụng chống viêm mãn tính lớn hơn [141]. Những phát hiện này hỗ trợ cho việc sử dụng tiềm năng chiết xuất *Hedera helix* trong điều trị các tình trạng viêm và bệnh về đường hô hấp. Những nghiên cứu gần đây đã làm rõ các đặc tính chống viêm của saponin liên quan đến các cơ chế như điều hòa con đường NF- κ B và ức chế enzyme COX/LOX [170].

Chống oxy hóa: Chiết xuất từ loài *Tribulus terrestris* (Bạch tật lê) đã được chứng minh hoạt động chống oxy hóa đáng kể trong các thử nghiệm *in vitro*. Các nghiên cứu đã chỉ ra rằng cả sản phẩm chiết xuất ethanol thô và chiết phân đoạn giàu saponin đều thể hiện đặc tính chống lại các gốc tự do như DPPH, ABTS, nitric oxide, hydroxyl và hydrogen

peroxide. Phân đoạn giàu saponin thường thể hiện hoạt tính chống oxy hóa mạnh hơn so với chiết xuất thô. Các chế phẩm từ loài *Tribulus terrestris* cũng cho tác dụng chống oxy hóa, khử sắt và ức chế quá trình peroxy hóa lipid. Các nghiên cứu *in vivo* trên chuột đã chứng minh được mức độ tăng lên của glutathione, ascorbic acid và superoxide dismutase tăng lên [61].

Hạ cholesterol: Saponin có tác dụng hạ cholesterol. Nghiên cứu đã chỉ ra rằng các saponin chiết xuất từ rễ như karaya, quillaja và sapindin làm giảm đáng kể nồng độ cholesterol trong huyết thanh của chuột cống và chuột nhắt [6], [159]. Dựa trên cơ chế tác động dịch chuyển các phân tử cholesterol khỏi các micelle muối mật đã tạo ra các cholesterol bị kết tủa (cholesterol kết tủa không thể đi qua lớp niêm mạc của ruột, nên không xảy ra sự hấp thụ chất này ở ruột) là cơ sở giúp cơ thể kiểm soát tình trạng tăng cholesterol máu, giảm nguy cơ mắc các bệnh tim mạch [159].

Điều hòa miễn dịch: Một số nghiên cứu đã chứng minh saponin có khả năng tăng cường phản ứng miễn dịch [27], [32]. Các hợp chất này có thể kích thích cả phản ứng miễn dịch Th1 và sản xuất tế bào lympho T gây độc, đây là cơ sở cho nhiều loại vaccine khác nhau, bao gồm vaccine chống lại các tác nhân gây bệnh nội bào cũng như vaccine điều trị ung thư [142]. Tác dụng hỗ trợ của saponin liên quan đến cấu trúc hóa học của chúng, nhờ đó chúng kích thích khả năng miễn dịch bẩm sinh thông qua tiết cytokine. Các nghiên cứu gần đây về saponin đã cho thấy về tiềm năng sử dụng qua đường hô hấp giúp làm tăng cường phản ứng miễn dịch, nghiên cứu này mở ra hướng phát triển các loại vaccine qua đường hô hấp [27].

Như vậy, saponin thực vật có cấu trúc đa dạng, có hoạt tính sinh học phong phú và có tiềm năng trong y học như: Chống viêm, kháng tế bào ung thư, kháng virus, điều hòa miễn dịch ... Mặc dù có nhiều triển vọng trong việc phát triển các loại thuốc dựa trên hoạt tính của saponin thì vẫn còn những lo ngại về độc tính của chúng, do đó cần tiếp tục có những nghiên cứu đầy đủ hơn nữa về cơ chế tác động của saponin và cách thức tối ưu hóa tiềm năng điều trị của chúng.

1.3.3.3. Phân lập saponin từ một số loài Hoya

Sắc ký lỏng cao áp (HPLC) và sắc ký khí khối phổ (GC-MS) là các kỹ thuật được sử dụng rộng rãi để xác định và phân tích saponin trong chiết xuất các hợp chất hóa học

của thực vật. Sắc kí lỏng hiệu năng cao kết hợp với các cột pha đảo và các máy dò khác nhau, có hiệu quả trong việc tách và tinh chế các hỗn hợp saponin phức tạp [94]. Phương pháp quang phổ khối, bao gồm ion hóa phun điện tử (ESI-MS) và giải hấp thụ ion hóa bằng laser được hỗ trợ bởi ma trận hấp thụ năng lượng (MALDI), đã được sử dụng rộng rãi để xác định đặc điểm cấu trúc của saponin [126]. Sắc kí lỏng và phổ khối song song (LC/MS/MS) là công cụ phân tích hiệu quả để xác định các saponin đồng phân. Các phương pháp này tạo nên các công cụ phân tích mạnh mẽ để sàng lọc nhanh và xác định cấu trúc của saponin trong chiết xuất thực vật [30]. Các phương pháp phân tích được kết hợp với nhau, chẳng hạn như GC-MS, LC-MS/MS và LC-SPE-NMR/MS, đã được sử dụng để sàng lọc và so sánh cấu hình saponin ở các loài thực vật khác nhau, cho phép xác định các hợp chất có hoạt tính sinh học [67]. Những kỹ thuật này đã nâng cao đáng kể hiệu quả phân tích saponin, khắc phục những thách thức đặt ra bởi sự đa dạng và phức tạp trong cấu trúc của chúng

Các hợp chất đã được xác định ở *H. lacunosa*, bao gồm triterpenol tự do và hợp chất este hóa như β -amyrin, α -amyrin và lupeol [103]. Lá *H. lacunosa* chứa một số C-glycosylflavonoid chính được xác định là: 6-C-arabinosyl 8-C-glucosylapigenein (isoschaftoside), 6-C-glucosyl 8-C-arabinosylapigenein (schaftoside-de) và 6,8-di-C-arabinosylapigenein. Trong nhựa mủ, triterpenol được tìm thấy có ở cả dạng tự do và ở dạng este, trong đó β -amyrin, α -amyrin, lupeol, 24-methylenecycloartanol, obtusifoliol và cycloeucalenol được xác định có trong thành phần các rượu tự do.

Trong nhựa mủ của loài *H. bella* người ta đã tìm thấy một số triterpenol, cả ở dạng tự do và dạng este. Các triterpenol chính là β -amyrin, cycloartenol, lupeol và isobauerenol với lượng nhỏ; một số este được xác định bao gồm cinnamic acid, propionic acid và isovaleric acid. Phân lập chất hữu cơ từ lá có sự xuất hiện của dạng acyl hóa flavonol glycoside và ferulic acid tự do; khi lá già nồng độ các chất này giảm xuống [13].

Thành phần mủ và phenolic trong lá, đặc biệt là flavon loại apigenein, được cho là đặc trưng cho từng loài và có khả năng hữu ích trong việc xác định phân loại các loài *Hoya* [16]. Nghiên cứu về khả năng sử dụng thành phần của latex triterpenes và phenolics từ lá của 25 loài *Hoya* đã được công bố cho thấy, cả phenolics (chủ yếu là

flavon loại apigenein) và latex triterpenes đều có tỷ lệ đặc trưng đối với các loài được nghiên cứu. Do tính ổn định, đặc trưng về thành phần các chất này nên có thể dùng chúng làm chất đánh dấu bổ sung trong việc xác định loài *Hoya*.

Fumiko và cs (1999) đã công bố kết quả nghiên cứu về 11 hợp chất được phân lập từ hạt của *H. carnososa*, trong số này có 6 hợp chất là saponin, các tác giả đã xác định được cấu trúc của 19-acetoxypurpurogenein II và xác định đây là loại chất mới. Nghiên cứu cũng đề xuất cấu trúc của 20 tetraosides và pentaosides, được đặt tên từ *H. carnosides* A đến T, 3 hợp chất còn lại đã được biết đến từ các nghiên cứu khoa học trước đây [3].

Các nghiên cứu về saponin của các loài thực vật thuộc chi *Hoya* cho thấy có triterpen và sterol (ở nhiều loài khác nhau). Lá *H. meliflua* chứa hỗn hợp gồm β -amyrin, α -amyrin và lupeol, cùng với các dẫn xuất và sterol của chúng là β -sitosterol và stirysterol [95]. Nghiên cứu này là một phần trong nghiên cứu về thành phần hóa học của các loài trong chi *Hoya* bản địa của Philippine. Các hợp chất được phân lập bằng phương pháp sắc ký silica gel và được xác định bằng kỹ thuật quang phổ cộng hưởng từ hạt nhân NMR. Kết quả phân tích hóa học dịch chiết điện di chloromethane từ lá cây *H. meliflua* thu được squalene và hỗn hợp β -amyrin, α -amyrin và lupeol với tỷ lệ khoảng 1:1:0,25; oleanon, ursenon và lupenon theo tỷ lệ khoảng 1:1:0,3; β -amyrin cinnamate (3a), α -amyrin cinnamate và lupenyl cinnamate theo tỷ lệ khoảng 0,5:0,3:1; β -sitosterol và stirysterol theo tỷ lệ khoảng 5:1.

Cinnamate triterpenoid, acetate và palmitate là các chất được phân lập từ lá *H. kerrii* [75]. Hai loại cinnamate triterpenoid mới, đó là 3β -O-D-friedours-11 α ,12 α -epoxy-14-enyl cinnamate và 3β -O-29-nor-20-oxolupanyl cinnamate, bổ sung vào dữ liệu 13 hợp chất đã được phân lập từ chiết xuất hexan lá của *H. kerrii*. Việc làm sáng tỏ cấu trúc hóa học của các hợp chất được thực hiện bằng phương pháp quang phổ. Một số cinnamate triterpenoid, cinnamate lupeol được cho là có đặc tính chống viêm mạnh. Sự hiện diện của các este 3β -O-D-friedours-11 α ,12 α -epoxy-14-enyl và 3β -O-D-friedoolean-11 α ,12 α -epoxy-14-enyl chỉ được tìm thấy ở một số loài trong họ Apocynaceae. Các nghiên cứu chứng minh saponin triterpenoid có tiềm năng trong điều trị các bệnh rối loạn chuyển hóa và bệnh về mạch máu [38].

Basir và cs (2022) trên cơ sở sử dụng kỹ thuật vi chiết pha rắn HS-SPME và sắc kí khí kết hợp với phổ khối GC-MS để đánh giá khả năng bay hơi các hợp chất của hoa *H. cagayanensis*, *H. lacunosa* và *H. coriacea*. Nghiên cứu góp phần làm sáng tỏ quá trình sinh tổng hợp các chất chuyển hóa thứ cấp, đặc biệt là terpenoid trong các loài *Hoya*. Nghiên cứu xác định được 23 chất chuyển hóa có ở *H. cagayanensis*, 14 chất có ở *H. lacunosa* và 36 chất ở *H. coriacea*. Nhiều chất chuyển hóa được xác định là riêng biệt cho mỗi loài, với 15 chất ở *H. cagayanensis*, 6 chất ở *H. lacunosa* và có 28 chất ở *H. coriacea*. Chỉ có 6 chất được phát hiện là β -ocimene, (Z)-3-hexen-1-ol, 1-octen-3-ol, butanoic acid (Z)-2-methyl-3-hexenyl ester, (Z)-butyric acid- 3-hexenyl ester và methyl salicylate chuyển hóa từ terpenoid, rượu, este và nhóm benzenoid được phát hiện ở cả 3 loài được nghiên cứu [18].

Đối với loài Lan tai cáo (*H. verticillata*) có rất ít nghiên cứu về định tính thành phần hóa học và đánh giá hoạt tính sinh học. Reza và cs (2007) nghiên cứu dịch chiết ethanol từ lá *H. verticillata* chứa flavonoid, đường khử, tannin và saponin. Dịch chiết từ lá có hoạt tính kháng khuẩn, chống lại cả vi khuẩn gram dương và gram âm, đồng thời có khả năng ức chế khối u ở chuột [114]. Ngoài ra, dịch chiết từ lá của Lan tai cáo cũng được đánh giá có hoạt tính chống oxy hóa. Như vậy có thể thấy rằng, đối với loài Lan tai cáo việc phân tích thành phần hóa học và phát hiện các hợp chất thứ cấp có giá trị dược học, đánh giá hoạt tính sinh học của các hợp chất đó là vấn đề cần được tiếp tục làm sáng tỏ. Nghiên cứu về saponin thuộc chi *Hoya* đã xác định được một số saponin mới. Từ dữ liệu quang phổ và phép đo phổ khối (HR-ESI-MS), cấu trúc của 3 glycoside steroid được phân lập từ cây *H. parasitica* (Wall. ex Hornem.) Wight đã được làm sáng tỏ. Đó là các saponin $3\beta,4\alpha$ -dihydroxy- 5β -spirost-(25)27-en- 1β -yl *O*- α -L-arabinopyranoside; 3β -hydroxy- 5β -spirost-25(27)- en- 1β -yl *O*- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-xylopyranoside; và (23S,24S,25S)- $3\beta,4\alpha,23,24$ -tetrahydroxy- 5β -spirostan - 1β -yl *O*- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-fucopyranoside [155].

1.3.4. Mô phỏng tương tác phân tử và động học phân tử

Tương tác phân tử (Molecular Docking) là mô hình tính toán của các phức hợp, được hình thành từ sự ghép nối dẫn đến tương tác của hai hoặc nhiều phân tử. Phương

pháp này cho phép dự đoán cấu trúc ba chiều của các chất cộng hợp, dựa trên các đặc tính liên kết của phối tử tham gia và các phân tử mục tiêu. Mô phỏng tương tác phân tử tạo ra các cấu trúc ứng viên có thể khác nhau, được xếp hạng và nhóm lại với nhau bằng cách sử dụng hàm tính điểm trong phần mềm của công cụ ghép nối phân tử. Mô phỏng tương tác dự đoán cấu trúc đồng dạng được ghép nối tối ưu dựa trên tổng năng lượng của hệ thống [7].

Dựa trên quá trình sắp xếp của phân tử nghiên cứu với một phối tử hoặc phân tử thụ thể để tạo ra một phức hợp ổn định (dự đoán ái lực liên kết, cường độ liên kết giữa phối tử và protein). Ái lực và hoạt động của một chất hóa học được dự đoán bởi tương tác chất nghiên cứu-thụ thể, điều này quan trọng trong việc khám phá và thiết kế thuốc. Việc tìm kiếm và phát triển các loại thuốc mới là một quá trình phức tạp và nhiều khó khăn. Các loại thuốc mới được khám phá bằng cách sử dụng phương pháp *in silico*, thiết kế thuốc dựa trên máy tính cần được tăng cường sử dụng để đẩy nhanh quá trình khám phá thuốc. Mô phỏng tương tác phân tử cho phép dự đoán hướng ưa thích mà hai phân tử sẽ thực hiện khi chúng tương tác với nhau để tạo thành một hợp chất ổn định theo mô hình hóa phân tử. Có thể dự đoán mức độ tham gia hoặc ái lực liên kết giữa hai phân tử bằng cách sử dụng hướng quay được chọn trong quá trình truyền tín hiệu, tương tác giữa các thành phần tương tự về mặt hóa học, chẳng hạn như protein, peptide, nucleic acid, carbohydrate và lipid. Do đó những ghép nối hữu ích để có thể dự đoán được cả hiệu lực và bản chất của tín hiệu. Do khả năng dự đoán cách các phối tử phân tử nhỏ sẽ gắn vào vị trí liên kết mục tiêu lý tưởng, tương tác phân tử là một trong những kỹ thuật được sử dụng nhiều nhất trong thiết kế thuốc dựa trên cấu trúc. Xác định hiệu suất liên kết đóng vai trò quyết định cho việc lập kế hoạch phát triển thuốc và để hiểu ý nghĩa của các quá trình sinh hóa cơ bản [34].

Rebeiro và cs nghiên cứu acetylcholinesterase và các alkaloid được phân lập từ cây *Tetrapteryx mucronata*, kết quả ghép nối cho kết quả tích cực, là cơ sở cho những đề xuất thực nghiệm về khám phá và phát triển dược phẩm. Nghiên cứu đóng góp tích cực về lý thuyết liên quan đến các sản phẩm tự nhiên này trong việc phòng chống các bệnh thoái hóa thần kinh [115].

Nghiên cứu về chứng đau do viêm là một trong những lĩnh vực nghiên cứu được phát triển và mở rộng nhanh nhất trong những năm gần đây. Các nghiên cứu từ phòng thí nghiệm đã chứng minh tiềm năng điều chỉnh cơn đau mãn tính của loài *Phyllanthus amarus* và khả năng tương tác của chúng với nhiều chất trung gian gây viêm liên quan đến các enzyme như COX-2 và PGE synthase, các cytokine như TNF-alpha và IL-1 beta và với thụ thể NMDA (Các thụ thể NMDA là các protein dẫn truyền các xung động thần kinh liên quan đến trí nhớ, ngôn ngữ và thần kinh tự chủ như nhịp tim, nhịp thở). Các chất trung gian gây viêm đóng vai trò quan trọng trong chứng tăng đau viêm mãn tính và tác động đến quá trình điều chỉnh, các chất này được lựa chọn và đánh giá tương tác của chúng với 86 thành phần thực vật có cấu trúc đa dạng được xác định từ loài *Phyllanthus amarus*. Phân tích tương tác các protein mục tiêu với các hợp chất hữu cơ thực vật được thực hiện bằng phần mềm VLifeMDS (Molecular Design Suite một công cụ phần mềm được sử dụng trong hóa học tính toán, khám phá thuốc và mô hình hóa phân tử do VLife Sciences phát triển). Thông qua các điểm ghép nối và phân tích tương tác của các hợp chất với các protein mục tiêu cho thấy các hợp chất quan trọng như lupeol, phyllanthin, hypopyllanthin, corilagin, epicatechin và hầu hết các hợp chất khác có khả năng liên kết với nhiều mục tiêu liên quan đến chứng tăng đau do viêm [28].

Với mục tiêu phát triển các liệu pháp an toàn hơn với ít tác dụng phụ hơn trong điều trị ung thư, nghiên cứu về các chất chiết xuất từ thực vật có ý nghĩa rất lớn. Các loài *Erythroxylum* (thuộc chi *Coca*) là nguồn cung cấp rộng rãi các hợp chất hữu cơ thực vật khác nhau, đặc biệt là terpenoid. Thực hiện các nghiên cứu mô hình tương tác một số terpenoid trên các mục tiêu đã được xác nhận như EGFR, VEGFR, CDK, protein tubulin (được biểu hiện quá mức trong nhiều loại ung thư) qua đó ước tính các đặc tính dược động học của các phân tử này bằng các kỹ thuật *in silico*. Các hợp chất cho thấy ái lực liên kết tốt và xác định được hằng số ức chế đối với tất cả các mục tiêu nghiên cứu [156].

Động học phân tử (MD-Molecular dynamics) là phương pháp mô phỏng bằng máy tính, giúp khám phá các nguyên lý của nhiều quá trình khác nhau thông qua mô phỏng chuyển động vĩnh cửu thực tế của các nguyên tử do lực giữa chúng gây ra. Trong

phân tích động học phân tử, các lực được thể hiện thông qua các hàm phân tích và các tham số liên quan, thường được gọi là trường lực. Tác động của các lực lên chuyển động của các nguyên tử được xác định bởi định luật thứ hai của Newton. Phương pháp nghiên cứu này có ưu điểm nổi bật đó là có thể theo dõi chính xác quỹ đạo của các nguyên tử tương tác với nhau nhờ sự biến đổi không gian của chúng theo thời gian từ cấu hình xác định ban đầu. Hệ thống có thể được mô tả và phân tích quỹ đạo tương tác bằng các biểu thức toán thống kê, liên hệ các trạng thái vi mô với các đặc tính vĩ mô có liên quan đến vấn đề đang xem xét. Động học phân tử là một công cụ tính toán được sử dụng thành công để khám phá các nguyên lý của nhiều quá trình khác nhau ở cấp độ nguyên tử. Nhờ những tiến bộ gần đây trong mô phỏng máy tính, động học phân tử có thể đạt đến thang thời gian mili giây, do đó đây là một phương pháp thay thế quan trọng bổ sung cho các phương pháp thực nghiệm trong sinh học. Hiệu quả tính toán của các thuật toán động học phân tử tiếp tục được nghiên cứu, bổ sung và nâng cấp. Một số phương pháp chuyên biệt bắt nguồn từ động học phân tử (MD) cho phép nghiên cứu nhiều quá trình khác nhau của các phân tử sinh học như: Động học phân tử trao đổi bản sao (REMD) động học phân tử mục tiêu (TMD), động học phân tử định hướng (SMD) [169].

Mô phỏng động học phân tử có vai trò quan trọng trong nhiều ngành khoa học, như vật lý, hóa học, khoa học vật liệu và sinh học. Phương pháp này đã chứng minh được sự liên quan sâu sắc trong việc mô tả chi tiết các hệ thống phân tử sinh học, bổ sung dữ liệu cho thực nghiệm, tối ưu hóa thiết kế thực nghiệm, giúp dự đoán các đặc tính có liên quan cho các hệ thống hóa học đất liền hoặc khó xử lý bằng thực nghiệm, được sử dụng trong các giai đoạn đầu của quá trình thiết kế và phát triển thuốc. Ứng dụng của mô phỏng động học phân tử vào việc mô tả các hệ thống phân tử sinh học rất rộng, bao gồm mô tả cấu trúc và tổ chức màng, tính thấm của màng, tương tác lipid với protein, tương tác lipid với thuốc, tương tác protein với hợp chất, động học của protein [45].

1.3.5. Nhận xét chung

Tổng quan tài liệu cho thấy các nghiên cứu hiện nay về thực vật chủ yếu tập trung vào việc xác định, định lượng và phân tích thành phần hóa học, cũng như đánh giá một số hoạt tính sinh học cơ bản của các hợp chất tự nhiên. Số lượng loài thực vật được khai

thác nhằm nghiên cứu thành phần hóa học và tiềm năng ứng dụng ngày càng gia tăng. Trong đó, các hợp chất hữu cơ có nguồn gốc thực vật nói chung và từ các loài thuộc chi *Hoya* nói riêng đã được nghiên cứu ở những mức độ khác nhau. Tuy nhiên, các nghiên cứu chuyên sâu và có hệ thống về hoạt tính sinh học của các loài thuộc chi *Hoya* vẫn còn khá hạn chế, đặc biệt là đối với nhóm hợp chất saponin, một nhóm chất có nhiều tiềm năng dược lý quan trọng.

Đáng chú ý, cho đến nay hầu như chưa có công trình nào tập trung vào việc phân lập, xác định cấu trúc và đánh giá hoạt tính sinh học của saponin từ Lan tai cáo (*H. verticillata* var. *verticillata*). Do đó, việc nghiên cứu phân lập và đặc trưng hóa các hợp chất, đặc biệt là saponin, từ loài này, kết hợp với đánh giá bước đầu về hoạt tính sinh học của chúng, là một hướng nghiên cứu có ý nghĩa khoa học và thực tiễn rõ rệt, đồng thời góp phần khai thác hợp lý tiềm năng nguồn dược liệu bản địa.

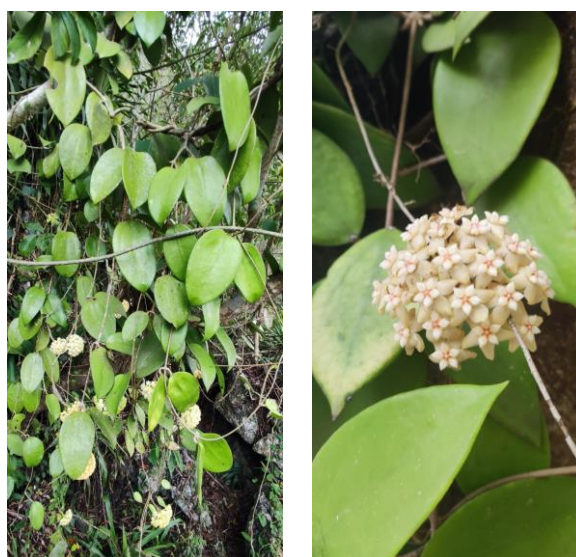
Bên cạnh đó, cho đến nay chưa có công trình nào công bố việc áp dụng các phương pháp *in silico* như mô phỏng docking phân tử (molecular docking) và mô phỏng động học phân tử (molecular dynamics) để đánh giá cơ chế tương tác giữa các hợp chất hữu cơ của *H. verticillata* var. *verticillata*, đặc biệt là saponin, với các protein đích liên quan đến quá trình chết theo chương trình của tế bào (apoptosis), một cơ chế sinh học then chốt gắn liền với sự hình thành và phát triển ung thư. Việc tích hợp các phương pháp tính toán hiện đại này với nghiên cứu thực nghiệm sẽ góp phần làm sáng tỏ cơ chế tác động ở mức độ phân tử, đồng thời nâng cao giá trị khoa học và định hướng ứng dụng của các kết quả nghiên cứu thu được.

Chương 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. VẬT LIỆU NGHIÊN CỨU

2.1.1. Vật liệu thực vật

Mẫu *H. verticillata* var. *verticillata* (Mẫu chứng từ HOYP202302TQ1) được thu thập tại xã Sơn Dương, tỉnh Tuyên Quang, Việt Nam vào tháng 2 năm 2023 tại tọa độ địa lí 21°41'04" N, 105°29'03" E và Mẫu chứng từ HOYP202305TN được thu thập tại khu vực Hang Trai, xã Quang Sơn, tỉnh Thái Nguyên, Việt Nam, ở tọa độ 21°43'34" N, 105°50'52" E. Hai mẫu HOYP202302TQ1 và HOYP202305TN được xác định thuộc loài *H. verticillata* var. *verticillata*. Việc định danh mẫu *H. verticillata* var. *verticillata* được thực hiện tại Bộ môn Thực vật học bởi PGS.TS. Sỹ Danh Thường và sau đó mẫu vật được lưu giữ tại Phòng Tiêu bản thực vật của Khoa Sinh học, Trường Đại học Sư phạm, Đại học Thái Nguyên. Lá Lan tai cáo được sử dụng để chiết rút cao chiết, phân lập saponin, tách DNA tổng số. Việc thu thập mẫu vật tuân thủ theo các quy định có liên quan, các hướng dẫn và luật pháp của Việt Nam và quốc tế.



A

B

Hình 2.1. Mẫu cây Lan tai cáo (*H. verticillata* var. *verticillata*). A: Cây Lan tai cáo mọc trong môi trường tự nhiên tại khu vực Hang Trai, xã Quang Sơn, tỉnh Thái Nguyên, Việt Nam; B: Cảnh Lan tai cáo mang hoa trên giá thể.

2.1.2. Các chủng vi khuẩn kiểm định

Vi khuẩn kiểm định sử dụng trong nghiên cứu để xác định hoạt tính kháng khuẩn gồm: *Escherichia coli* (ATCC25922), *Staphylococcus aureus* (ATCC13709), *Pseudomonas aeruginosa* (ATCC15442), *Bacillus subtilis* (ATCC6633).

Các chủng vi khuẩn do Bảo tàng chủng giống vi sinh vật, Bệnh viện Đa khoa Trung ương Thái Nguyên và Bộ môn Vi sinh, Trường Đại học Y - Dược, Đại học Thái Nguyên cung cấp. Các chủng vi khuẩn được nuôi cấy và lưu giữ tại Khoa Sinh học, Trường Đại học Sư phạm - Đại học Thái Nguyên để tiến hành thử nghiệm. Vi khuẩn kiểm định sử dụng trong nghiên cứu đều là những loài gây bệnh ở người, trong đó vi khuẩn *S. aureus*, *B. subtilis* thuộc nhóm Gram dương, các chủng vi khuẩn còn lại thuộc nhóm Gram âm. Các thực nghiệm được tiến hành trong phòng thí nghiệm và tuân thủ đầy đủ các quy tắc an toàn trong phòng thí nghiệm.

2.1.3. Các dòng tế bào ung thư

Các dòng tế bào thử nghiệm được sử dụng trong nghiên cứu gồm: Tế bào ung thư phổi ở người (TA549); Tế bào ung thư vú ở người (MCF7); Tế bào ung thư gan ở người (HepG2). Các dòng tế bào do GS.TS. J. M. Pezzuto, Trường Đại học Long-Island, US và GS. Jeanette Maier, Trường Đại học Milan, Italia cung cấp, được lưu giữ tại Phòng Thử nghiệm sinh học, Viện Sinh học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

2.1.4. Dữ liệu

Sử dụng dữ liệu hệ gene lục lạp và trình tự các vùng DNA của hệ gene lục lạp từ các loài đã được công bố trên GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>); dữ liệu thông tin và cấu trúc protein trên Ngân hàng dữ liệu protein (<https://www.rcsb.org/>).

2.2. HÓA CHẤT, THIẾT BỊ

2.2.1. Hóa chất

DPPH (2,2-diphenyl-1-picrylhydrazyl), ascorbic acid (Vitamin C), DMSO (Dimethyl sulfoxide), SRB (Sulforhodamine B), Acetic acid; Enzyme α -glucosidase chiết xuất từ nấm men, p-nitrophenyl- α -D-glucopyranoside (pNPG), 4-Nitrophenol, Ellipticine, sulforhodamine B, trichloroacetic acid, acetic acid, đệm tris base, đệm phosphate, enzyme α -glucosidase (CAS No 9001-42-7, Sigma), p-Nitrophenyl- α -D-

glucopyranoside (CAS No 3767-28-0, Sigma), 4-Nitrophenol (CAS No 100-02-7, Sigma), dimethyl sulfoxide (CAS No 67-68-5), DMEM (Dulbecco's Modified Eagle Medium), MEME (Minimum Essential Medium with Eagle salt), L-glutamine, sodium pyruvate, NaHCO₃, penicillin/streptomycin, 10% FBS (Fetal Bovine Serum), Trypsin-EDTA (0.05%) do Hãng Sigma, Mỹ cung cấp.

Silica gel pha thường (230-400 mesh), silica gel pha đảo RP-18 (150 µm), Sephadex LH-20, nhựa trao đổi ion Diaion HP-20 (Nhật Bản).

2.2.2. Thiết bị

Thiết bị sử dụng trong thí nghiệm đều mới, hiện đại và có độ chính xác cao thuộc các phòng thí nghiệm của Khoa Sinh học, Trường Đại học Sư phạm - Đại học Thái Nguyên; Viện Sinh học thuộc Viện Hàn lâm Khoa học và Công nghệ Việt Nam; Trung tâm Nghiên cứu Vật liệu Sinh học Quốc tế (IBMRC), Viện Nghiên cứu Khoa học Sinh học và Công nghệ sinh học Hàn Quốc (KRIBB).

2.3. PHƯƠNG PHÁP NGHIÊN CỨU

2.3.1. Nhóm phương pháp nghiên cứu hệ gene lục lạp

2.3.1.1. Tách chiết DNA tổng số và giải trình tự hệ gene lục lạp

DNA tổng số được chiết xuất từ lá Lan tai cáo khô bằng phương pháp sử dụng CTAB (cetyl trimethylammonium bromide), CTAB là phương pháp được sử dụng rộng rãi để phân lập nucleic acid từ thực vật giàu polysaccharide [82]. Phương pháp đo huỳnh quang (Qubit, Invitrogen) và điện di trên gel được sử dụng để đánh giá chất lượng và độ tinh khiết của DNA. DNA tổng số đảm bảo chất lượng được sử dụng để phân lập hệ gene lục lạp, quy trình được thực hiện như sau, từ 100ng DNA tổng số được phân mảnh bằng sóng siêu âm trên thiết bị Qsonica 800 R2 tạo ra các đoạn 350 bp, các đoạn này được nối với bộ chuyển đổi của Illumina và được khuếch đại bằng PCR. Các đoạn có từ 500-600 bp là kích thước phù hợp để thiết lập thư viện mẫu. Thư viện mẫu được định lượng bằng thiết bị TapeStation 4200 (Agilent Technologies) và Bộ định lượng thư viện KAPA (KK4824, Kapa Biosystems). Các thư viện tinh khiết thu được được gắn với các đoạn mồi trên chip giải trình tự và được khuếch đại thành cụm phân tử giống hệt nhau (khuếch đại dòng vô tính), công nghệ giải trình tự sử dụng dNTP được gắn thuốc nhuộm huỳnh quang với mạch bổ sung nhờ DNA polymerase (máy giải trình tự Illumina

NovaSeq 6000), dNTP không liên kết sẽ bị rửa trôi. Máy giải trình tự Illumina dựa trên hình ảnh bước sóng phát xạ và cường độ huỳnh quang của nucleotide hợp nhất từ đó xác định trình tự của các đoạn mẫu.

2.3.1.2. Lắp ráp và chú thích hệ gene

Khoảng 5 GB dữ liệu thô thu được được lắp ráp để xây dựng một hệ gene hoàn toàn mới từ rất nhiều đoạn trình tự ngắn nhờ máy NOVOPlasty v4.3.1 [33], dựa trên trình tự tham chiếu là hệ gene lục lạp hoàn chỉnh của loài *H. angustifolia* (OL754669.1). Quá trình lắp ráp, căn chỉnh, chuẩn hóa dữ liệu hệ gene nhờ phần mềm Geneious Prime v. 2023.2.1 [66], BBDuk v. 38.84 và BBNorm v. 38.84 mà không cần sử dụng hệ gene tham chiếu. Một hệ gene lục lạp hoàn chỉnh của loài *H. verticillata* var. *verticillata* được tạo ra bằng cách sử dụng các vùng tiếp giáp thu được bằng hai phương pháp: Vùng mã hóa protein của các gene, các gene mã hóa rRNA và tRNA được xác định bằng công cụ tìm kiếm trong Geneious; các gene tRNA được chú thích lại trong máy chủ web tRNAscan-SE v. 2.0.12. Bản đồ hệ gene lục lạp dạng vòng được xây dựng bằng OrganellarGenomeDRAW (OGDRAW) v. 1.3.1 [52], [63]. Độ bao phủ của hệ gene lục lạp hoàn chỉnh là 1843,9×.

2.3.1.3. Phân tích hệ gene

Các yếu tố lặp lại trong hệ gene của *H. verticillata* var. *verticillata* được tìm thấy bằng hai phương pháp, các đoạn lặp lại trình tự đơn giản (SSR) được phát hiện bằng Phobos v. 3.3.12 [20], [80], [87], bao gồm đơn vị lặp lại cho mono-, di-, tri-, tetra- và penta-nucleotide; các đoạn lặp lại trình tự lớn (LSR) được nghiên cứu bằng REPuter [77].

Thu thập và sử dụng thêm 5 hệ gene lục lạp có sẵn của các loài *H. carnosa* (OK539590), *H. kerrii* (OL754673), *H. lyi* (MW719055), *H. meliflua* (OL754674) và *H. ovalifolia* (OL754666) từ GenBank trên NCBI để làm tham chiếu. Toàn bộ trình tự hệ gene lục lạp hoàn chỉnh của 6 loài thuộc chi *Hoya* đã được căn chỉnh và trực quan hóa bằng cách sử dụng căn chỉnh Mauve với thuật toán progressiveMauve. Sử dụng IRscope (irscope.shinyapps.io/irapp/) để hiển thị trực quan và so sánh các ranh giới của các vùng LSC, SSC và IR của sáu loài *Hoya* [9]. Thực hiện sự phân kỳ trình tự của 29 loài thuộc chi *Hoya* thông qua phân tích cửa sổ trượt tính toán (hệ số Pi) giữa các hệ gene lục lạp nhờ phần mềm phân tích đa hình trình tự DNA (DnaSP v. 6.12.03) [9],

[120]. Đối với phân tích sự phân kỳ thì trình tự được áp dụng có kích thước cửa sổ là 600 bp và kích thước bước là 200 bp.

2.3.2. Nhóm phương pháp phân tích phát sinh chủng loại

2.3.2.1. Phương pháp phân tích phát sinh chủng loại trình tự hệ gene lục lạp

Sơ đồ phát sinh chủng loại được xây dựng dựa trên 79 trình tự gene mã hóa protein từ 50 trình tự hệ gene lục lạp hoàn chỉnh của họ Apocynaceae, bao gồm 49 trình tự hệ gene lục lạp hoàn chỉnh của họ Apocynaceae từ cơ sở dữ liệu GenBank và của Lan tai cáo. Mỗi hệ gene được căn chỉnh bằng chương trình căn chỉnh MAFFT v. 7.490 [65] trong Geneious v. 2023.2.1 và được nối lại, bộ dữ liệu dài 75.302 bp được đưa vào jModeltest (mô hình phát sinh chủng loại) trong cổng thông tin khoa học CIPRES (phylo.org) để xác định mô hình phù hợp nhất [31], [87]. Mô hình phù hợp nhất là GTR+I+G, sơ đồ phát sinh loài với khả năng tối đa ML được xây dựng bằng cách sử dụng RAxML-HPC2 trên XSEDE trong Cổng thông tin khoa học CIPRES với hệ số bootstrap là 1000 [137] và được trực quan hóa bằng Geneious v. 2023.2.1.

2.3.2.2. Phương pháp phân tích phát sinh chủng loại dựa trên trình tự một số vùng DNA lục lạp

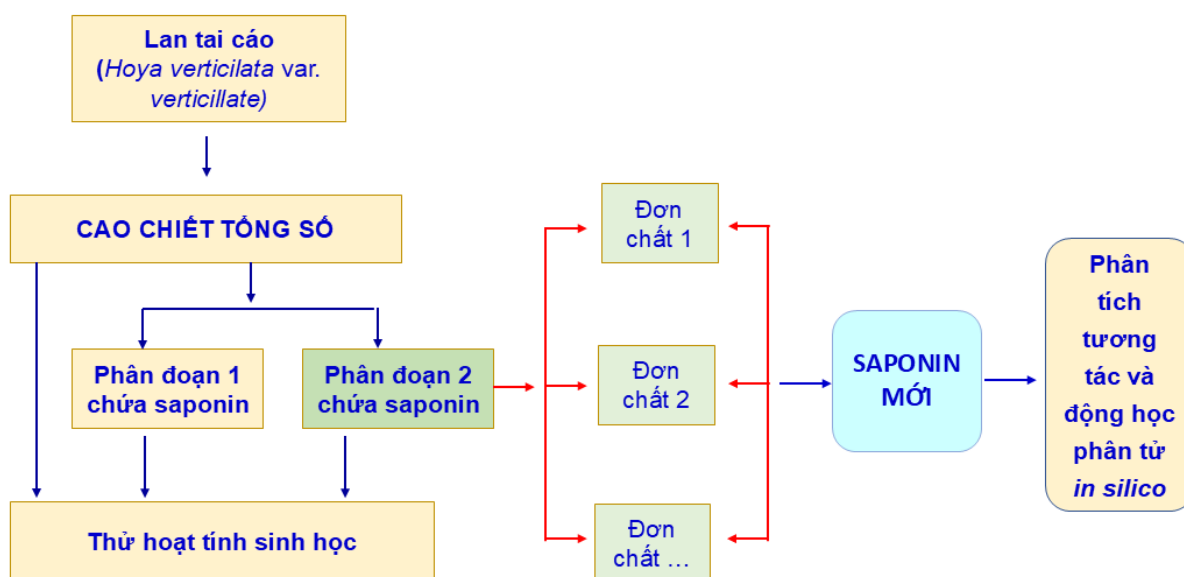
Dữ liệu hệ gene lục lạp của loài Lan tai cáo (*H. verticillata* var. *verticillata*) lưu giữ trên GenBank mang mã số OR475244.1. Trình tự nucleotide của các gene, vùng đệm của hệ gene lục lạp được nhận diện và tải từ dữ liệu hệ gene lục lạp của Lan tai cáo trên GenBank.

Sử dụng BLAST để phân tích sự tương đồng đối với mỗi trình tự vùng đệm của *H. verticillata* var. *verticillata* với các loài khác trên GenBank. Chọn các trình tự có tổng điểm và độ che phủ cao làm dữ liệu cho phân tích phát sinh chủng loại.

Phân tích phát sinh chủng loại dựa trên trình tự vùng đệm bằng MEGA11 [144], theo phương pháp Maximum Likelihood (ML) và mô hình Tamura-Nei [143], với hệ số bootstrap được suy ra từ 1000 lần lặp lại [44].

2.3.3. Nhóm phương pháp phân tích hoạt tính sinh học, phân lập, xác định chất

Sơ đồ thí nghiệm nghiên cứu phân lập, xác định chất saponin mới và phân tích hoạt tính sinh học được mô tả khái quát ở Hình 2.2.



Hình 2.2. Sơ đồ thí nghiệm phân lập và xác định saponin từ lá của Lan tai cáo

2.3.3.1. Định tính các hợp chất hữu cơ có trong Lan tai cáo

Mẫu lá Lan tai cáo sau khi được thu hái rửa sạch, loại bỏ mẫu hỏng, phơi/sấy khô tới khối lượng không đổi và nghiền thành bột mịn. Lấy 50 mg bột khô cho vào bình tam giác, sau đó thêm 100 ml dung môi EtOH/H₂O theo tỷ lệ 70:30, sử dụng phương pháp chiết có sử dụng sóng siêu âm hỗ trợ, chiết trong thời gian 30 phút ở nhiệt độ 50°C. Lọc dịch chiết thu được dịch chiết của lá Lan tai cáo [1].

Tiến hành xác định thành phần alkaloid bằng cách nhỏ lần lượt vào ống nghiệm có chứa 2ml dịch chiết lá Lan tai cáo từ 1-2 giọt thuốc thử Wagner, Dragendorff, Mayer; quan sát kết tủa các ống nghiệm so với đối chứng (không nhỏ thuốc thử).

Sử dụng phản ứng Shinoda để xác định thành phần flavonoid: Nhỏ vào ống nghiệm chứa dịch chiết lá Lan tai cáo 2-3 giọt HCl đậm đặc và thêm bột magnesium (Mg), kết quả dung dịch có màu từ vàng nhạt, nếu kết luận trong ống nghiệm có flavonoid.

Dung dịch sắt chloride (FeCl₃) được sử dụng để định tính tannin; tiến hành nhỏ vào thêm vào ống nghiệm vài giọt FeCl₃ 5%, nếu kết quả có kết tủa màu xanh (lá cây) đen, kết luận trong ống nghiệm có tannin.

Phản ứng Liebermann – Burchardt được sử dụng để định tính steroid; thêm vào ống nghiệm 1 mL chloroform, để nghiêng ống 45⁰ rồi cho từ từ từng giọt H₂SO₄ đậm đặc, nếu dung dịch trong ống nghiệm có màu xanh lá, chứng tỏ trong dịch chiết lá cây có steroid.

Định tính glycoside tim bằng phản ứng Keller-Kiliani, thêm 2 mL glacial acetic acid vào ống nghiệm, sau đó nhỏ từng giọt H₂SO₄ đậm đặc theo thành ống nghiệm. Quan sát có lớp chất lỏng màu xanh nhạt xuất hiện, kết luận ống nghiệm có glycoside tim.

Sử dụng NaOH để định tính coumarin, tiến hành thêm 3 mL NaOH 10% vào ống nghiệm chứa dịch Lan tai cáo, đun sôi ống nghiệm, để nguội. Quan sát thấy hiện tượng có kết tủa màu vàng chứng tỏ trong ống nghiệm có coumarin.

Định tính saponin bằng phản ứng tạo bọt, lắc mạnh ống nghiệm chứa dịch lá Lan tai cáo, sau đó để yên trong năm phút, quan sát thấy có bọt bền vững chứng tỏ trong dung dịch có chứa saponin.

2.3.3.2. Phương pháp tạo cao chiết

Lá Lan tai cáo được rửa sạch và sấy khô ở nhiệt độ 50°C (độ ẩm < 6%) và đem nghiền nhỏ thành bột. Bột lá được ngâm chiết bằng phương pháp có sự hỗ trợ siêu âm sử dụng dung môi EtOH/H₂O (tỷ lệ 70:30) ở điều kiện: 200 W, 60°C, 45 phút bằng máy siêu âm Elmasonic S10H (Elma, Thụy Sĩ), 500 mL mỗi lần × 2 lần. Sau khi lọc và bay hơi, thu được cao chiết tổng số.

2.3.3.3. Phương pháp xác định hoạt tính sinh học

Phương pháp thử hoạt tính kháng khuẩn

Khảo sát khả năng kháng vi sinh vật kiểm định của cao chiết tổng số bằng phương pháp khuếch tán đĩa thạch [24]. Các loài vi khuẩn kiểm định được hoạt hóa từ 4-8 giờ trong môi trường LB lỏng ở 28°C, lắc 200 vòng/phút trước khi tiến hành thử nghiệm. Hút 70 µL dịch nuôi mỗi loài vi khuẩn kiểm định đã được hoạt hóa và cấy trải lên đĩa chứa môi trường LB đặc. Dùng khoan nút chai tạo năm giếng có đường kính 1 cm trên đĩa thạch. Sau đó, hút 100 µL mẫu nghiên cứu ở các nồng độ 100; 200 và 400 µg/mL (cao chiết được pha trong DMSO 1%), kháng sinh amoxicillin nồng độ 200 µg/mL (đối chứng dương) và DMSO 1% (đối chứng âm) vào các giếng thạch. Các đĩa petri đã bổ sung mẫu nghiên cứu được đặt vào tủ lạnh 4°C khoảng 1-2 giờ rồi đặt vào tủ ấm nuôi ở

28°C từ 18-24 giờ. Đo đường kính vòng kháng khuẩn, chụp hình và ghi lại kết quả, mỗi thí nghiệm được lặp lại ba lần.

Khả năng kháng vi sinh vật kiểm định của cao chiết được xác định dựa trên đường kính vòng ức chế vi sinh vật của cao chiết. Công thức tính đường kính vòng ức chế vi sinh vật = $D - d$, trong đó D là đường kính vòng ức chế vi sinh vật, d là đường kính đục lỗ.

Phương pháp thử hoạt tính kháng oxy hóa

Hoạt tính chống oxy hóa của phân đoạn chứa saponin được tiến hành theo phương pháp của Abramović và cs (2018) [4] có chỉnh sửa cho phù hợp với điều kiện của phòng thí nghiệm, cụ thể như sau: Mẫu thử được tạo thành bằng cách pha dung dịch chiết trong methanol (100%) nồng độ 0,25 μM , sau đó được pha thành dải nồng độ từ 0,8-100 $\mu\text{g/mL}$ với nước cất khử ion. L-ascorbic acid được sử dụng như đối chứng và được pha thành dải nồng độ từ 0,8-100 $\mu\text{g/mL}$ với nước cất khử ion; DPPH pha trong methanol (100%) nồng độ 0,25 μM . Hút mẫu nghiên cứu (100 μL) đã pha ở các nồng độ vào các giếng của đĩa 96 giếng. Thêm dung dịch DPPH đã chuẩn bị ở trên vào các giếng đã có sẵn mẫu nghiên cứu (tỉ lệ 1:1), giếng đối chứng không có mẫu thử chỉ có 100 μL nước và 100 μL DPPH. Ủ ở nhiệt độ phòng trong 30 phút, sau đó xác định độ hấp thụ (OD) của dung dịch sau phản ứng tại bước sóng 517 nm. Khả năng trung hòa gốc oxy hóa tự do (Scavenging Activities - SA) sinh ra từ DPPH của mẫu thử được tính theo công thức sau:

$$\% \text{ SA} = (\text{OD}_{\text{đối chứng}} - \text{OD}_{\text{mẫu thử}}) * 100 / \text{OD}_{\text{đối chứng}} (\%)$$

Trong đó: $\text{OD}_{\text{đối chứng}} = \text{OD tại giếng không chứa chất thử} - \text{OD giếng trắng}$

$$\text{OD}_{\text{mẫu thử}} = \text{OD tại giếng chứa chất thử} - \text{OD giếng trắng}$$

Giá trị SC_{50} (Scavenging Concentration at 50% - nồng độ trung hòa được 50% gốc tự do của DPPH) sẽ được xác định nhờ vào phần mềm máy tính TableCurve 2Dv4.

Phương pháp thử hoạt tính ức chế enzyme α -glucosidase

Hoạt tính ức chế enzyme α -glucosidase của phân đoạn chứa saponin được thực hiện như sau: Chất thử được hòa tan trong DMSO 100% để có dung dịch gốc (20 mg/mL) và pha loãng trong dung dịch đệm phosphate 10 mM (pH 6.8) để tạo dải nồng độ. Tiếp theo, sử dụng 50 μL mẫu đã pha loãng được đưa vào các giếng của khay 96 giếng để chuẩn bị thử nghiệm; 20 μL α -glucosidase (0,5U/mL) và 130 μL đệm

phosphate 100 mM (pH 6.8) được thêm vào mỗi giếng, trộn đều và ủ ở 37°C trong 15 phút [154]. Như vậy, nồng độ mẫu thử đạt được cuối cùng trong giếng lần lượt là 500-100-20-4 µg/mL; 50 µL p-nitrophenyl- α -D-glucopyranoside (pNPG) nồng độ 5 mM được đưa tiếp vào từng giếng thí nghiệm rồi ủ tiếp ở 37°C trong 60 phút. Giếng thí nghiệm có mẫu thử, đệm đệm phosphate và pNPG được sử dụng làm đối chứng trắng (blank). Giếng có DMSO 10%, đệm phosphate, enzyme và pNPG được sử dụng làm đối chứng. Thí nghiệm được lặp lại 3 lần/khay để đảm bảo sự chính xác. Dừng thí nghiệm bằng cách thêm vào 80 µL Na₂CO₃ 0,2M và đo OD (Optical Density - đo mức độ hấp thụ ánh sáng của mẫu) ở bước sóng 405 nm bằng máy đo ELISA Plate Reader (Biotek). Khả năng ức chế α -glucosidase của mẫu thử được xác định theo công thức sau:

$$\% \text{ Ức chế} = 100 - [(OD_{\text{mẫu thử}} / OD_{\text{đối chứng}}) \times 100]$$

$$\text{Trong đó: } OD_{\text{đối chứng}} = OD_{\text{đối chứng}} - OD_{\text{blank}}$$

$$OD_{\text{mẫu thử}} = OD_{\text{mẫu thử}} - OD_{\text{blank}}$$

Giá trị IC₅₀ (nồng độ ức chế 50%) sẽ được xác định nhờ vào phần mềm máy tính TableCurve2Dv4.

Phương pháp thử hoạt tính gây độc tế bào ung thư

Phương pháp xác định hoạt tính gây độc tế bào (cytotoxic assay) ung thư của phân đoạn chứa saponin đối với tế bào nuôi cấy dạng đơn lớp: Phương pháp thử độ độc tế bào *in vitro* được Viện Ung thư Quốc gia Hoa Kỳ (National Cancer Institute – NCI) xác nhận là phép thử độ độc tế bào chuẩn nhằm sàng lọc, phát hiện các chất có khả năng kìm hãm sự phát triển hoặc diệt tế bào ung thư ở điều kiện *in vitro*. Phép thử này được thực hiện theo phương pháp của Skehan [136]. Phép thử tiến hành xác định hàm lượng protein tế bào tổng số dựa vào mật độ quang học đo được khi thành phần protein của tế bào được nhuộm bằng sulforhodamine B (SRB). Giá trị OD máy đo được tỷ lệ thuận với lượng SRB gắn với phân tử protein, do đó lượng tế bào càng nhiều (lượng protein càng nhiều) thì giá trị OD càng lớn. Phép thử được thực hiện trong điều kiện cụ thể như sau:

i) Trypsin hóa tế bào thí nghiệm để làm rời tế bào và đếm trong buồng đếm để điều chỉnh mật độ cho phù hợp với thí nghiệm. Tiến hành đưa 190 µL tế bào vào đĩa 96 giếng để thử nghiệm.

ii) Mẫu thử được hòa tan trong DMSO 100% để có nồng độ ban đầu (stock) là 20 mM. Tiến hành pha loãng mẫu trên đĩa 96 giếng bằng môi trường nuôi cấy tế bào (không có FBS- fetal bovine serum) thành 4 dãy nồng độ từ cao xuống thấp. Chất thử đã pha loãng ở các nồng độ (10 μ L) được đưa vào các giếng của đĩa 96 giếng đã chuẩn bị tế bào ở trên. Giếng không có chất thử nhưng có tế bào ung thư (190 μ L) + DMSO 1% (10 μ L) sẽ được sử dụng làm đối chứng ngày 0. Sau 1 giờ, giếng đối chứng ngày 0 tế bào sẽ được cố định bằng trichloroacetic acid (TCA) 20%.

iii) Ủ trong tủ ấm 72 giờ. Sau 72 giờ, tế bào được cố định bằng TCA trong 1 giờ, được nhuộm bằng SRB trong 30 phút ở 37 °C, sau đó rửa 3 lần bằng acetic acid rồi để khô ở nhiệt độ phòng.

iv) 10 mM dung dịch đệm (Tris base) để hòa tan lượng SRB, lắc nhẹ trong 10 phút rồi đọc kết quả OD ở bước sóng 540 nm trên máy ELISA Plate Reader (BioTek).

v) Phần trăm ức chế sự phát triển của tế bào khi có mặt chất thử sẽ được xác định thông qua công thức sau:

$$\% \text{ Ức chế} = 100 - \frac{OD(\text{mẫu}) - OD(\text{ngày0})}{OD(\text{DMSO}) - OD(\text{ngày0})} \times 100$$

vi) Phép thử được lặp lại 3 lần để đảm bảo tính chính xác. Ellipticine ở các nồng độ 10 - 2 - 0,4 - 0,08 μ g/mL được sử dụng như là chất đối chứng tham khảo.

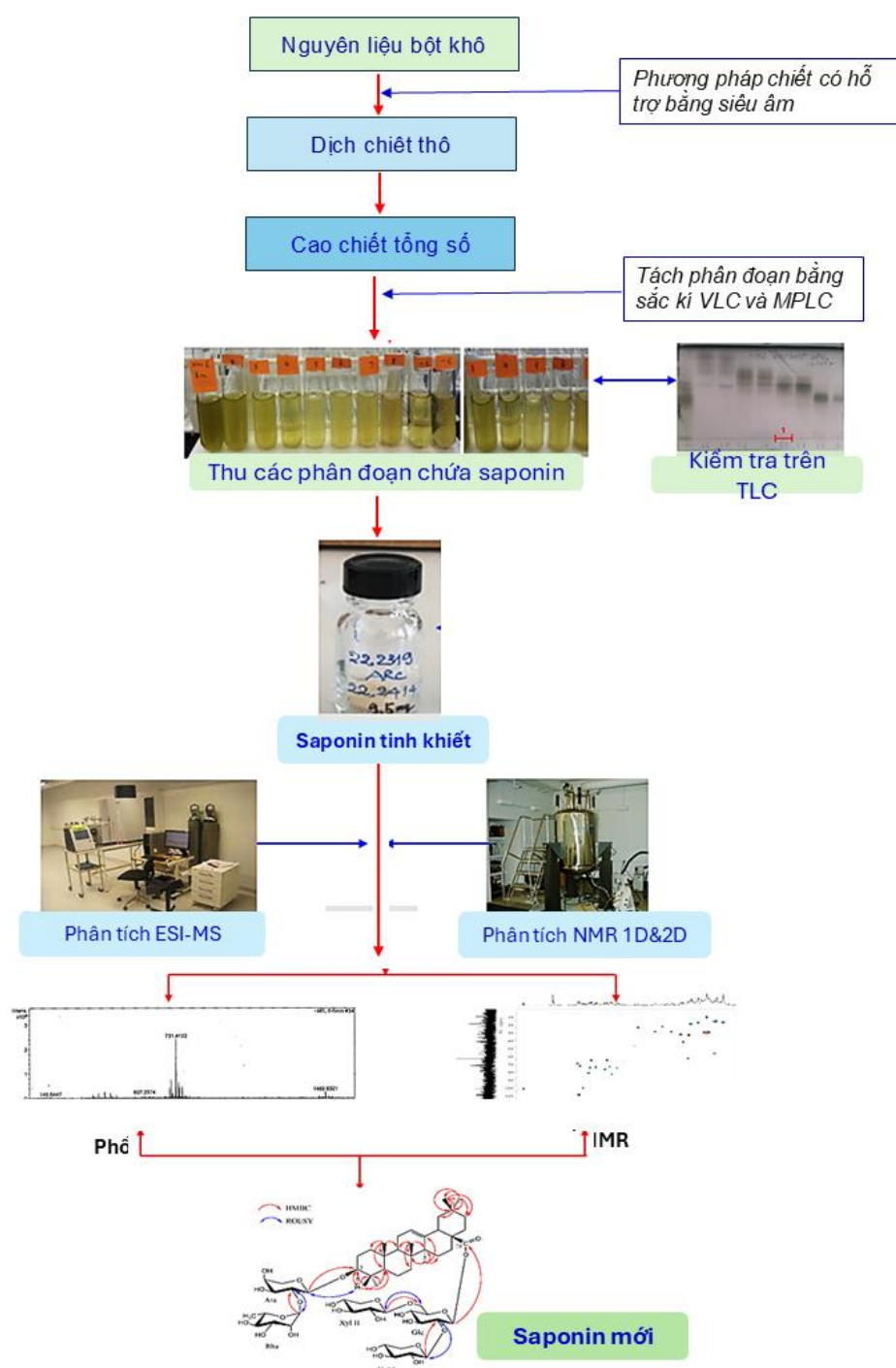
DMSO 1% luôn được sử dụng như đối chứng âm (nồng độ cuối cùng trong giếng thử là 0,05%). Giá trị IC₅₀ (nồng độ ức chế 50% sự phát triển) sẽ được xác định nhờ vào phần mềm máy tính TableCurve 2Dv4. Theo tiêu chuẩn của Viện ung thư quốc gia Hoa Kỳ (NCI), căn chiết được coi có hoạt tính tốt với IC₅₀ $\leq$ 20 μ g/ml, trong khi chất tinh khiết được coi có hoạt tính tốt khi IC₅₀ $\leq$ 5 μ M [60].

2.3.3.4. Phân lập saponin từ lá của Lan tai cáo

Các thí nghiệm phân lập saponin từ Lan tai cáo được thực hiện theo sơ đồ ở Hình 2.3.

(1) Phương pháp phân lập saponin

Từ cao chiết tổng số của Lan tai cáo và các phân đoạn chứa saponin được tách bằng sắc kí VLC và MPLC được sử dụng để phân lập saponin theo sơ đồ khái quát ở Hình 2.3.



Hình 2.3. Sơ đồ phân lập saponin từ Lan tai cao

(2) Phương pháp sắc ký

Việc tinh chế các hợp chất được thực hiện trên hệ thống sắc ký lỏng chân không trên silica gel RP-18 (75-200 μm , Silicycle, Canada) và silica gel 60 (63-200 μm , Sigma Aldrich, Pháp). Việc phân lập và tinh chế thêm saponin từ các phân đoạn được thực hiện

trên hệ thống sắc ký lỏng trung áp (MPLC) sử dụng pha tĩnh trên Merck silica gel 60, cỡ hạt 15-40 μm (Đức) bằng máy bơm Gilson M 305 và máy thu thập phân đoạn (Biopharmacia). Cột thủy tinh Büchi (460 mm $\times$ 25 mm, 110 mm $\times$ 15 mm) đã được sử dụng trong hệ thống MPLC. Sắc ký lớp mỏng (TLC) và sắc ký lớp mỏng hiệu năng cao (HPTLC) được thực hiện trên Merck silica gel 60F254 (Đức), với dung môi CHCl_3 -MeOH- H_2O -AcOH (70:30:5:1). Các vết TLC và HPTLC được xác định bằng cách phun lên các đĩa hỗn hợp EtOH H_2SO_4 (50:1), đun nóng ở 100°C trong 10 phút. Độ quay quang học được xác định trên máy đo phân cực tự động AA-10R (Anh). Phổ khối ion hóa phun điện tử có độ phân giải cao (HR-ESI-MS) được thực hiện trên máy quang phổ khối Bruker micrOTOF II (Bruker, Đức) ở chế độ ion dương. Sắc ký khí (GC) được thực hiện trên thiết bị sắc ký khí ThermoQuest Trace (Thermo Scientific, Hoa Kỳ) sử dụng OV-17 trên cột Silanox (0,3 mm $\times$ 50 m) [1], [96].

Bột lá khô được tách chiết bằng phương pháp có sự hỗ trợ siêu âm sử dụng dung môi EtOH/ H_2O (75:25) ở điều kiện 200 W, 60°C , 45 phút, thu được 500 mL cho mỗi lần tách chiết. Cao chiết tổng số được tách tiếp bằng VLC silica gel RP-18 và rửa giải lần lượt bằng H_2O 100%, EtOH/ H_2O 50:50, EtOH 100%, các phân đoạn thu được được phân tích bằng sắc ký lớp mỏng TLC (Thin Layer Chromatography). Các phân đoạn thu được được sử dụng để tách phân đoạn bằng sắc ký MPLC silica gel 60 (CHCl_3 /MeOH/ H_2O 80:20:2, 75:25:3, v/v/v), gel 60 (CHCl_3 /MeOH/ H_2O 75:25:3, v/v/v), tiến hành tách phân đoạn nhiều lần và kiểm tra bằng sắc ký lớp mỏng để tạo đơn chất.

(3) Phương pháp xác định đơn chất

Quá trình thủy phân 3 mg mỗi glycoside được thực hiện với 5 mL dung dịch 2N. CF_3COOH (4 mL) và nước (1 mL) trong 3 giờ ở 95°C . Hỗn hợp phản ứng sau đó được pha loãng trong nước (10 mL) và được chiết bằng 5 mL CH_2Cl_2 (3 lần). Lóp CH_2Cl_2 được làm bay hơi bằng MeOH cho đến khi dung môi trở nên trung tính để tạo ra cặn đường và được sắc ký tiếp bằng TLC trên silica gel sử dụng hệ rửa giải CHCl_3 /MeOH/ H_2O 8:5:1 bằng cách so sánh với các mẫu xác thực. Hỗn hợp monosacharide được pha loãng với 100 μL pyridin khan và xử lý với 0,06 mol/L $\text{HSCH}_2\text{CH}(\text{NH}_2)\text{COOCH}_3\cdot\text{HCl}$ ở 60°C trong 1 giờ, sau đó thêm 150 μL HMDS-TMCS

(3:1) ở 60°C trong 30 phút. Sau khi dừng phản ứng bằng cách thêm 2 mL H₂O, phần đường còn lại được chiết bằng 0,1 mL n-hexan. Lớp n-hexan được thu hồi và sau đó thành phần đường được phân tích (1 µL) bằng sắc ký khí. Việc xác định đường được thực hiện bằng cách so sánh thời gian lưu của chúng với thời gian lưu của đường tiêu chuẩn (Sigma-Aldrich) [96]. Thời gian lưu (phút): 11,07 (rhamnose), 13,82 (fucose) và 17,51 (glucose).

(4) Phương pháp phân tích phổ cộng hưởng từ hạt nhân

Các hợp chất được xác định cấu trúc hóa học trên cơ sở sử dụng các phương pháp đo phổ bằng các thiết bị hiện đại đồng thời kết hợp với phân tích và tra cứu tài liệu tham khảo. Các phương pháp đo phổ được sử dụng gồm:

i) Phổ khối lượng phân giải cao (HR-ESI-MS) được đo trên máy Agilent 6530 Accurate-Mass Q-TOF LC/MS (Agilent Technologies, Mỹ) tại Viện Hóa sinh biển, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

ii) Phổ cộng hưởng từ hạt nhân (NMR) được đo trên máy Bruker AM500 FT-NMR (Bruker, Đức) tại Viện Hóa học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam. Hợp chất chuẩn đã được thiết lập để hiệu chuẩn phổ ¹H và ¹³C là tetramethylsilane (TMS). Các kỹ thuật phổ cộng hưởng từ hạt nhân được sử dụng bao gồm: Phổ cộng hưởng từ hạt nhân một chiều (¹H-NMR, ¹³C-NMR và DEPT); phổ cộng hưởng từ hạt nhân hai chiều (HSQC, HMBC, COSY và NOESY). Dung môi sử dụng bao gồm DMSO-d₆, CD₃OD, CDCl₃ (việc lựa chọn dung môi đo phụ thuộc vào bản chất của từng mẫu, theo nguyên tắc dung môi phải hòa tan hoàn toàn mẫu đo và không che lấp các tín hiệu phân tích).

iii) Phổ sắc ký khí (GC) được đo trên máy GCMS-QP2010 Plus (Shimadzu, Nhật Bản) tại Viện Hóa học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

(5) Phương pháp xác định chất mới

Sử dụng thư viện SciFinder là một nguồn tài nguyên từ Chemical Abstracts Service (www.cas.org) để tìm kiếm tài liệu và tìm thông tin cơ bản về cấu trúc hóa học, tính chất và phản ứng của chất. SciFinder hỗ trợ tìm kiếm trực tiếp các chất và phản ứng hóa học, trên cơ sở công thức và một số tính chất trong cấu trúc hóa học chất đã được

xác định. Trên cơ sở nguồn dữ liệu tham khảo sẽ xác định cấu trúc không gian, xác định các liên kết giữa các nguyên tử của hợp chất cần nghiên cứu. Các tệp cấu trúc từ các bản ghi SciFinder hiện có có thể được nhập vào ChemDraw, để chia sẻ với những người khác để tra cứu và tìm kiếm cấu trúc của chất.

2.3.3.5. Phân tích tương tác phân tử và động học phân tử *in silico*

Nghiên cứu về tiềm năng, cơ chế và ái lực liên kết của saponin steroid phân lập từ Lan tai cáo (*H. verticillata* var. *verticillata*) với Bcl-2 (B-cell lymphoma 2). Bcl-2 là loại protein đóng vai trò quan trọng trong việc điều hòa chu trình sống của tế bào. Bcl-2 thúc đẩy ung thư bằng cách ngăn chặn quá trình chết theo chương trình (apoptosis) của tế bào. Kỹ thuật mô phỏng quá trình tương tác phân tử và động học phân tử *in silico* được sử dụng để phân tích hoạt tính gây độc tế bào ung thư của saponin phân lập được từ Lan tai cáo.

(1) Mô phỏng tương tác phân tử

Chuẩn bị protein

Cấu trúc protein mục tiêu Bcl-2 (ID: 6GL8 là cấu trúc ba chiều của protein mục tiêu Bcl-2) được thu thập từ Ngân hàng dữ liệu protein (<https://www.rcsb.org/>). Biovia Discovery Studio Visualizer v24.1 (Dassault Systemes BIOVIA, Hoa Kỳ) là phần mềm giúp hình ảnh hóa 3D phân tử nghiên cứu. Các hướng liên kết và cấu hình thích hợp của phân tử được thực hiện bằng AutoDock Tools v1.5.7. Việc xác nhận giao thức tương tác phân tử thông qua tính độ lệch bình phương căn bậc hai (RMSD- root mean square deviation). Xác định tương tác giữa các phân tử nghiên cứu và thụ thể của protein Bcl-2 được thực hiện bằng phần mềm Biovia Discovery Studio Visualizer.

Mô tả saponin

Có tổng số 5 saponin từ *H. verticillata* var. *verticillata* gồm 3 saponin trong các nghiên cứu trước và 2 saponin trong nghiên cứu này [155]. Parasiticoside A và parasiticoside B có công thức phân tử là $C_{33}H_{54}O_{11}$ và $C_{39}H_{64}O_{15}$, khối lượng phân tử lần lượt là 626,3666 và 772,4245 m/z. 3 saponin steroid đã được nghiên cứu trước đây là

3 β ,4 α -dihydroxy-5 β -spirost-(25)27-en-1 β -yl O- α -L-arabinopyranoside (1), 3 β -hydroxy-5 β -spirost-25(27)-en-1 β -yl O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-xylopyranoside (2), (23S,24S,25S)-3 β ,4 α ,23,24-tetrahydroxy-5 β -spirostan-1 β -yl O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-fucopyranoside (3), có công thức phân tử là C₃₂H₅₀O₉, C₃₈H₆₀O₁₂, C₃₉H₆₄O₁₅ và khối lượng phân tử là lần lượt là 578,3455; 708,4085; 772,4245 m/z.

Chuẩn bị phối tử

Cấu trúc 2D của saponin được chuẩn bị ở định dạng *.sdf bằng phần mềm ChemDraw Prime v23.1 (Perkin Elmer, Hoa Kỳ) và được chuyển đổi thành cấu trúc 3D ở định dạng *.pdb bằng phần mềm Biovia Discovery Studio Visualizer. Các loại liên kết gồm: Liên kết hydrogen phân cực, điện tích riêng phần và tất cả các lực xoắn được thực hiện bằng AutoDock Tools. Cấu trúc 2D của Navitoclax được lấy từ cơ sở dữ liệu PubChem ở định dạng *.sdf (<https://pubchem.ncbi.nlm.nih.gov>) và được tối ưu hóa cấu trúc 3 chiều bằng phần mềm Avogadro v1.2.0 [56].

(2) Mô phỏng động học phân tử

Thực hiện việc ghép nối phân tử giữa các hợp chất và thụ thể bằng ứng dụng AutoDock Tools. Thuật toán di truyền Lamarckian được sử dụng để xác định các cấu hình liên kết giữa các phân tử theo hướng có lợi nhất về mặt năng lượng. Tất cả các tham số được đặt mặc định và Navitoclax được chọn làm hợp chất kiểm soát, đối chứng.

Mô phỏng động học phân tử để xác nhận khả năng duy trì ghép nối tốt nhất với protein 6GL8 được thực hiện trong 10 ns bằng phần mềm GROMACS v2024. Protein được cố định cấu trúc để không làm thiếu đi một số nguyên tử thành phần và có thêm dư lượng khác bằng phần mềm Swiss-PdbViewer [55]. Swiss PARAM được sử dụng để tạo ra các cấu trúc của phối tử. Hệ thống cân bằng được thiết lập để chạy mô phỏng động học phân tử lên đến 50000 bước trên 100 ps, sử dụng máy tính CPU Intel(R) Xeon(R) @ 2,20 GHz được trang bị GPU NVIDIA Tesla T4 16GB GDDR6. Dữ liệu mô phỏng động học phân tử được phân tích ở 100 ns bằng UCSF Chimera v1.17.3 và phần mềm Grace [109]. Dữ liệu mô phỏng được phân tích về số liên kết hydrogen trong mỗi khung theo thời gian, biến động bình phương trung bình gốc (RMSF) của mỗi phân tử trong cấu trúc và độ lệch bình phương trung bình căn bậc hai (RMSD) của cấu trúc đã cho theo thời gian.

2.3.4. Phương pháp phân tích số liệu

Các phần mềm Tin sinh học được sử dụng trong phân tích hệ gene như MEGA11, BioEdit, BLAST trong NCBI.

Các số liệu được phân tích bằng phần mềm SPSS (Statistical software suite) của Windows phiên bản 29.0.2.0. So sánh sự khác biệt bằng kiểm định kiểm định Duncan ở mức ý nghĩa $\alpha = 0,05$.

2.4. ĐỊA ĐIỂM NGHIÊN CỨU VÀ HOÀN THÀNH LUẬN ÁN

Phân tích Tin sinh học và một số phân tích hóa sinh được thực hiện tại các phòng thí nghiệm thuộc Bộ môn Di truyền học & Công nghệ sinh học, Khoa Sinh học, Trường Đại học Sư phạm - Đại học Thái Nguyên. Phân tích hoạt tính sinh học và hóa học thực vật được thực hiện tại Viện Sinh học thuộc Viện Hàn lâm Khoa học và Công nghệ Việt Nam. Giải trình tự DNA lục lạp và phân tích hệ gene lục lạp của Lan tai cáo được thực hiện tại Trung tâm nghiên cứu vật liệu sinh học quốc tế (International Biological Materials Reserch Center - IBMRC), Viện nghiên cứu khoa học sinh học và công nghệ sinh học Hàn Quốc (Korea Research Institute of Bioscience and Biotechnology - KRIBB), Hàn Quốc (Republic of Korea).

Luận án được hoàn thành tại Bộ môn Di truyền học & Công nghệ sinh học, Khoa Sinh học, Trường Đại học Sư phạm - Đại học Thái Nguyên.

Chương 3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

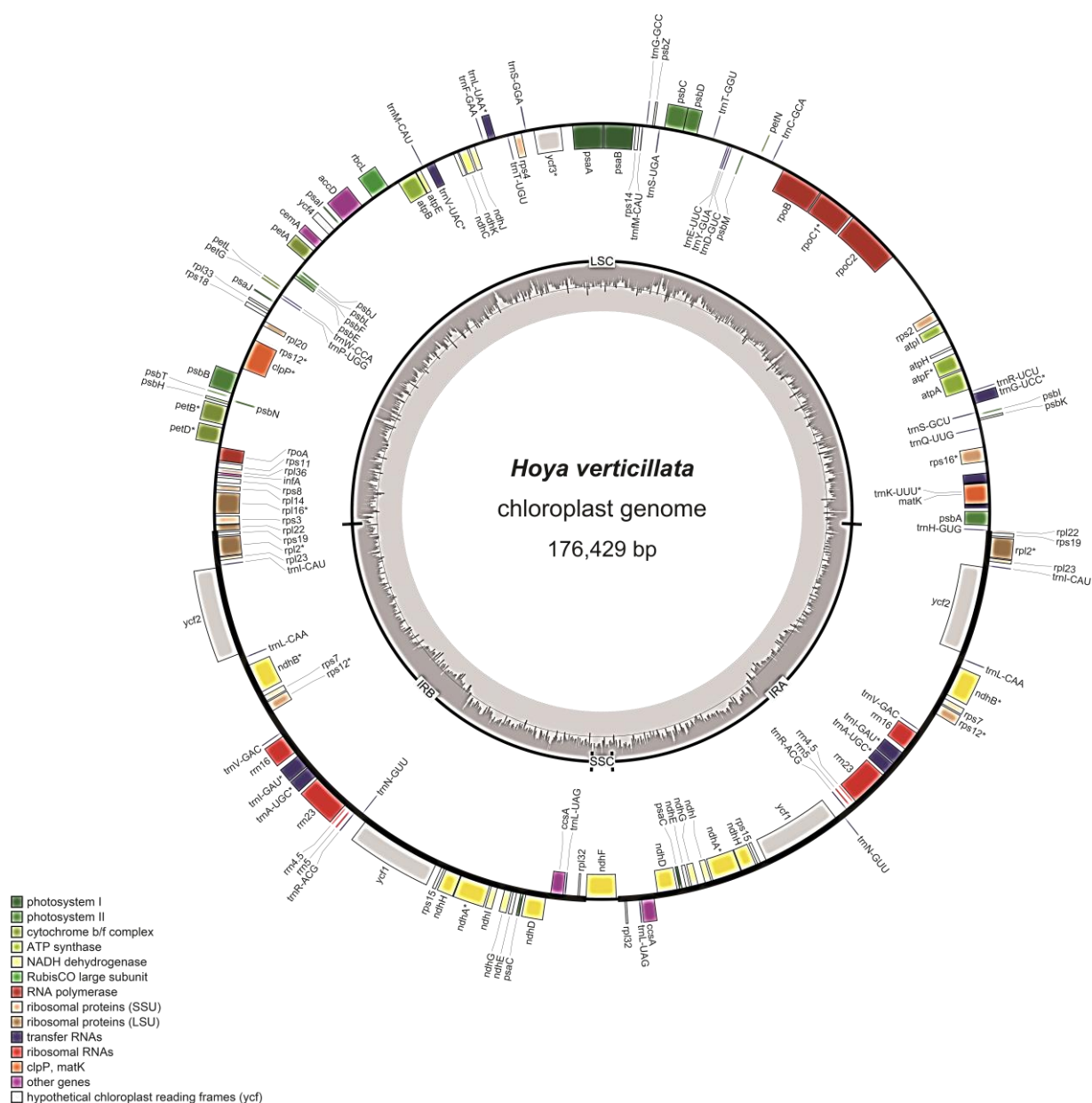
3.1. ĐẶC ĐIỂM HỆ GENE LỤC LẠP CỦA LOÀI LAN TAI CÁO

3.1.1. Cấu trúc hệ gene

Hệ gene lục lập hoàn chỉnh của Lan tai cáo (*H. verticillata* var. *verticillata*) có kích thước 176.429 bp (Hình 3.1). Bốn vùng điển hình của phân tử DNA dạng vòng, bao gồm vùng sao chép đơn lớn LSC (91.020 bp), sao chép đơn nhỏ SSC (2.297 bp) và 2 vùng lặp lại đảo ngược IR, mỗi vùng IR có kích thước 41.556 bp, trong đó hàm lượng GC chiếm 37,1% (Bảng 3.1).

Kết quả hiển thị thông tin hệ gene lục lập từ GeSeq và tRNAscan-SE cho thấy hệ gene lục lập của Lan tai cáo có tổng cộng 144 gene, bao gồm 98 gene mã hóa protein, 38 gene tRNA và 8 gene rRNA (Bảng 3.1). Các gene được phân loại thành ba nhóm chính dựa trên chức năng của chúng. Nhóm gene liên quan đến quang hợp có 44 gene, gồm các nhóm gene mã hóa các tiểu đơn vị của ATP synthase, phức hợp cytochrome, hệ thống quang hợp I và II, NADPH dehydrogenase và các tiểu đơn vị lớn của các enzyme tham gia vào quá trình quang hợp và tùy thuộc vào ánh sáng; 60 gene còn lại thuộc các nhóm chức năng liên quan đến phiên mã và dịch mã. Chín gene còn lại được phân thành các nhóm gene khác, bao gồm 5 gene có chức năng liên quan đến quá trình xử lý RNA (*matK* – tổng hợp protein tham gia vào quá trình cắt nối các intron khỏi các RNA thông tin), 4 gene bảo tồn khung đọc mở, tổng hợp cytochrome loại c (*ccsA*), tổng hợp acid béo (*accD*), chuyển hóa carbon (*cemA*) và phân giải protein. Trình tự hệ gene lục lập của Lan tai cáo có chú thích đã được đăng kí trên GenBank mang mã truy cập là OR475244.1.

Phân tích cấu trúc vùng IR của hệ gene lục lập *H. verticillata* var. *verticillata* cho thấy 2 vùng này có số lượng và cấu trúc gene tương tự nhau nhưng theo hướng ngược nhau; mỗi vùng IR bao gồm 26 gene, cụ thể là 4 gene rRNA, 8 gene tRNA, 4 gene mã hóa protein ribosome, 7 gene mã hóa protein NADH-dehydrogenase, 2 gene bảo tồn khung đọc mở và 1 gene tiểu đơn vị quang hợp I. Các gene chứa 1 intron gồm *ndhA*, *ndhB*, *rps12*, *trnA-UGC* và *trnI-GAU* (Bảng 3.2).



Hình 3.1. Bản đồ hệ gene lục lạp của Lan tai cáo (*Hoya verticillata* var. *verticillata*) thu tại Việt Nam. Các gene ở bên trong vòng tròn ngoài được phiên mã theo chiều kim đồng hồ, các gene ở bên ngoài được phiên mã ngược chiều kim đồng hồ. Các gene thuộc các nhóm chức năng khác nhau được mã hóa màu. Vòng tròn đồng tâm bên trong chỉ ra cấu trúc bốn phần của hệ gene lục lạp, bao gồm các vùng LSC, SSC và IR. Vùng màu xám nhạt biểu thị hàm lượng AT và vùng màu xám đậm biểu thị hàm lượng GC.

Hệ gene lục lạp của thực vật hạt kín thường có thành phần và hàm lượng gene được bảo tồn cao. Hệ gene lục lạp của Lan tai cáo (*H. verticillata* var. *verticillata*) có 144 gene, trong đó 17 gene bao gồm 1 hoặc 2 intron (15 gene chứa 1 intron và 2 gene chứa 2 intron). Vùng SSC chỉ chứa 1 gene *ndhF*, không có intron. Vùng LSC có 10 gene chứa 1 intron, trong đó có 3 gene quan trọng trong việc tổng hợp tiểu đơn vị của phức hợp cytochrome là *petB*, *petD* và *atpF* (*petD* và *atpF* đều chứa 2 intron) (Bảng 3.2).

Bảng 3.1. Tóm tắt hệ gene lục lạp của loài *H. verticillata* var. *verticillata*

Thành phần	Số lượng
Kích thước hệ gene (bp)	176.429
Kích thước vùng sao chép đơn lớn (LSC) (bp)	91.020
Kích thước vùng sao chép đơn nhỏ (SSC) (bp)	2.297
Kích thước vùng lặp lại đảo ngược IR (bp)	41.556
Tỷ lệ GC (%)	37,1
Tổng số gene	144
Tổng số gene mã hóa protein	98
Tổng số gene mã hóa tRNA	38
Tổng số gene mã hóa rRNA	8

Bảng 3.2. Thành phần gene của hệ gene lục lạp *H. verticillata* var. *verticillata*

Loại gene	Chức năng	Tên gene
Quang hợp (52 gene)	Các tiểu đơn vị của ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF^a</i> , <i>atpH</i> , <i>atpI</i>
	Các tiểu đơn vị của NADH-dehydrogenase	<i>ndhA^a</i> (×2), <i>ndhB</i> (×2) ^a , <i>ndhC</i> , <i>ndhD</i> (×2), <i>ndhE</i> (×2), <i>ndhF</i> , <i>ndhG</i> (×2), <i>ndhH</i> (×2), <i>ndhI</i> (×2), <i>ndhJ</i> , <i>ndhK</i>
	Các tiểu đơn vị của phức hợp cytochrome b/f	<i>petA</i> , <i>petB^a</i> , <i>petD^a</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	Các đơn vị của hệ thống quang hợp I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> (×2), <i>psaI</i> , <i>psaJ</i>

	Các đơn vị của hệ thống quang hợp II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>
	Tiểu phần nhỏ của rubisco	<i>rbcL</i>
Phiên mã và dịch mã (80 gene)	Đơn vị lớn của ribosome	<i>rpl14, rpl16^a, rpl2^a, rpl20, rpl22, rpl23, rpl32 (×2), rpl33, rpl36</i>
	RNA polymerase	<i>rpoB, rpoA, rpoC1^a, rpoC2</i>
	Đơn vị nhỏ của protein ribosome	<i>rps11, rps12^a (×2), rps14, rps15 (×2), rps16^a, rps18, rps19, rps2, rps3, rps4, rps7 (×2), rps8</i>
	Các gene rRNA	<i>rrn23 (×2), rrn16 (×2), rrn5 (×2), rrn4.5 (×2)</i>
	Các gene tRNA	<i>trnA-UGC^a (×2), trnC-GCA, trnD-GUC, trnE-UUC^a, trnF-GAA, trnG^a-CAU, trnG-GCC, trnH-GUG, trnI-CAU (×2), trnI-GAU (×2), trnG-UCC, trnK-UUU^a, trnL-CAA (×2), trnL-UAA^a, trnL-UAG (×2), trnM-CAU, trnN-GUU (×2), trnP-UGG, trnQ-UUG, trnR-ACG (×2), trnR-UCU, trnS-CGA^a, trnS-GCU, trnS-GGA, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC (×2), trnV-UAC, trnW-CCA, trnY-GUA</i>
	Yếu tố khởi đầu dịch mã	<i>infA</i>
Các gene khác (12 gene)	Đơn vị phụ của Acetyl-CoA-carboxylase (tổng hợp acid béo)	<i>accD</i>
	Gene tổng hợp cytochrome loại c	<i>ccsA (×2)</i>
	Vận chuyển hóa carbon	<i>cemA</i>
	Protease	<i>clpP^b</i>
	Xử lý RNA	<i>matK</i>
	Bảo tồn khung đọc mở	<i>ycf1 (×2), ycf2 (×2), ycf3^b, ycf4</i>

(a): Gene chứa 1 intron, (b): Gene chứa 2 intron, (×2): Gene ở vùng lặp lại đảo ngược IR.

Tần suất sử dụng mã di truyền của 98 gene mã hóa protein đã được đánh giá trong Bảng 3.3. Tần suất sử dụng mã di truyền được tìm thấy trong các vùng mã hóa là 29.223, trong đó mã kết thúc bằng U- và A- được tìm thấy thường xuyên hơn G- và C-. AUU (mã hóa izoleusine) là mã di truyền được tìm thấy phổ biến nhất trong số 64 mã di truyền với tần suất RSCU là 4,27, tiếp theo là GAA (mã hóa glutamic acid) với RSCU là 4,24, AAA (mã hóa lysine) với RSCU là 4,19. Trong khi đó, hiếm nhất là các codon AUC, CUG, GUG với tần suất RSCU là 0,01. Tổng cộng có 37 codon được sử dụng thường xuyên hơn với tần suất RSCU > 1. Codon AUA được sử dụng một lần.

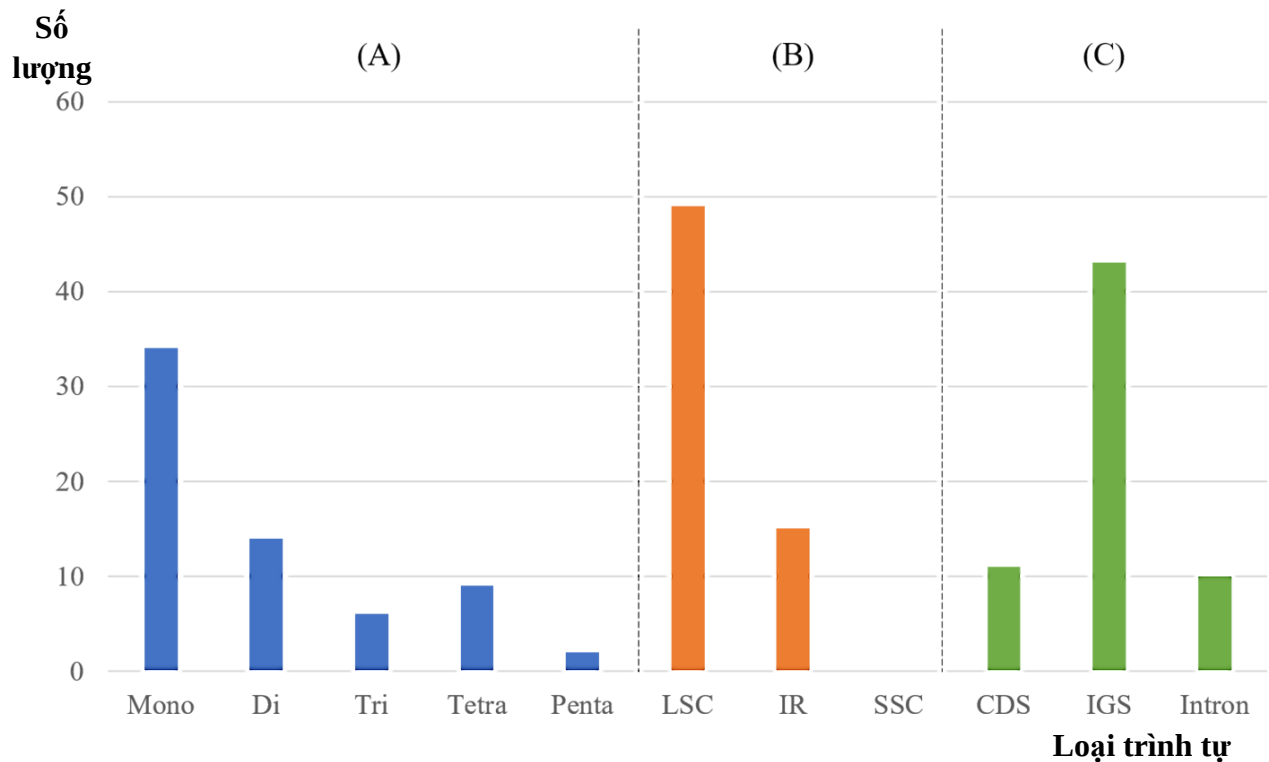
Trình tự lặp lại của hệ gene lục lạp loài *H. verticillata* var. *verticillata* được xác định bao gồm các trình tự lặp lại mono-nucleotide với tần suất 34 lần (53,12%), trình tự lặp lại di-nucleotide là 14 lần (21,88%), trình tự lặp lại tri-nucleotide là 6 lần (9,38%), trình tự lặp lại tetra-nucleotide là 9 lần (14,06%), trình tự lặp lại penta-nucleotide là 1 lần (1,56%). Do đó, các trình tự lặp lại mono-nucleotide phổ biến nhất là A/T, với tần suất 34 lần (53,12%), vì A và T bổ sung cho nhau (Hình 3.2). Trình tự lặp lại đơn giản gồm di-nucleotide như TA (6 lần), GA (1 lần); trình tự lặp lại tri-nucleotide như TTC (2 lần) và GAA, ATT, TAT (1 lần), trình tự lặp lại tetra-nucleotide như AAAT, ATTA, AATT, CTTT, TATT, TTAG, TTTA (1 lần) và trình tự lặp lại pentanucleotide như TTTTA (1 lần).

Tần suất lặp lại đơn giản ở các vùng khác nhau của hệ gene cũng có sự phân hóa rõ ràng, trong đó vùng LSC có tần suất 49 trình tự lặp lại, bao gồm 27 mono-nucleotide, 11 di-nucleotide, 4 tri-nucleotide, 6 tetra-nucleotide, 1 penta-nucleotide; vùng IR có tần suất lặp lại là 15, trong đó có 7 mono-nucleotide, 3 di-nucleotide, 2 tri-nucleotide, 3 tetra-nucleotide và không có trình tự lặp lại penta-nucleotide (Hình 3.2A); vùng SSC chỉ có một gene và không chứa trình tự lặp lại đơn (Hình 3.2B). Trong các vùng gene mã hóa (CDS), vùng đệm xen kẽ (IGS) và vùng không mã hóa (Intron), tần suất các trình tự lặp lại đơn giản chủ yếu tồn tại ở các vùng đệm với tần suất 43 lần (vùng CDS là 11, intron là 10). Vùng đệm xen kẽ IGS chủ yếu chứa các trình tự lặp lại 1 nucleotide với tần suất 24 lần, bao gồm 9 di-nucleotide, 2 tri-nucleotide, 7 tetra-nucleotide, 1 penta-nucleotide; Vùng đệm và vùng intron có các trình tự lặp lại tương đương (Hình 3.2C).

Bảng 3.3. Tần số sử dụng codon đồng nghĩa tương đối (RSCU) cho các gene mã hóa protein ở *H. verticillata* var. *verticillata*.

Codon	AA	ObsFreq	RSCU	Codon	AA	ObsFreq	RSCU	Codon	AA	ObsFreq	RSCU
UAA	*	53	0,17	AUC	I	488	1,59	CGC	R	133	0,43
UAG	*	23	0,08	AUU	I	1308	4,27	CGG	R	167	0,55
UGA	*	21	0,07	AAA	K	1284	4,19	CGU	R	387	1,26
GCA	A	454	1,48	AAG	K	419	1,37	AGC	S	154	0,50
GCC	A	271	0,89	CUA	L	492	1,61	AGU	S	460	1,50
GCG	A	172	0,56	CUC	L	209	0,68	UCA	S	464	1,52
GCU	A	687	2,24	CUG	L	214	0,70	UCC	S	361	1,18
UGC	C	86	0,28	CUU	L	723	2,36	UCG	S	268	0,88
UGU	C	266	0,87	UUA	L	1015	3,32	UCU	S	695	2,27
GAC	D	258	0,84	UUG	L	622	2,03	ACA	T	499	1,63
GAU	D	1000	3,27	AUA	M	1	0,00	ACC	T	304	0,99
GAA	E	1298	4,24	AUC	M	2	0,01	ACG	T	167	0,55
GAG	E	399	1,30	AUG	M	695	2,27	ACU	T	582	1,90
UUC	F	580	1,89	CUG	M	2	0,01	GUA	V	605	1,98
UUU	F	1185	3,87	GUG	M	2	0,01	GUC	V	202	0,66
GGA	G	783	2,56	AAC	N	325	1,06	GUG	V	223	0,73
GGC	G	225	0,74	AAU	N	1176	3,84	GUU	V	585	1,91
GGG	G	382	1,25	CCA	P	355	1,16	UGG	W	520	1,70
GGU	G	645	2,11	CCC	P	279	0,91	UAC	Y	228	0,74
CAC	H	185	0,60	CCG	P	191	0,62	UAU	Y	906	2,96
CAU	H	483	1,58	CCU	P	458	1,50				
AUA	I	804	2,63	CAA	Q	793	2,59				

Ghi chú: * codon kết thúc; AA: amino acid; ObsFreq: Số lượng codon; RSCU: Mức sử dụng codon tương đối.

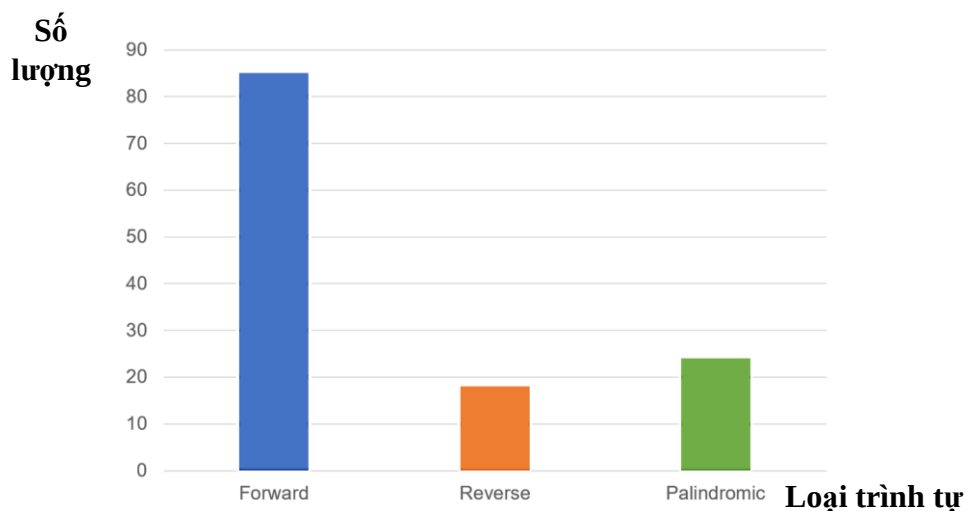


Hình 3.2. Phân tích các trình tự lặp lại đơn giản (SSR) của hệ gene lục lập hoàn chỉnh của *H. verticillata* var. *verticillata*. Trục hoành biểu thị các loại trình tự lặp lại đơn giản hoặc các vùng của hệ gene lục lập, trục tung biểu thị số lượng. A: Số lượng các loại trình tự lặp lại đơn giản (Mono-đơn, Di-đôi, Tri-ba, Tetra-bốn, Penta-năm); B: Tần suất các loại lặp lại trong các vùng sao chép đơn lớn, sao chép đơn nhỏ, lặp lại đảo ngược; C: Tần suất các loại lặp lại trong các vùng của hệ gene gồm CDS là vùng mã hóa protein, IGS là vùng đệm giữa các gene, Intron là vùng không mã hóa của gene.

Tần số lặp lại đơn giản SSR trong các vùng khác nhau của hệ gene cũng có sự phân hóa rõ rệt, trong đó vùng sao chép đơn lớn (LSC) có tần số 49 trình tự lặp lại (mono-nucleotide=27, di-nucleotide=11, tri-nucleotide=4, tetra-nucleotide=6, penta-nucleotide=1); vùng lặp lại đảo ngược IR có tần số trình tự lặp lại là 15 (mono-nucleotide=7, di-nucleotide=3, tri-nucleotide=2, tetra-nucleotide=3, không có trình tự lặp lại penta-nucleotide); vùng sao chép đơn nhỏ (SSC) chỉ có 01 gene và không chứa trình tự lặp lại đơn giản.

Xác định số lần lặp lại xuôi (F), đối xứng (P) và ngược (R) được thực hiện bằng chương trình REPuter4, với kích thước lặp lại tối thiểu là 30 bp với mã Hamming là

332. Trong số 127 trình tự lặp lại lớn (LSR), có 85 lần lặp lại xuôi, 24 lần lặp lại đối xứng và 18 lần lặp lại ngược, như thể hiện trong Hình 3.3.



Hình 3.3. Phân tích trình tự lặp lại lớn (LSR) trên quy mô hệ gene ở *H. verticillata* var. *verticillata*. Trục hoành là các trình tự lặp lại lớn (Forward-lặp xuôi, Reverse-lặp ngược, Palindromic- lặp đối xứng), trục tung là số lượng.

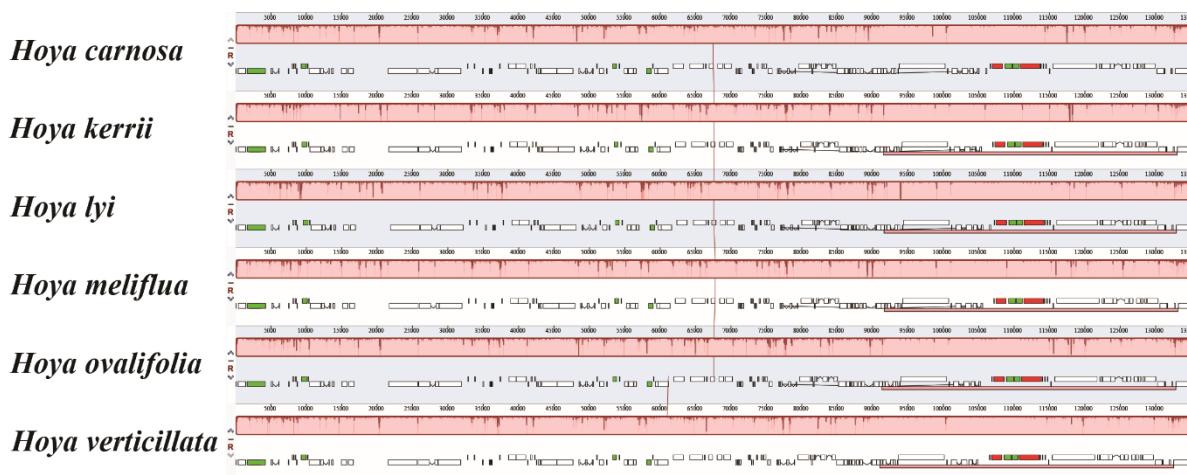
3.1.2. So sánh hệ gene lục lập hoàn chỉnh của *H. verticillata* var. *verticillata* thu thập ở Việt Nam với các loài thuộc chi *Hoya* khác trên GenBank

Trong nghiên cứu này, chúng tôi phân tích hệ gene lục lập hoàn chỉnh của sáu loài, bao gồm *H. verticillata* var. *verticillata* (Lan tai cáo) trong nghiên cứu này và 5 hệ gene lục lập đã công bố trên GenBank (*H. carnos*a-OK539590, *H. ovalifolia*-OL754666, *H. kerrii*-OL754673, *H. meliflua*-OL754674 và *H. lyi*-MW719055). Các trình tự được lựa chọn đều có chất lượng chú thích cao và đã được xác nhận rộng rãi, bảo đảm tính tin cậy cho việc so sánh cấu trúc bộ gene, biến đổi tại ranh giới vùng lặp lại đảo ngược (IR) và mức độ phân kỳ trình tự. Mặc dù gần đây có thêm một số loài thuộc chi *Hoya* được báo cáo, nhiều loài vẫn thiếu dữ liệu hệ gene lục lập hoàn chỉnh hoặc chú thích chưa đầy đủ, điều này có thể làm giảm độ chính xác và tính nhất quán trong phân tích so sánh.

Các hệ gene được căn chỉnh lại bằng thuật toán Mauve, trong đó các khối màu đại diện cho những vùng tương đồng giữa các loài (Hình 3.4). Kết quả cho thấy sáu hệ gene lục lập của *Hoya* có mức độ bảo tồn cao, với các khối đồng dạng được sắp xếp

thẳng hàng rõ rệt. Các số đánh dấu trên sơ đồ thể hiện vị trí nucleotide, trong khi các khối màu biểu thị những đặc điểm đặc trưng của từng hệ gene. Đáng chú ý, các vùng lặp lại đảo ngược (IR) có sự mở rộng về phía vùng sao chép đơn lớn (LSC) và vùng sao chép đơn nhỏ (SSC).

Phân tích cấu trúc cho thấy sáu hệ gene lục lập có mức độ tương đồng rất cao; các gene được sắp xếp liên tục dọc toàn bộ chiều dài hệ gene, và hầu như không xuất hiện hiện tượng tái sắp xếp ở quy mô lớn như đảo đoạn hoặc chuyển đoạn.



Hình 3.4. So sánh cấu trúc hệ gene lục lập hoàn chỉnh của sáu loài thuộc chi *Hoya* sử dụng phương pháp căn chỉnh Mauve. Sáu đường ngang biểu thị 6 hệ gene với tên loài ở đầu hàng; đối với mỗi hệ gene gồm có kích thước, số nucleotide và các gene tương ứng.

So sánh tỷ lệ tương quan giữa các vùng trong cấu trúc hệ gene lục lập của các loài thuộc chi *Hoya* có đặc trưng về mở rộng của vùng lặp lại đảo ngược IR ảnh hưởng đến cả vùng sao chép đơn lớn LSC và vùng sao chép đơn nhỏ SSC (Bảng 3.4). Tỷ lệ vùng IR tăng tương ứng với vùng LSC và SSC giảm.

Độ đa dạng nucleotide (Pi) được tính bằng phân tích cửa sổ trượt được thực hiện trong DnaSP v. 6.12.03 (phần mềm mạnh mẽ để phân tích đa hình nucleotide từ dữ liệu trình tự DNA), sử dụng độ dài phân tích là 600 bp và kích thước bước là 200 bp. Kết quả phân tích cho thấy cả hai vùng sao chép đơn đều được xác định là có sự khác biệt

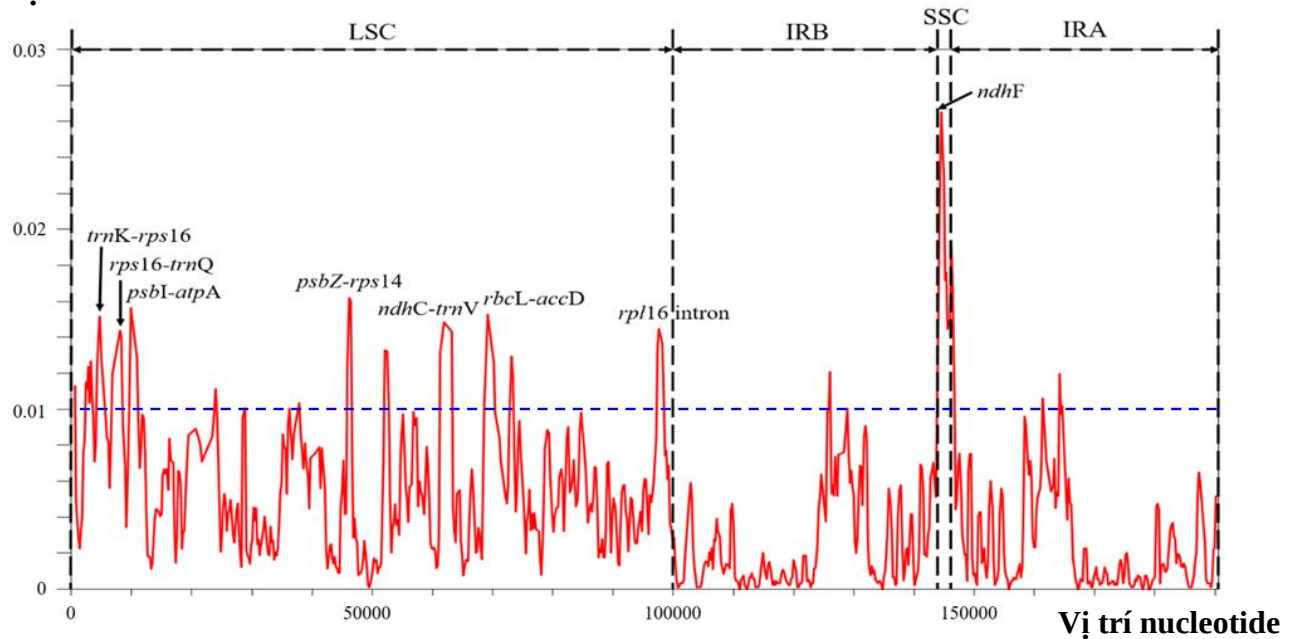
trình tự đáng kể hơn so với vùng lặp lại đảo ngược IR. Với giá trị giới hạn P_i là 0,03, tám vùng biến đổi cao đã được xác định gồm 6 vùng đệm giữa các gene (*trnK-rps16*, *rps16-trnQ*, *psbI-atpA*, *psbZ-rps14*, *ndhC-trnV*, *rbcL-accD*) và 2 vùng intron thuộc các gene *rpl16*, *ndhF* (Hình 3.5). Sáu vùng biến đổi cao thuộc về vùng LSC, một vùng nằm giữa vùng LSC và vùng IR và gene *ndhF* thuộc về vùng SSC.

Bảng 3.4. Tỷ lệ phân bố các vùng trong hệ gene lục lạp của 6 loài thuộc chi *Hoya*

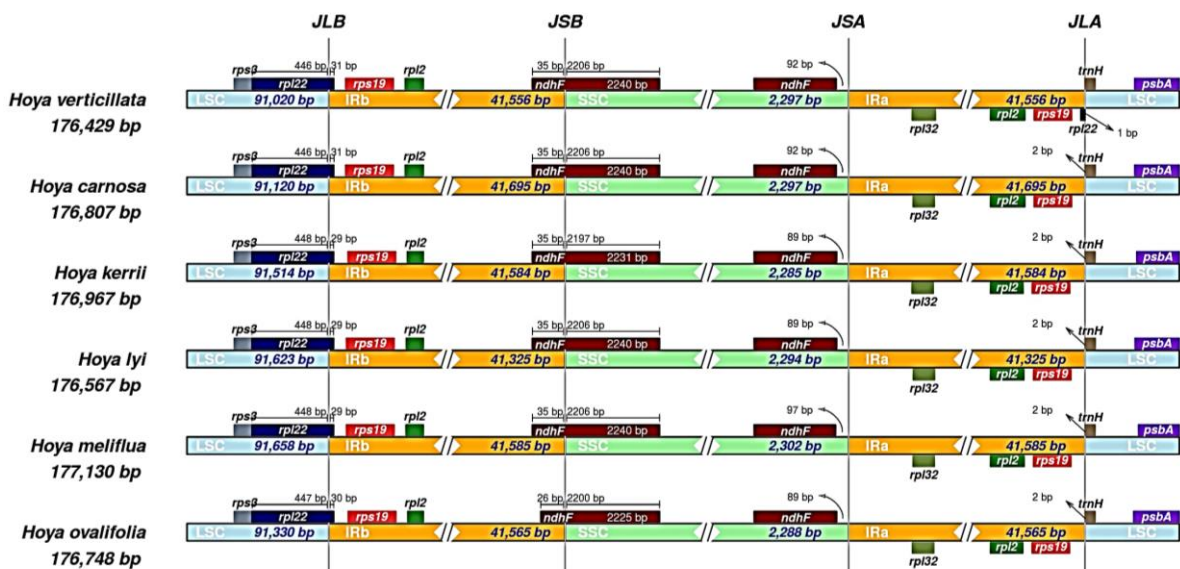
Loài	Tỷ lệ vùng LSC	Tỷ lệ vùng SSC	Tỷ lệ vùng IR
<i>H. carnosa</i> (OK539590)	0,516729	0,013026	0,472893274
<i>H. verticillata</i> var. <i>verticillata</i>	0,515902	0,013019	0,471079018
<i>H. ovalifolia</i> (OL754666)	0,516719	0,012945	0,470325318
<i>H. kerrii</i> (OL754673)	0,517116	0,012912	0,46995536
<i>H. meliflua</i> (OL754674)	0,517462	0,012996	0,469542144
<i>H. lyi</i> (MW719055)	0,518905	0,012992	0,468086311

Ranh giới LSC/IRB luôn có vị trí bên dưới gene *rpl22* ở đầu 3' của gene *rpl22*; điểm nối vùng LSC/IRB nằm ở vị trí 29-31 bp của gene *rpl22*, điểm nối IRb/SSC nằm ở vị trí cách gene *ndhF* 26-35 bp. Điểm nối SSC/IRA nằm ở khoảng cách 89-97 bp từ vị trí gene *ndhF*, điểm nối IRA/LSC nằm ở khoảng cách *rps19-trnH*, sự khác biệt được thể hiện đối với khoảng cách *rps19-trnH* của *H. verticillata* var. *verticillata* là chứa gene *rpl22* với khoảng 30 bp và điểm nối chồng lên nhau 1bp tại gene *rpl22*. Sự thu hẹp của vùng SSC và mở rộng của vùng IR trong hệ gene *Hoya* dẫn đến những thay đổi tương quan về chiều dài của 4 vùng cấu trúc và toàn bộ trình tự hệ gene lục lạp, tuy nhiên, sự thay đổi này là rất nhỏ (Hình 3.6).

Giá
trị Pi



Hình 3.5. Đồ thị so sánh các giá trị đa dạng nucleotide (Pi) thông qua phân tích cửa sổ trượt (chiều dài cửa sổ là 600 bp; kích thước bước là 200 bp) trên toàn bộ hệ gene lục lạp của 30 loài *Hoya*. Trục Y biểu thị giá trị Pi tại điểm giữa của cửa sổ trượt và trục X biểu thị vị trí nucleotide. Đường chấm màu xanh biểu thị giá trị ngưỡng Pi là 0,01. Các gene vượt quá ngưỡng này được dán nhãn phía trên các đỉnh của đồ thị.



Hình 3.6. So sánh điểm nối các vùng LSC, SSC và hai vùng IR trên 6 hệ gene lục lạp hoàn chỉnh của *Hoya*. Các chữ viết tắt JLB, JSB, JSA và JLA lần lượt biểu thị các điểm nối giữa LSC/IRB, SSC/IRB, SSC/IRA và LSC/IRA.

Cho đến thời điểm nghiên cứu, 45 hệ gene lục lạp của chi *Hoya* đã được công bố trên GenBank, bao gồm một số loài được công bố một phần của hệ gene, ví dụ như *H. latifolia*, *H. hamiltoniorum*, *H. insularis* và *H. coronaria*...; một số hệ gene lục lạp không đầy đủ như ở một số loài: *H. sp.* 3 WO-2022, *H. sp.* 11 WO-2022, *H. sp.* 8 WO-2022, *H. sp.* 4 WO-2022 và *H. sp.* 2 WO-2022. Kết quả phân tích hệ gene lục lạp hoàn chỉnh của *H. verticillata* var. *verticillata* và 5 trình tự hệ gene lục lạp hoàn chỉnh khác của *Hoya* trên GenBank cho thấy cấu trúc có nhiều điểm tương đồng (Hình 3.4 và Hình 3.6). Vùng SSC bị thu hẹp và vùng IR được mở rộng đáng kể. Tuy nhiên, sự khác biệt giữa các hệ gene lục lạp các loài này thể hiện rõ nhất ở kích thước hệ gene của chúng, dao động từ 175.404 bp ở *H. lacunosa* đến 179.069 bp ở *H. ariadna*, chủ yếu là do sự thay đổi chiều dài của vùng LSC (80.795 bp ở *H. liangii* đến 92.072 bp ở *H. sp2* ZCF6006). Vùng SSC chỉ bao gồm một gene duy nhất (*ndhF*), dao động từ 2.265 bp ở *H. pandurata* đến 2.306 bp ở *H. caudata*. Trong nghiên cứu này, hệ gene *H. verticillata* var. *verticillata* thu thập được ở Việt Nam có kích thước là 176.429 bp, vùng LSC dài 91.020 bp, vùng SSC dài 2.297 bp. Vùng SSC cũng chỉ có 1 gene *ndhF* với kích thước 1.182 bp.

Về trình tự nucleotide, kết quả so sánh cấu trúc hệ gene lục lạp của *H. verticillata* var. *verticillata* (trong nghiên cứu của chúng tôi) với cấu trúc hệ gene đã công bố của 29 loài khác [104], [119] thuộc chi *Hoya* cho thấy những điểm tương đồng cơ bản thể hiện ở thành phần, số lượng và sự phân bố các gene trong hệ gene. Điều này chứng tỏ mối quan hệ chặt chẽ giữa các loài thuộc chi *Hoya* (Hình 3.5). Theo Rodda và Niissalo (2021), trong các công bố trước đây về hệ gene lục lạp của nhóm *Hoya* còn chưa đầy đủ, hoặc hệ gene lục lạp được lắp ráp không chính xác, việc lắp ráp hệ gene lục lạp ở nhóm *Hoya* phức tạp hơn do sự mở rộng của hai vùng IR và sự thu hẹp của vùng SSC. Trong nghiên cứu này, trình tự gene và cấu trúc tổng thể của hệ gene lục lạp của *H. verticillata* var. *verticillata* tương tự như hệ gene lục lạp của *H. carnosa* [167]. Tuy nhiên, có những điểm khác biệt so với hệ gene lục lạp của *H. verticillata* trong báo cáo của Tan và cs (2018), nghiên cứu đã bỏ qua các vùng CDS của các gene *accD*, *ndhH*, *ycf1* và *ycf2* và những thay đổi ranh giới giữa các vùng trong hệ gene lục lạp [146]. Rodda và Niissalo (2021) cho rằng sự thay đổi ranh giới quan sát được là duy nhất chỉ

có ở nhóm *Hoya* (Rodda và Niissalo 2021) và *H. verticillata* cũng không ngoại lệ. Kết quả nghiên cứu của chúng tôi và kết quả của Odago và Niissalo (2021) là khác nhau vì trong nghiên cứu của chúng tôi, các phân tích dựa trên hệ gene lục lạp hoàn chỉnh của *H. verticillata* var. *verticillata* (số truy cập: OR475244.1) [69], trong khi nghiên cứu của Odago và Niissalo, các phân tích dựa trên hệ gene lục lạp của *H. verticillata* chỉ được giải trình tự một phần (GenBank: MW719071.1). Cho đến nay, kết quả giải trình tự hệ gene lục lạp của Lan tai cáo trong nghiên cứu này là báo cáo đầy đủ đầu tiên được công bố [69].

Cấu trúc hệ gene lục lạp hoàn chỉnh của *H. verticillata* var. *verticillata* đã được làm sáng tỏ dựa trên kết quả giải trình tự gene và phân tích cấu trúc của hệ gene. Hệ gene lục lạp của *H. verticillata* var. *verticillata* có cấu trúc điển hình như hệ gene lục lạp của các loài thực vật bậc cao khác, cấu trúc dạng vòng gồm 4 vùng (Sao chép đơn lớn, sao chép đơn nhỏ và 2 vùng lặp lại đảo ngược). Hệ gene có sự mở rộng đáng kể của vùng IR và vùng SSC bị thu hẹp mạnh (chỉ còn lại một gene là *ndhF*). Cấu trúc hệ gene lục lạp hoàn chỉnh của *H. verticillata* var. *verticillata* cho thấy tần số sử dụng codon đồng nghĩa tương đối, các đoạn lặp lại trình tự đơn giản, trình tự lặp lại lớn và xác định được độ đa dạng nucleotide (Pi). Sáu vùng biến đổi cao thuộc về vùng LSC, một vùng nằm giữa LSC và IR, và gene *ndhF* thuộc về SSC, đây là cơ sở để nghiên cứu tiến hóa phân tử và xác định sự phát sinh chủng loại của loài *H. verticillata* var. *verticillata* và chi *Hoya*.

3.2. PHÂN TÍCH SỰ PHÁT SINH CHỦNG LOẠI DỰA TRÊN TRÌNH TỰ HỆ GENE LỤC LẠP VÀ MỘT SỐ VÙNG DNA THUỘC HỆ GENE LỤC LẠP

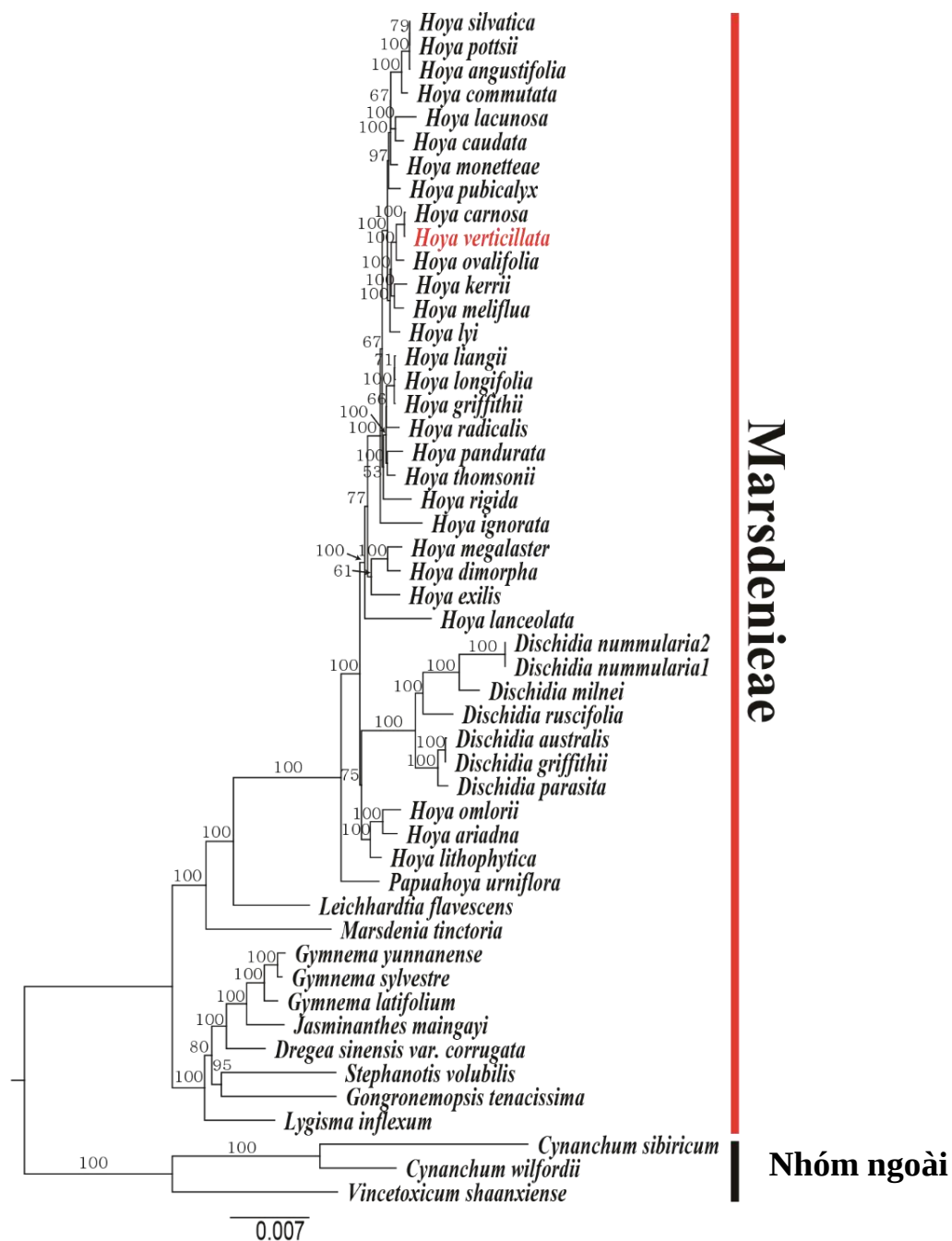
3.2.1. Phân tích phát sinh chủng loại dựa trên hệ gene lục lạp hoàn chỉnh

Trong nghiên cứu này, thay vì sử dụng toàn bộ trình tự hệ gene lục lạp, chúng tôi lựa chọn 79 gene mã hóa protein (protein-coding genes, PCGs) phổ biến và được bảo tồn cao để xây dựng cây phát sinh chủng loại. Cách tiếp cận này có cơ sở khoa học như sau: (i) PCG cung cấp tín hiệu phát sinh chủng loại mạnh mẽ và đáng tin cậy, đồng thời có chú thích chức năng rõ ràng; (ii) các vùng không mã hóa và khoảng cách giữa các gene thường chứa chèn/xóa, yếu tố lặp hoặc sai khác trong căn chỉnh, có thể tạo nhiễu

và làm giảm độ chính xác của phân tích; và (iii) bộ 79 PCG chuẩn hóa đã được áp dụng rộng rãi trong các nghiên cứu phát sinh loài ở thực vật, nhờ đó đảm bảo tính thống nhất và khả năng so sánh giữa các nghiên cứu. Do đó, việc sử dụng 79 PCG giúp nâng cao cả độ tin cậy lẫn tính ổn định của quá trình tái tạo quan hệ phát sinh loài.

Dữ liệu phục vụ phân tích được thu thập từ GenBank, bao gồm trình tự DNA lục lạp của 50 hệ gene thuộc chi *Hoya*, họ Apocynaceae, cùng một số loài nhóm ngoài (Phụ lục PL1). Cây phát sinh chủng loại được xây dựng (Hình 3.7) cho thấy hai nhánh chính: Nhánh thứ nhất gồm ba loài thuộc phân họ Asclepiadeae, trong khi nhánh thứ hai bao gồm các loài thuộc chi *Hoya* cùng một số chi khác trong tông Marsdenieae. Đáng chú ý, các loài *Hoya* được chia thành hai cụm: Một cụm lớn gồm 26 loài và một cụm nhỏ gồm ba loài. Trong cụm lớn, *H. verticillata* var. *verticillata* (Lan tai cáo) hình thành một nhóm đơn ngành, có quan hệ gần gũi với các loài *Hoya* khác và nằm cùng nhánh với *H. carnosa*, với giá trị bootstrap đạt 100%.

Để so sánh, phân biệt và xác định chính xác nguồn gốc của các loài trong chi *Hoya*, nhóm nghiên cứu đã bổ sung 3 loài khác là *Vincetoxicum shaanxiense*, *Cynanchum wilfordii* và *Cynanchum sibiricum*; nhánh tách ra sớm nhất là *Lygisma inflexum*, nhánh tách ra sớm thứ hai gồm 7 loài (*Gymnema yunnanense*, *Gymnema sylvestre*, *Gymnema latifolium*; *Jasminathes maingayi*, *Dregea sinensis* var. *corrugata*, *Stephanotis volubilis*, *Gongronemopsis tenacissima*), trong đó *Gymnema yunnanense*, *Gymnema sylvestre*, *Gymnema latifolium* có quan hệ gần gũi thể hiện giá trị bootstrap = 100; nhánh thứ ba gồm 3 loài là *Marsdenia tinctoria*, *Leichhardtia flavesceus*, *Papuahoya urniflora*; nhánh thứ tư gồm 10 loài và phân làm 2 lớp, *Dischidia* 7 loài (*Dischidia nummularia* 2, *Dischidia nummularia* 1, *Dischidia milnei*, *Dischidia rusifolia*, *Dischidia australis*, *Dischidia griffithii*, *Dischidia erticil*) và *Hoya* với 3 loài (*H. omlorii*, *H. ariadna*, *H. lithophytica*); nhánh thứ năm lớn nhất gồm có 26 loài *Hoya* (*H. silvatica*, *H. pottsii*, *H. angustifolia*, *H. cummutata*, *H. lacunosa*, *H. caudata*, *H. monetteae*, *H. pubicalyx*, *H. carnosa*, *H. verticillata* var. *verticillata*, *H. ovalifolia*, *H. kerrii*, *H. meliflua*, *H. lyi*, *H. liangii*, *H. longifolia*, *H. griffithii*, *H. radicalis*, *H. pandurata*, *H. thomsonii*, *H. rigida*, *H. ignorata*, *H. megalaster*, *H. dimorpha*, *H. exilis*, *H. lanceolata*).



Hình 3.7. Sơ đồ phát sinh chủng loại của 50 loài thuộc họ Apocynaceae dựa trên 79 trình tự gene mã hóa protein. Trình tự được căn chỉnh có chiều dài 75.252 bp. Cây xác suất tối đa (ML) được xác định bằng phần mềm RaxML với giá trị $-\ln L = 196716.416607$. Giá trị của các nút trên biểu thị giá trị bootstrap. Các loài *Cynanchum sibiricum*, *Cynanchum wilfordii* và *Vincetoxicum shaanxiense*, thuộc họ Asclepiadeae, được đặt làm nhóm ngoài (không thuộc tông Marsdenieae).

Lan tai cáo (*H. verticillata* var. *verticillata*) có đặc điểm giống và nằm chung nhánh với *H. carnosa*, có quan hệ gần gũi với *H. ovalifolia*, giá trị bootstrap là 100; trong khi đó *H. lyi*, *H. meliflua*, *H. kerrii* ở một nhánh riêng. Trong cây phát sinh chủng loại cho thấy ba loài *H. omlorii*, *H. ariadna*, *H. lithophytica* phân bố ở nhóm cận đơn ngành, chứng tỏ hệ gene của chúng đang trong quá trình tiến hóa.

Trong nghiên cứu này, hệ gene lục lạp hoàn chỉnh của Lan tai cáo (*H. verticillata* var. *verticillata*) thu thập tại Việt Nam đã được xác định với kích thước 176.429 bp, bao gồm có 4 vùng cấu trúc điển hình: LSC, SSC và hai vùng IR. Cấu trúc và số lượng gene của hệ gene lục lạp của các loài *Hoya* khác trong tông Marsdenieae (họ Apocynaceae) tương tự nhau và có tính bảo tồn cao. Hệ gene có sự mở rộng của hai vùng IR và sự thu hẹp của vùng SSC. Vùng gene *ndhF* có giá trị Pi cao là ứng cử viên mã vạch DNA tiềm năng để hỗ trợ nhận diện chi *Hoya* và phân biệt *H. verticillata* var. *verticillata* với các loài khác trong chi *Hoya*. Thông tin hệ gene lục lạp của *H. verticillata* var. *verticillata* và các phân tích so sánh trong nghiên cứu này rất hữu ích cho các nghiên cứu hệ gene lục lạp của các loài thuộc chi *Hoya* và khám phá các mối quan hệ phát sinh chủng loại trong họ Apocynaceae.

3.2.2. Phân tích phát sinh chủng loại dựa trên một số vùng DNA của hệ gene lục lạp

Dữ liệu hệ gene lục lạp của loài Lan tai cáo (*H. verticillata* var. *verticillata*) được lưu giữ trên GenBank mang mã truy cập OR475244.1. Trong nghiên cứu này, một số gene và vùng đệm giữa các gene của của hệ gene *H. verticillata* var. *verticillata* có giá trị đa dạng nucleotide (Pi) cao như *ndhF* (vùng SSC), *trnK-rps16*, *rps16-trnQ*, *psbI-atpA*, *psbZ-rps14*, *ndhC-trnV*, *rbcL-accD* (vùng LSC) được sử dụng để phân tích sự phát sinh chủng loại. Đây là những vùng DNA tiềm năng có thể được sử dụng để lựa chọn làm mã vạch DNA. Trình tự nucleotide của vùng gene *ndhF* và các vùng đệm *trnK-rps16*, *rps16-trnQ*, *ndhC-trnV*, *psbI-atpA*, *psbZ-rps14* và *rbcL-accD* của hệ gene lục lạp được nhận diện và tải từ dữ liệu hệ gene lục lạp của Lan tai cáo trên GenBank.

Sử dụng BLAST để phân tích sự tương đồng đối với mỗi trình tự vùng gene và vùng đệm của *H. verticillata* var. *verticillata* với các loài khác trên GenBank. Chọn các trình tự có tổng điểm và độ che phủ cao làm dữ liệu cho phân tích phát sinh chủng loại.

3.2.2.1. Phát sinh chủng loại dựa trên một số vùng gene

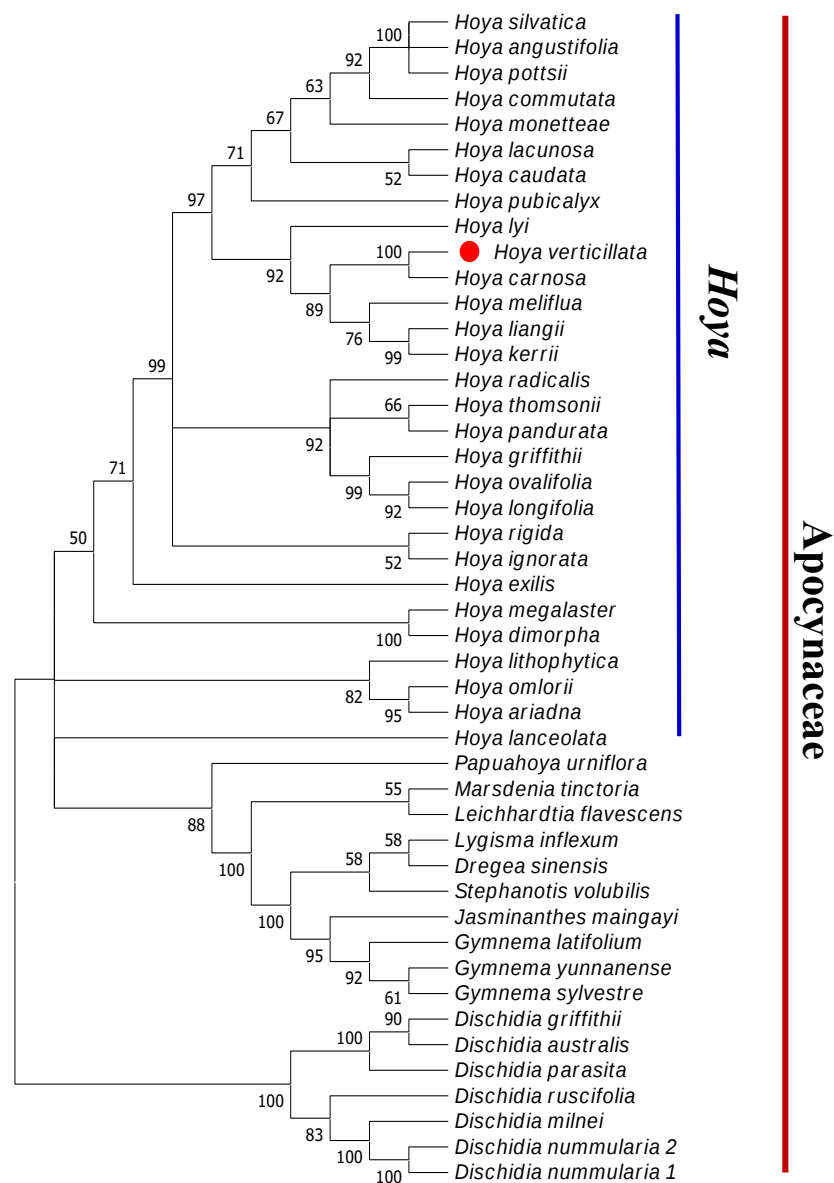
Trong số các vùng gene có giá trị Pi cao, *ndhF*, *rpl16* và *matK* (mã vạch DNA phổ biến cho nhiều loài thực vật) đã được chọn để phân tích phát sinh chủng loại. Các trình tự vùng gene *ndhF*, *matK* và *rpl16* được tải từ trình tự DNA lục lạp của Lan tai cáo (Mã số truy cập trên GenBank: OR475244.1). Trên cơ sở dữ liệu trình tự gene *ndhF* của các loài trên GenBank, cây phát sinh chủng loại dựa trên trình tự gene *ndhF* được thiết lập ở Hình 3.8.

Trên Hình 3.8, Lan tai cáo (*H. verticillata* var. *verticillata*) và *H. carnosa* được phân cùng nhóm với giá trị bootstrap là 100%. Đối chiếu với Hình 3.7, trình tự hệ gene lục lạp và vùng gene *ndhF* của Lan tai cáo và *H. carnosa* của chi *Hoya* phân bố trong một nhóm có giá trị bootstrap là 100%. Ngược lại với trình tự vùng gene *ndhF*, tập dữ liệu gene *matK* và *rpl16* của Lan tai cáo và *H. carnosa* (OK539590.1) có độ phân giải phát sinh chủng loại thấp, với giá trị bootstrap là 28-80% (*matK*) và 30-77% (*rpl16*) (Phụ lục PL3 và PL4). Vì vậy, trình tự gene *ndhF* có độ phân giải phát sinh chủng loại cao có thể được sử dụng để hỗ trợ phân biệt Lan tai cáo với các loài khác thuộc chi *Hoya*.

Việc xây dựng phát sinh chủng loại ở thực vật dựa trên hệ gene lục lạp hoàn chỉnh đã cung cấp thông tin đầy đủ và có tính thuyết phục trong việc phân loại ở các cấp phân loại thấp hơn [57]. Tuy nhiên, đối với mỗi vùng DNA, nghiên cứu cũng chỉ ra sự khác biệt về số lượng và thành phần nucleotide để thấy được các vùng gene trong hệ gene *Hoya* đang biến đổi. Vì vậy, các vùng gene hoặc vùng đệm của DNA lục lạp có khả năng biến đổi cao có thể được sử dụng làm chỉ thị DNA để hỗ trợ phân biệt các loài khác nhau [100], [102].

3.2.2.2. Phát sinh chủng loại dựa trên một số vùng đệm

Dữ liệu về trình tự của 6 vùng đệm giữa các gene (*trnK-rps16*, *rps16-trnQ*, *ndhC-trnV*, *psbI-atpA*, *psbZ-rps14*, *rbcL-accD*) (Phụ lục PL5) đã được sử dụng làm dữ liệu để phân tích sự phát sinh chủng loại bằng MEGA11 [144], theo phương pháp Maximum Likelihood (ML). Lịch sử tiến hóa được suy luận bằng phương pháp khả năng tối đa và mô hình Tamura-Nei [143] và kết quả phân tích phát sinh chủng loại với độ tin cậy cao nhất được suy ra từ 1000 lần lặp lại đã thể hiện lịch sử tiến hóa của các đơn vị phân loại được phân tích [44].

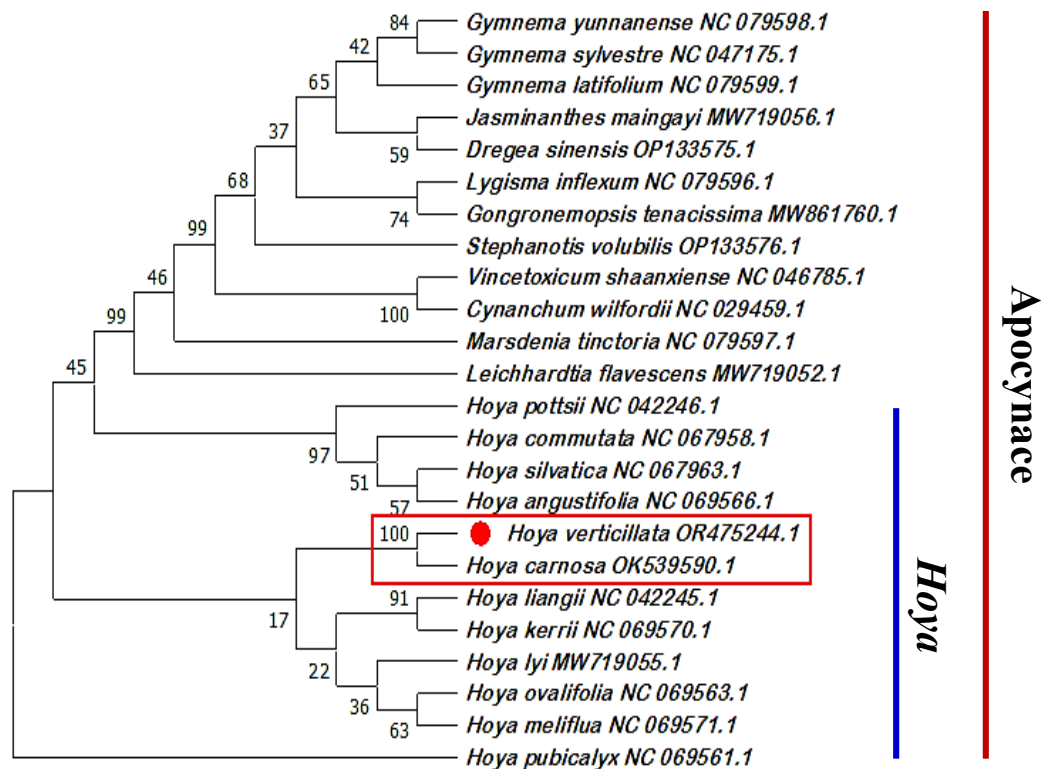


Hình 3.8. Mối quan hệ phát sinh chủng loại được suy ra bằng phương pháp Maximum Likelihood dựa trên trình tự vùng gene *ndhF*. Lan tai cáo (*H. verticillata* var. *verticillata*) và *H. carnosa* nhóm phân bố trong cùng một nhóm với hỗ trợ bootstrap mạnh (100%). Các loài thuộc chi *Marsdenia*, *Lygisma*, *Leichhardtia*, *Dregea*, *Stephanotis*, *Jasminanthes*, *Gymnema* và *Dischidia* của họ Apocynaceae được đặt làm nhóm ngoài. Các số trên mỗi nút là giá trị hỗ trợ bootstrap.

Sơ đồ phát sinh chủng loại được xây dựng dựa trên trình tự vùng đệm của 24 loài (*trnK-rps16*), 44 loài (*rps16-trnQ*), 47 loài (*ndhC-trnV*), 46 loài (*psbI- atpA*), 33 loài (*psbZ-rps14*), và 34 loài (*rbcL-accD*) (Phụ lục PL6).

Phân tích phát sinh chủng loại dựa trên trình tự vùng đệm *trnK-rps16*

Sử dụng vùng đệm tham chiếu *trnK-rps16* với 793 bp trong phân tích BLAST trực tuyến trên NCBI đã thu được 24 trình tự được lựa chọn để phân tích sự phát sinh chủng loại với tổng số điểm từ 848-1465 (đối với nhóm *Hoya*) và từ 889-1243 (nhóm ngoài *Hoya*), độ che phủ là 100%. Độ tương đồng 98,55-100% (Nhóm *Hoya*) và 87,77-100% (nhóm ngoài *Hoya*). Cây phát sinh chủng loại được thiết lập dựa trên trình tự vùng đệm *trnK-rps16* nhờ phần mềm MEGA11 cho thấy 12 loài thuộc chi *Hoya* có quan hệ rất gần với 10 loài thuộc các chi khác (nhóm ngoài) như *Leichhardtia*, *Marsdenia*, *Cynanchum*, *Vincetoxicum*, *Stephanotis*, *Gongronemopsis*, *Lygisma*, *Dregea*, *Jasminanthes*, *Gymnema*. Trong nhóm *Hoya*, loài *Lan* tai cáo phân bố cùng nhóm với loài *H. carnosa*, giá trị bootstrap là 100% (Hình 3.9).



Hình 3.9. Cây phát sinh chủng loại sử dụng vùng đệm *trnK-rps16* của hệ gene lục lạp *H. verticillata* var. *verticillata* và của các loài khác thuộc họ Apocynaceae

Phân tích phát sinh chủng loại dựa trên trình tự vùng đệm rps16-trnQ

Sử dụng vùng đệm tham chiếu *rps16-trnQ* với 1328 bp trong phân tích BLAST trực tuyến trên NCBI đã thu được 44 trình tự được lựa chọn để phân tích sự phát sinh chủng loại với tổng số điểm từ 1443-2453 (đối với nhóm *Hoya*) và từ 889-1243 (nhóm ngoài *Hoya*), độ che phủ là 87-100%. Độ tương đồng 98,98-100% (Nhóm *Hoya*) và 91,43-94,35% (nhóm ngoài *Hoya*). Cây phát sinh chủng loại được thiết lập dựa trên trình tự vùng đệm *rps16-trnQ* cho thấy 21 loài thuộc chi *Hoya* có quan hệ rất gần với 23 loài thuộc các chi khác (nhóm ngoài) như *Leichhardtia*, *Marsdenia*, *Cynamchum*, *Vincetoxicum*, *Stephanotis*, *Gongronemopsis*, *Lygisma*, *Dregea*, *Jasminanthes*, *Gymnema*. Trong nhóm *Hoya*, loài *Lan tai cáo* phân bố cùng nhóm với loài *H. carnosa*, giá trị bootstrap là 99% (Hình 3.10).

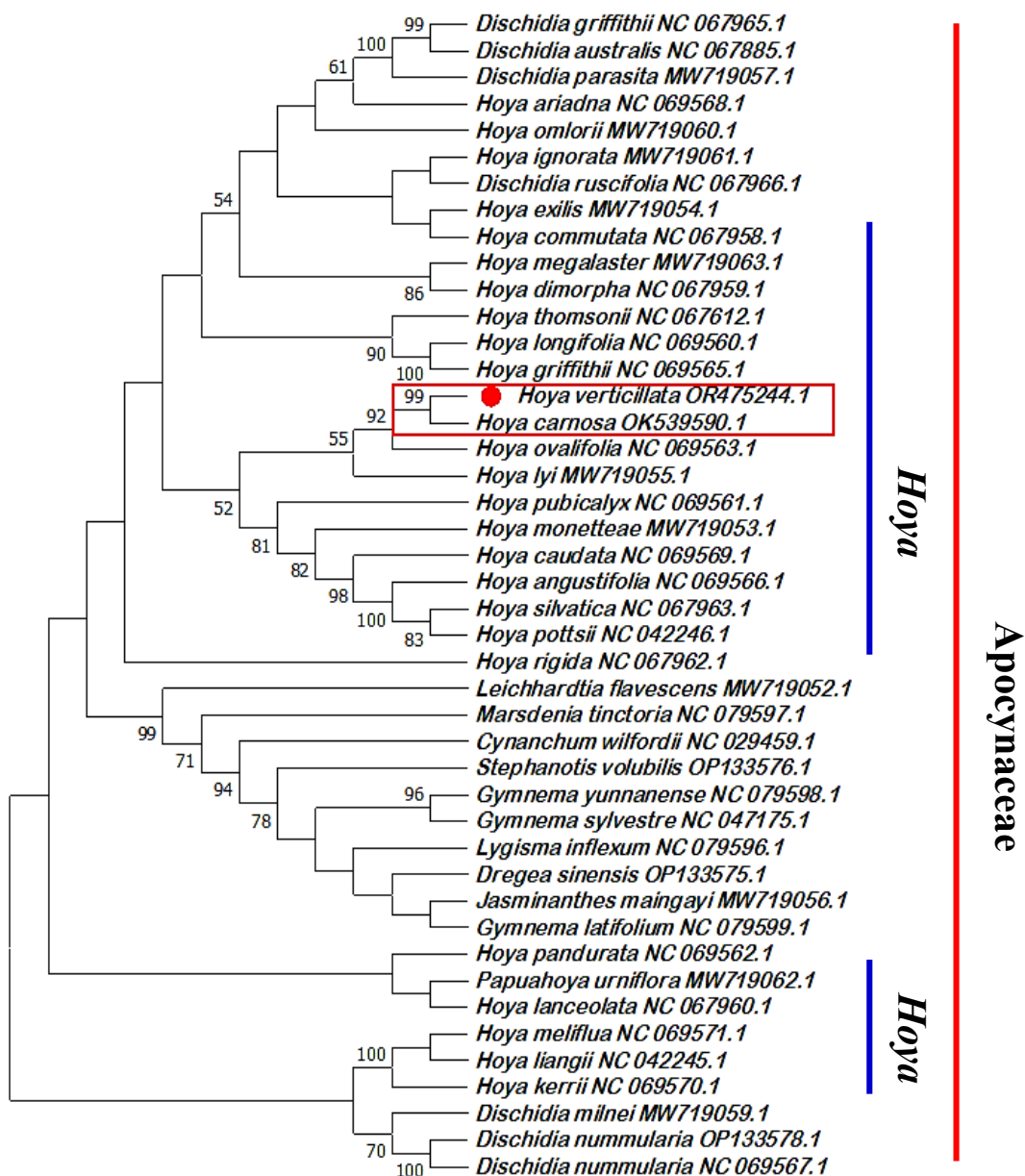
Phân tích phát sinh chủng loại dựa trên trình tự vùng đệm ndhC-trnV

Sử dụng vùng đệm tham chiếu *ndhC-trnV* với 1492 bp trong phân tích BLAST trực tuyến trên NCBI đã thu được 47 trình tự được lựa chọn để phân tích sự phát sinh chủng loại với tổng số điểm từ 867-2756 (đối với nhóm *Hoya*) và từ 774-1613 (nhóm ngoài), độ che phủ là 84-100%. Độ tương đồng 94,6-100% (Nhóm *Hoya*) và 88,79-93.75% (nhóm ngoài). Cây phát sinh chủng loại được thiết lập dựa trên trình tự vùng đệm *ndhC-trnV* cho thấy 29 loài thuộc chi *Hoya* có quan hệ rất gần với 19 loài thuộc các chi khác (nhóm ngoài) như *Leichhardtia*, *Dischidia*, *Marsdenia*, *Cynamchum*, *Stephanotis*, *Gongronemopsis*, *Lygisma*, *Dregea*, *Jasminanthes*, *Gymnema*. Trong nhóm *Hoya*, loài *Lan tai cáo* phân bố cùng nhóm với loài *H. carnosa*, giá trị bootstrap là 99% (Hình 3.11).

Phân tích phát sinh chủng loại dựa trên trình tự vùng đệm psbI-atpA

Sử dụng vùng đệm tham chiếu *psbI-atpA* với 1878 bp trong phân tích BLAST trực tuyến trên NCBI đã thu được 46 trình tự được lựa chọn để phân tích phát sinh chủng loại với tổng số điểm từ 2285-3469 (đối với nhóm *Hoya*) và từ 2279-3022 (nhóm ngoài), độ che phủ là 98-100%. Độ tương đồng 97,27-100% (Nhóm *Hoya*) và 95,65-97,12% (nhóm ngoài). Cây phát sinh chủng loại được thiết lập dựa trên trình tự vùng đệm *psbI-*

atpA cho thấy 24 loài thuộc chi *Hoya* có quan hệ rất gần với 12 loài thuộc các chi khác (nhóm ngoài) như *Leichhardtia*, *Dischidia*, *Marsdenia*, *Stephanotis*, *Lygisma*, *Dregea*, *Jasminanthes*, *Gymnema*, *Vincetoxicum*. Trong nhóm *Hoya*, loài Lan tai cáo phân bố cùng nhóm với loài *H. carnosa*, giá trị bootstrap là 100% (Hình 3.12).



Hình 3.10. Cây phát sinh chủng loại sử dụng vùng đệm *rps16-trnQ* của hệ gene lục lạp *H. verticillata* var. *verticillata* và của các loài khác thuộc họ Apocynaceae

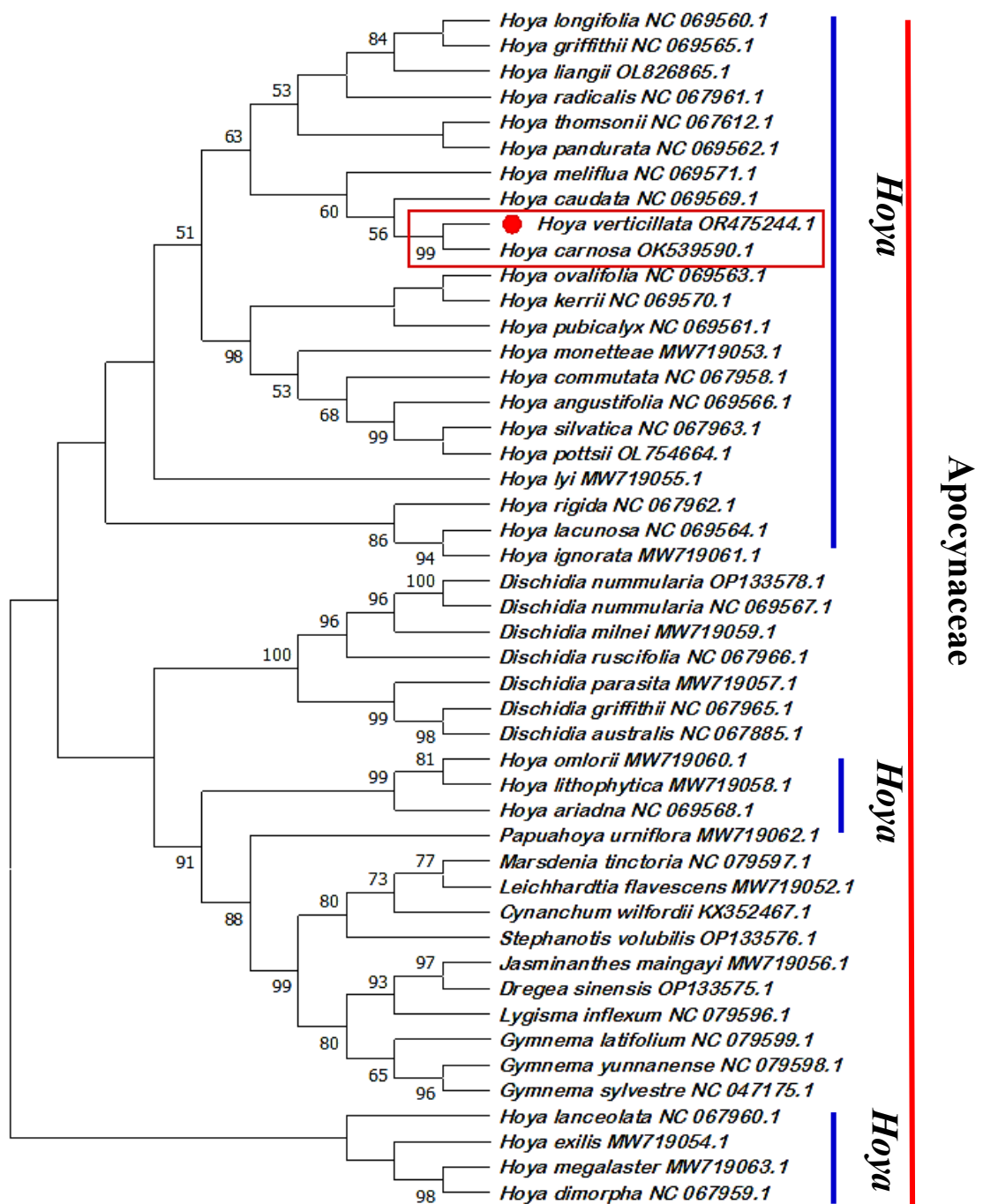
Phân tích phát sinh chủng loại dựa trên trình tự vùng đệm *rbcL-accD*

Sử dụng BLAST phân tích tương đồng giữa vùng đệm *rbcL-accD* trong hệ gene lục lạp của loài Lan tai cáo (*H. verticilata* var. *verticillata*) với các trình tự trên GenBank kết quả đã thu được 100 trình tự, trong đó có 34 trình tự được lựa chọn để phân tích sự phát sinh chủng loại với tổng điểm là 1274 - 1818 (nhóm *Hoya*) và 716 - 1092 (nhóm ngoài), độ che phủ là 98 - 100% (nhóm *Hoya*) và 77 - 82%. Cây phát sinh chủng loại được thiết lập dựa trên 34 trình tự vùng đệm *rbcL-accD* cho thấy 26 loài thuộc chi *Hoya* có quan hệ gần với 8 loài thuộc các chi khác (nhóm ngoài) như *Leichhardtia*, *Marsdenia*, *Vincetoxicum*, *Lygisma*, *Gymnema*, *Stephanotis*, *Dregea*, *Synanchum* với giá trị bootstrap là 77%. Trong nhóm *Hoya*, loài Lan tai cáo có quan hệ rất gần với loài *H. carnos*a (mã số NC_045868.1), giá trị bootstrap là 96%, với loài *H. ovalifolia* (NC_069563.1), giá trị bootstrap là 100% (Hình 3.13).

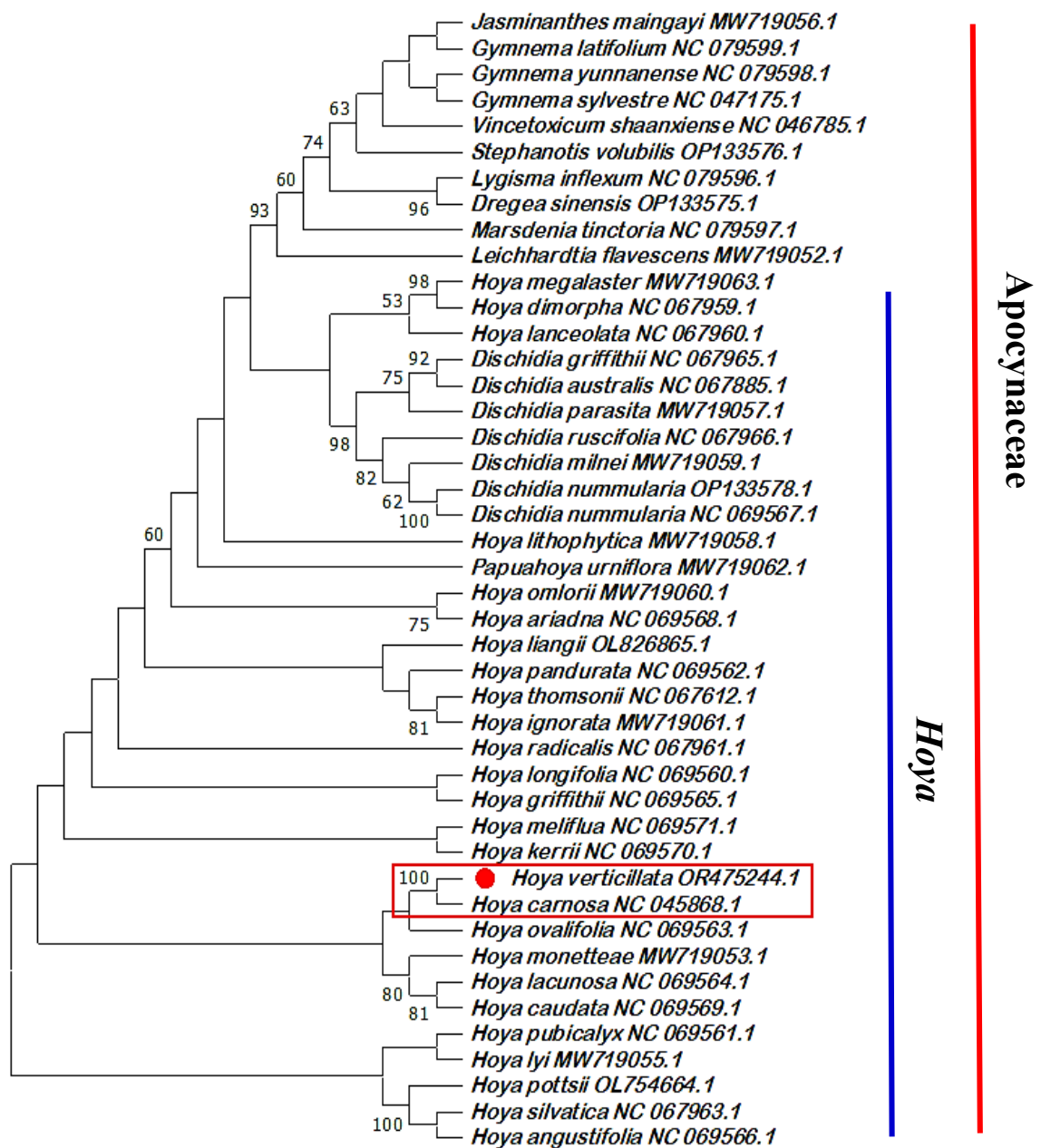
Phân tích phát sinh chủng loại dựa trên trình tự vùng đệm *psbZ-rps14*

Phân tích phát sinh chủng loại dựa trên trình tự các vùng đệm giữa các gene *psbZ-rps14* cho thấy *H. verticilata* var. *verticillata* có liên quan rất chặt chẽ với nhau và cùng nhóm với *H. carnos*a với giá trị bootstrap 76%. Trong khi đó kết quả phân loại các loài thuộc chi khác nhau như: *Stephanotis*, *Jasminanthes*, *Gymnema*, *Lygisma*, *Dregea*, *Vincetoxicum*, *Marsdenia*, *Leichhardtia*, *Dischidia*, *Hoya* cho giá trị bootstrap thấp, do đó độ tin cậy không cao (Phụ lục PL7).

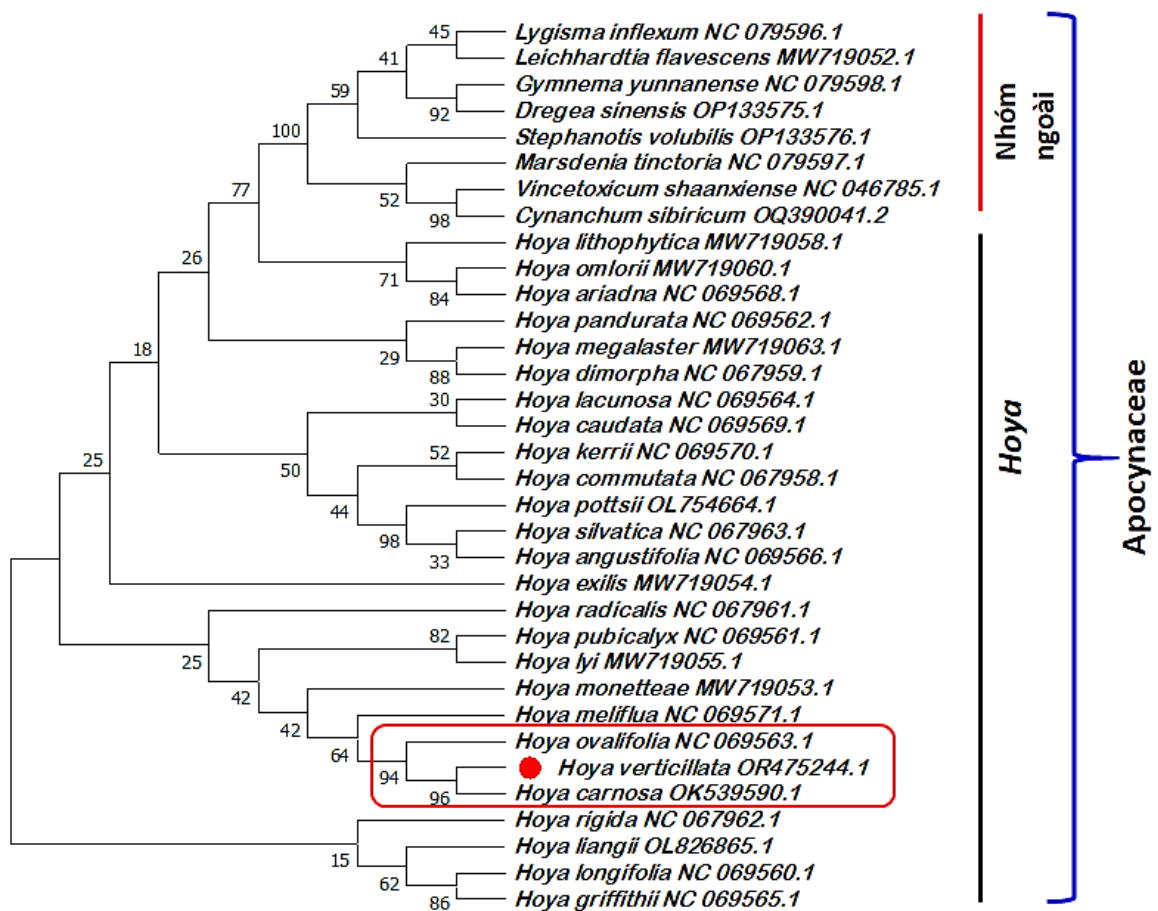
Như vậy, kết quả phân tích sự phát sinh chủng loại dựa trên 6 vùng đệm (*trnK-rps16*, *rps16-trnQ*, *psbI-atpA*, *ndhC-trnV*, *psbZ-rps14*, *rbcL-accD*) đã cho thấy Lan tai cáo (*H. verticillata* var. *verticillata*) có quan hệ họ hàng rất gần và phân bố trong cùng nhóm với *H. carnos*a với giá trị Bootstrap là 75-100%. Trong số các trình tự vùng đệm trên, 5 trình tự *trnK-rps16*, *rps16-trnQ*, *psbI-atpA*, *ndhC-trnV* và *rbcL-accD* là những chỉ thị DNA tiềm năng có thể sử dụng để hỗ trợ phân biệt các loài thuộc chi *Hoya*.



Hình 3.11. Cây phát sinh chủng loại sử dụng vùng đệm *ndhC-trnV* của hệ gene lục lạp *H. verticillata* var. *verticillata* và của các loài khác thuộc họ Apocynaceae



Hình 3.12. Cây phát sinh chủng loại dựa trên vùng đệm *psbI-atpA* của hệ gene lục lạp *H. verticillata* var. *verticillata* và của các loài khác thuộc họ Apocynaceae



Hình 3.13. Cây phát sinh chủng loại dựa trên vùng đệm *rbcL-accD* của hệ gene lục lạp *H. verticillata* var. *verticillata* và của các loài khác thuộc họ Apocynaceae

Tóm lại, kết quả phân tích phát sinh chủng loại dựa trên 3 vùng gene và 6 vùng đệm có hệ số đa dạng nucleotide cao của hệ gene lục lạp Lan tai cáo (*H. verticillata* var. *verticillata*) đã xác định được vùng gene *ndhF* và các vùng đệm *trnK-rps16*, *rps16-trnQ*, *psbI-atpA*, *ndhC-trnV*, *rbcL-accD* là những ứng cử viên mã vạch tiềm năng có thể sử dụng để hỗ trợ phân biệt các loài thuộc chi *Hoya*, phân biệt chi *Hoya* với các chi khác thuộc họ Apocynaceae. Như vậy, ngoài những gene đã được chứng minh là những ứng cử viên tiềm năng sử dụng để hỗ trợ phân biệt loài là *matK*, *rbcL* thì nghiên cứu này đã bổ sung thêm vùng gene *ndhF* và các vùng đệm *trnK-rps16*, *rps16-trnQ*, *psbI-atpA*, *ndhC-trnV*, *rbcL-accD*. Tuy nhiên, cần có những thực nghiệm tiếp theo để xác nhận các trình tự này là mã vạch DNA để có thể sử dụng trong thực tiễn.

3.3. PHÂN TÍCH HOẠT TÍNH SINH HỌC CỦA SAPONIN PHÂN LẬP TỪ LAN TAI CÁO

3.3.1. Thành phần hóa học của Lan tai cáo

Cao chiết tổng số thu được từ lá Lan tai cáo được sử dụng trong thí nghiệm khảo sát thành phần hóa học bằng thuốc thử (Phụ lục PL8), kết quả đã xác định được 5 hợp chất là flavonoid, steroid, coumarin, saponin và tannin (Bảng 3.5).

Bảng 3.5. Kết quả định tính các nhóm hợp chất hữu cơ có trong Lan tai cáo

STT	Nhóm chất	Phản ứng định tính	Kết quả	Nhận xét
1	Alkaloid	Thuốc thử Dragendorff	-	Không có
		Thuốc thử Wagner	-	
		Thuốc thử Mayer	-	
2	Flavonoid	Phản ứng Shinoda	+	Có
3	Steroid	Phản ứng Liebermann-Burchard	+	Có
4	Glycoside tim	Phản ứng Keller-Kilian	-	Không có
5	Coumarin	NaOH 10%	+	Có
6	Saponin	Phản ứng tạo bọt	+	Có
7	Tannin	FeCl ₃ 5%.	+	Có

3.3.2. Hoạt tính kháng khuẩn

Cao chiết tổng số từ Lan tai cáo được thử nghiệm đánh giá hoạt tính sinh học thông qua phân tích khả năng kháng khuẩn. Các thang nồng độ cao chiết được sử dụng là 25 mg/mL, 50 mg/mL, 100 mg/mL bằng phương pháp khuếch tán giếng thạch trên 4 chủng vi khuẩn (*E. coli*, *S. aureus*, *P. aeruginosa*, *B. subtilis*). Sử dụng kháng sinh amoxicillin 50 mg/mL làm đối chứng. Kết quả phân tích hoạt tính kháng vi khuẩn của cao chiết Lan tai cáo được thể hiện ở Phụ lục PL9 và Bảng 3.6.

Bảng 3.6. Hoạt tính kháng khuẩn của cao chiết tổng số từ Lan tai cáo

Chủng	Amoxicillin (50 mg/mL)	Cao chiết Lan tai cáo		
		100 mg/mL	50 mg/mL	25 mg/mL
<i>Vì khuẩn</i>	<i>Đường kính kháng khuẩn (mm)</i>			
<i>E. coli</i>	20,00 ^b ± 0,58	27,00 ^c ± 0,58	18,67 ^b ± 0,88	15,00 ^a ± 0,58
<i>S. aureus</i>	21,00 ^c ± 0,58	19,00 ^b ± 0,58	19,00 ^b ± 0,58	14,00 ^a ± 0,58
<i>P. aeruginosa</i>	21,00 ^c ± 0,58	22,67 ^c ± 0,88	18,00 ^b ± 0,58	15,00 ^a ± 0,58
<i>B. subtilis</i>	24,33 ^b ± 0,67	25,00 ^b ± 0,58	17,00 ^a ± 0,58	16,00 ^a ± 0,58

Ghi chú: Các chữ cái khác nhau trong cùng một dòng biểu thị sự khác biệt có ý nghĩa thống kê với $p < 0,05$.

Kết quả Bảng 3.6 cho thấy, hoạt tính kháng khuẩn của cao chiết Lan tai cáo phụ thuộc vào nồng độ. Nồng độ càng cao (100 mg/mL), vòng kháng khuẩn càng lớn. Phổ kháng khuẩn của cao chiết rộng, có tác dụng lên cả vi khuẩn Gram âm (*E. coli*, *P. aeruginosa*) và Gram dương (*S. aureus*, *B. subtilis*). Ở nồng độ 100 mg/mL, cao chiết có hiệu quả hơn kháng sinh Amoxicillin trong việc chống lại *E. coli* và có hiệu quả tương đương với kháng sinh để chống lại *P. aeruginosa* và *B. subtilis*. Hoạt tính kháng khuẩn của cao chiết thể hiện mạnh nhất đối với *E. coli* ($p < 0,05$).

3.3.3. Hoạt tính sinh học của các phân đoạn chứa saponin từ cao chiết Lan tai cáo

Trên cơ sở biểu hiện hoạt tính sinh học của cao chiết tổng số, tiến hành tách tiếp bằng sắc kí lỏng chân không thu được 3 phân đoạn (A, B, C), Phân đoạn B được tách bằng sắc kí lỏng áp suất trung bình được 5 phân đoạn (B1-B5); dựa trên kích thước phân tử tương ứng lựa chọn Phân đoạn B2 và Phân đoạn B3 để thử nghiệm đánh giá khả năng kháng oxy hóa, gây độc dòng tế bào ung thư và khả năng ức chế enzyme α -glucosidase.

3.3.3.1. Hoạt tính chống oxy hóa

Sử dụng 1,1-Diphenyl-2-picrylhydrazyl (DPPH), một gốc tự do có màu tím, có độ hấp thụ cực đại ở bước sóng 517 nm để phân tích hoạt tính chống oxy hóa. Khi có mặt chất chống oxy hóa, DPPH sẽ bị khử thành 2,2-Diphenyl-1-picrylhydrazine (DPPH-H) có màu

vàng. Sử dụng phương pháp đo mức độ giảm độ hấp thụ của DPPH ở bước sóng 517 nm để xác định khả năng khử của chất chống oxy hóa. Kết quả phân tích hoạt tính kháng oxy hóa thông qua phép thử trung hòa gốc tự do của DPPH được trình bày ở Bảng 3.7.

Bảng 3.7. Hoạt tính trung hòa gốc tự do (DPPH) của các phân đoạn chứa saponin

Nồng độ Phân đoạn 1 (µg/mL)	Tỷ lệ % trung hòa gốc tự do		Nồng độ Phân đoạn 2 (µg/mL)	% trung hòa gốc tự do của L- ascorbic acid
	Phân đoạn B2	Phân đoạn B3		
100	42,54 ± 1,83	60,28 ± 1,23	100	91,23 ± 1,58
20	15,74 ± 1,07	15,74 ± 1,05	20	73,88 ± 1,22
4	3,11 ± 0,34	4,75 ± 0,45	4	33,20 ± 1,29
0,8	0,55 ± 0,07	1,12 ± 0,11	0,8	12,60 ± 0,81
SC₅₀	>100	79,92 ± 2,63	SC₅₀	7,78 ± 0,38

Kết quả cho thấy ở nồng độ 100 µg/mL, phân đoạn 2 đạt 60,28% trong khi phân đoạn B2 chỉ đạt 42,54%. Giá trị SC₅₀ của phân đoạn 1 là >100 µg/mL, có nghĩa là ngay cả ở nồng độ cao nhất (100 µg/mL), mẫu này vẫn không thể đạt được hiệu quả trung hòa 50% do đó hoạt tính của nó rất yếu. Đối với phân đoạn B3 giá trị SC₅₀ là 79,92 ± 2,63 µg/mL, vậy hiệu quả trung hòa 50% gốc tự do ở nồng độ ~80 µg/mL. Như vậy, phân đoạn B3 có hoạt tính chống oxy hóa mạnh hơn so với phân đoạn 1. Đối chứng ascorbic acid (Vitamin C) hoạt động ổn định trong thí nghiệm.

3.3.3.2. Hoạt tính gây độc tế bào ung thư

Các phân đoạn chứa saponin được thử nghiệm hoạt tính gây độc tế bào với nồng độ tối đa 500 µg/mL trên 3 dòng tế bào ung thư ở người gồm: Ung thư biểu mô phổi (A549), ung thư biểu mô gan (HepG2), ung thư vú (MCF7) và ellipticine được sử dụng làm đối chứng.

Kết quả thí nghiệm được thể hiện trong Bảng 3.8 cho thấy cả 2 phân đoạn đã thể hiện hoạt tính ức chế sự phát triển của các dòng tế bào ung thư nghiên cứu với giá trị IC₅₀ từ 153,26 – 482,83 µg/mL (Bảng 3.8). Phân đoạn B3 có hoạt tính gây độc tế bào ung thư mạnh hơn phân đoạn B2, trong đó biểu hiện mạnh đối với hoạt tính gây độc tế bào ung thư biểu mô gan (HepG2) và ung thư vú (MCF7). Chất đối chứng dương ellipticine hoạt động ổn định trong thí nghiệm. Như vậy, cả hai phân đoạn đều thể hiện

hoạt tính gây độc 3 dòng tế bào ung thư được sử dụng nghiên cứu với giá trị IC_{50} từ 153,26 – 482,83 $\mu\text{g/mL}$, trong đó phân đoạn B3 thể hiện hoạt tính ức chế các dòng tế bào ung thư tốt hơn phân đoạn B2.

Bảng 3.8. Khả năng gây độc tế bào ung thư của các phân đoạn chứa saponin

Nồng độ phân đoạn B2 ($\mu\text{g/mL}$)	Tỷ lệ % gây độc tế bào của phân đoạn B2		
	A549	HepG2	MCF7
500	51,61 $\pm$ 1,12	63,03 $\pm$ 2,53	51,43 $\pm$ 1,78
200	25,32 $\pm$ 1,02	31,58 $\pm$ 1,31	25,49 $\pm$ 1,36
100	14,41 $\pm$ 2,29	22,27 $\pm$ 1,18	16,39 $\pm$ 0,86
20	8,83 $\pm$ 0,44	15,32 $\pm$ 2,16	6,85 $\pm$ 0,57
4	2,43 $\pm$ 0,18	7,00 $\pm$ 0,67	2,11 $\pm$ 0,12
IC_{50}	480,99 $\pm$ 12,62	379,03 $\pm$ 17,97	482,83 $\pm$ 21,81
Nồng độ phân đoạn B3 ($\mu\text{g/mL}$)	Tỷ lệ % gây độc tế bào của phân đoạn B3		
	A549	HepG2	MCF7
500	67,81 $\pm$ 1,82	97,07 $\pm$ 2,61	89,57 $\pm$ 1,65
200	31,22 $\pm$ 0,93	58,70 $\pm$ 2,16	51,93 $\pm$ 1,39
100	20,31 $\pm$ 1,02	38,63 $\pm$ 1,01	28,94 $\pm$ 1,41
20	7,17 $\pm$ 0,52	15,71 $\pm$ 0,86	10,33 $\pm$ 1,06
4	6,16 $\pm$ 0,48	1,56 $\pm$ 0,17	4,40 $\pm$ 0,41
IC_{50}	347,16 $\pm$ 10,52	153,26 $\pm$ 8,94	201,59 $\pm$ 8,71
Nồng độ ellipticine ($\mu\text{g/mL}$)	Tỷ lệ % gây độc tế bào của ellipticine		
	A549	HepG2	MCF7
10	94,17 $\pm$ 2,65	97,41 $\pm$ 2,09	99,05 $\pm$ 1,99
2	78,60 $\pm$ 2,24	79,22 $\pm$ 1,70	80,13 $\pm$ 0,99
0,4	50,24 $\pm$ 1,00	51,39 $\pm$ 0,95	50,18 $\pm$ 1,70
0,08	21,13 $\pm$ 0,91	23,19 $\pm$ 0,86	22,98 $\pm$ 1,41
IC_{50}	0,37 $\pm$ 0,02	0,34 $\pm$ 0,02	0,35 $\pm$ 0,03

3.3.3.3. Hoạt tính ức chế α -glucosidase

Trong cơ thể, màng tế bào ruột non tiết ra α -glucosidase thủy phân carbohydrate trong thức ăn thành các oligosaccharide, thủy phân oligosaccharide thành glucose và thẩm thấu vào máu qua màng ruột non để nuôi các tế bào của cơ thể. Khi cơ thể bị rối loạn chuyển hóa carbohydrate thì lượng đường trong máu cao sẽ dẫn đến bệnh đái tháo đường. Acarbose là một loại thuốc chống tiểu đường được sử dụng để điều trị bệnh tiểu đường typ 2, acarbose là chất ngăn chặn thủy phân tinh bột thông qua ức chế α -glucosidase.

Tiến hành sử dụng cao chiết của 2 phân đoạn nghiên cứu ở những nồng độ khác nhau 4 $\mu\text{g/mL}$, 20 $\mu\text{g/mL}$, 100 $\mu\text{g/mL}$, 500 $\mu\text{g/mL}$ nhằm đánh giá khả năng ức chế α -glucosidase với tham chiếu là chất đối chứng acarbose. Kết quả thí nghiệm được trình bày ở Bảng 3.9.

Bảng 3.9. Tác động ức chế α -glucosidase của các phân đoạn chứa saponin

Nồng độ ($\mu\text{g/mL}$)	Phân đoạn B2	Phân đoạn B3	Acarbose
500	$17,42 \pm 1,41$	$97,85 \pm 2,01$	$77,77 \pm 2,39$
100	$9,30 \pm 0,56$	$33,98 \pm 1,93$	$51,72 \pm 1,05$
20	$5,59 \pm 0,36$	$10,25 \pm 1,09$	$15,51 \pm 0,95$
4	$3,82 \pm 0,03$	$6,93 \pm 0,21$	$6,91 \pm 0,18$
IC ₅₀	> 500	$186,57 \pm 11,35$	$122,80 \pm 6,40$

Kết quả trên cho thấy phân đoạn B2 chưa thể hiện hoạt tính ức chế 50% α -glucosidase (IC₅₀) ngay cả với nồng độ 500 $\mu\text{g/mL}$; phân đoạn B3 đã thể hiện hoạt tính ức chế α -glucosidase với IC₅₀ = $186,57 \pm 11,35$ $\mu\text{g/mL}$; chất đối chứng dương acarbose hoạt động ổn định trong thí nghiệm.

Như vậy, các phân đoạn chứa saponin tách chiết từ lá Lan tai cáo (cao tổng số) đều có hoạt tính kháng khuẩn, ức chế α -glucosidase, kháng oxy hóa, gây độc tế bào ung thư, trong đó hoạt tính của phân đoạn B3 biểu hiện hoạt tính mạnh hơn phân đoạn B2. Các phân đoạn này được sử dụng để phân tích xác định chất có hoạt tính sinh học mạnh là cơ sở đề xuất sử dụng chất trong y học.

3.3.4. Phân lập saponin

3.3.4.1. Chiết xuất và phân lập saponin

Bột lá khô (53,6 g) *H. verticillata* var. *verticillata* được tách chiết bằng phương pháp có sự hỗ trợ siêu âm sử dụng dung môi EtOH/H₂O (75:25) ở điều kiện 200 W, 60°C, 45 phút, thu được 500 mL cho mỗi lần tách chiết (có 2 lần tách chiết). Cao chiết tổng số được tách tiếp bằng VLC silica gel RP-18 và rửa giải lần lượt bằng H₂O 100%, EtOH/H₂O 50:50, EtOH 100%, thu được 3 phân đoạn (A-C) dựa trên phân tích TLC. Phân đoạn B (5,60 g) được tách phân đoạn bằng sắc ký MPLC silica gel 60 (CHCl₃/MeOH/H₂O 80:20:2, 75:25:3, v/v/v) thu được 5 phân đoạn (B1-B5).

Phân đoạn B2 (356,5 mg) được sắc ký bằng MPLC trên silica gel 60 (CHCl₃/MeOH/H₂O 80:20:2, v/v/v) thu được 4 phân đoạn (B2.1-B2.4). Phân đoạn B2.2 (42,3 mg) tiếp tục được tách phân đoạn bằng sắc ký MPLC trên silica gel 60 (CHCl₃/MeOH/H₂O 75:25:3, v/v/v) thu được 5 phân đoạn (B2.2.1-B2.2.5). Phân đoạn B2.2.3 (15,3 mg) tiếp tục được tách phân đoạn bằng sắc ký MPLC trên silica gel 60 (CHCl₃/MeOH/H₂O 75:25:3, v/v/v) thu được Hợp chất 1 (3,3 mg).

Tiếp tục tách phân đoạn B3 (255,6 mg) bằng sắc ký MPLC trên silica gel 60 (CHCl₃/MeOH/H₂O 75:25:3, v/v/v) thu được 4 phân đoạn (B3.1-B3.4). Phân đoạn B3.3 (85,2 mg) được tách phân đoạn bằng sắc ký MPLC trên silica gel 60 (CHCl₃/MeOH/H₂O 75:25:3, v/v/v) thu được 5 phân đoạn (B3.3.1-B3.3.5). Phân đoạn B3.3.3 (29,8 mg) được tách bằng sắc ký MPLC trên silica gel 60 (CHCl₃/MeOH/H₂O 70:30:5, v/v/v) thu được Hợp chất 2 (3,1 mg).

3.3.4.2. Thủy phân acid và sắc kí khí

Quá trình thủy phân 3 mg mỗi glycoside mới được thực hiện với 5 mL dung dịch 2N.CF₃COOH (4 mL) và nước (1 mL) trong 3 giờ ở 95°C. Hỗn hợp phản ứng sau đó được pha loãng trong nước (10 mL) và được chiết bằng 5 mL CH₂Cl₂ (3 lần). Lóp CH₂Cl₂ được làm bay hơi bằng MeOH cho đến khi dung môi trở nên trung tính để tạo ra cặn đường và được sắc ký tiếp bằng TLC trên silica gel sử dụng hệ rửa giải CHCl₃/MeOH/H₂O tỷ lệ 8:5:1. Hỗn hợp monosacharide được pha loãng với 100 µL pyridin khan và xử lý với 0,06 mol/L HSCH₂CH(NH₂)COOCH₃·HCl ở 60°C trong 1

giờ, sau đó thêm 150 μL HMDS-TMCS (tỷ lệ 3:1) ở 60°C trong 30 phút. Sau khi dừng phản ứng tiến hành tách thêm 2 mL H_2O , phần đường còn lại được chiết bằng 0,1 mL n-hexan. Lớp n-hexan được thu hồi và sau đó phần đường được phân tích (1 μL) bằng sắc ký khí (GC). Việc xác định đường được thực hiện bằng cách so sánh thời gian lưu của chúng với thời gian lưu của đường tiêu chuẩn (Sigma-Aldrich) [96]. Thời gian lưu (phút): 11,07 (rhamnose), 13,82 (fucose) và 17,51 (glucose).

Trên cơ sở hoạt tính sinh học của cao chiết Lan tai cáo, tiến hành chiết phân đoạn để phân lập chất mới. Lấy kết quả sắc ký của phân đoạn B2, phân đoạn B3 tiếp tục chiết phân đoạn và tiến hành đo bằng máy phổ hạt nhân cộng hưởng từ hạt nhân được 2 hợp chất mới là:

Parasiticoside A (Hợp chất 1): Bột trắng vô định hình, $[\alpha]_{25} = -38^\circ$ (MeOH, c 0,30). HR-ESI-MS: m/z 649.3563 $[\text{M}+\text{Na}]^+$ (được xác định là $\text{C}_{33}\text{H}_{54}\text{NaO}_{11}$, 649.3558). Dữ liệu NMR ^1H và ^{13}C được đưa ra trong Bảng 3.10.

Parasiticoside B (Hợp chất 2): Bột trắng vô định hình, $[\alpha]_{25} = -45^\circ$ (MeOH, c 0,22). HR-ESI-MS: m/z 795.4142 $[\text{M}+\text{Na}]^+$ (được xác định là $\text{C}_{39}\text{H}_{64}\text{NaO}_{15}$, 795.4137). Dữ liệu NMR ^1H và ^{13}C được nêu trong Bảng 3.10.

Bảng 3.10. Dữ liệu phổ ^{13}C và ^1H NMR của hợp chất 1 và 2 (dung môi pyridine- d_5 , δ tính bằng ppm, J tính bằng Hz)

1			2		
Position	δC	δH	Position	δC	δH
1	37,5	1,06 m; 1,62 m	1	37,4	1,08 m; 1,66 m
2	31,3	1,70 m; 2,14 m	2	32,3	1,75 m; 2,18 m
3	78,1	3,93 ov	3	77,9	3,97 ov
4	39,3	2,28 m; 2,72 m	4	39,2	2,40 m, 2,82 m
5	140,9	-	5	140,9	-
6	125,0	5,67 br s	6	125,0	5,72 br s
7	32,1	1,40 m; 1,95 m	7	32,7	1,43 m; 1,97 m
8	31,8	1,30 ov	8	32,2	1,35 ov
9	50,3	0,80 m	9	50,3	0,87 m
10	36,9	-	10	36,3	-
11	21,0	1,27 m; 1,45 m	11	23,9	1,18 m; 1,48 m
12	39,1	1,00 m; 1,82 m	12	39,3	1,11 m; 1,83 m
13	40,5	-	13	40,6	-
14	56,7	1,08 ov	14	56,7	1,09 ov
15	28,0	1,91 m; 2,40 m	15	28,0	1,92 m; 2,48 m
16	26,4	1,05 m; 1,53 m	16	26,8	1,09 m; 1,54 m

17	62,4	1,60 ov	17	62,3	1,64 ov
18	16,9	0,85 s	18	16,9	0,91 s
19	16,0	1,28 s	19	15,9	1,26 s
20	81,3	3,98 ov	20	81,2	4,01 ov
21	23,5	1,12 d (6,0)	21	23,6	1,28 d (6,0)
Fuc-1	99,8	4,78 d (7,6)	Fuc-1	99,9	4,80 d (7,6)
2	74,2	4,52 dd (9,4; 7,6)	2	74,4	4,53 dd (9,4; 7,6)
3	76,6	4,08 dd (9,4; 2,9)	3	76,4	4,12 dd (9,4; 2,9)
4	72,9	3,93 br d (2,9)	4	73,1	3,96 br d (2,9)
5	70,8	3,62 br q (6,4)	5	70,9	3,68 br q (6,4)
6	17,0	1,46 d (6,4)	6	16,8	1,49 d (6,4)
Glc-1	103,4	4,90 d (7,6)	Glc-1	103,5	4,88 d (7,6)
2	74,7	4,05 ov	2	74,6	4,06 ov
3	78,0	4,26 ov	3	77,9	4,28 ov
4	72,2	4,18 ov	4	72,3	4,18 ov
5	78,1	3,96 m	5	78,1	3,96 m
6	62,4	4,33 m; 4,54 m	6	62,3	4,30 m; 4,52 m
			Rha-1	101,2	6,35 br s
			2	72,1	4,79 br s
			3	72,2	4,61 dd (9,4; 3,5)
			4	73,8	4,35 dd (9,9; 9,4)
			5	69,0	4,83 dq (9,9; 5,9)
			6	18,7	1,78 d (5,9)

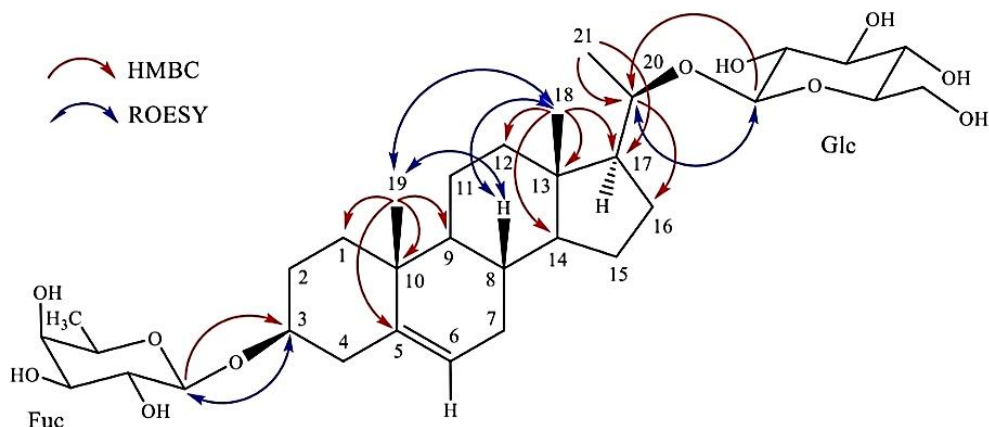
ov: Overlapped signal (tín hiệu chồng chất) không xác định được bội số rõ ràng;
br s: Broad singlet (tín hiệu đơn rộng) - thường gặp ở proton anomeric của một số đường
hoặc proton olefinic; m: Multiplet (tín hiệu bội phức tạp); d, t, dd, dq, ...: Các bội số đã
xác định được, với hằng số coupling J (Hz) trong ngoặc đơn.

3.3.4.3. Xác định cấu trúc saponin

Quá trình thủy phân bằng acid và phân tích phổ cộng hưởng từ hạt nhân 2D mở rộng của các Hợp chất **1** và **2** đã cho thấy các gốc đường được xác định là D-fucopyranosyl (Fuc), D-glucopyranosyl (Glc) cho **1** và D-fucopyranosyl (Fuc), D-glucopyranosyl (Glc) và α -rhamnopyranosyl (Rha) cho **2**. Các monosacharide, bao gồm cả cấu hình tuyệt đối của chúng, được xác định bằng phân tích sắc kí khí là D cho Fuc và Glc cho **1**, D cho Fuc và Glc, L cho Rha của hợp chất **2** (thời gian lưu là 11,07 phút đối với Rha, 13,82 phút đối với Fuc và 17,51 phút đối với Glc) [96]. Định hướng di thường β của Fuc và đơn vị Glc được xác định bằng cách tính toán các giá trị $3J_{H-1,H-2}$ tương đối lớn ở 7,6 Hz [97], [98]. Dạng pyranoid α xích đạo được đặt thành đơn vị Rha được chứng minh bằng $3J_{H-1,C-1}$ lớn, giá trị ở 166 Hz [99].

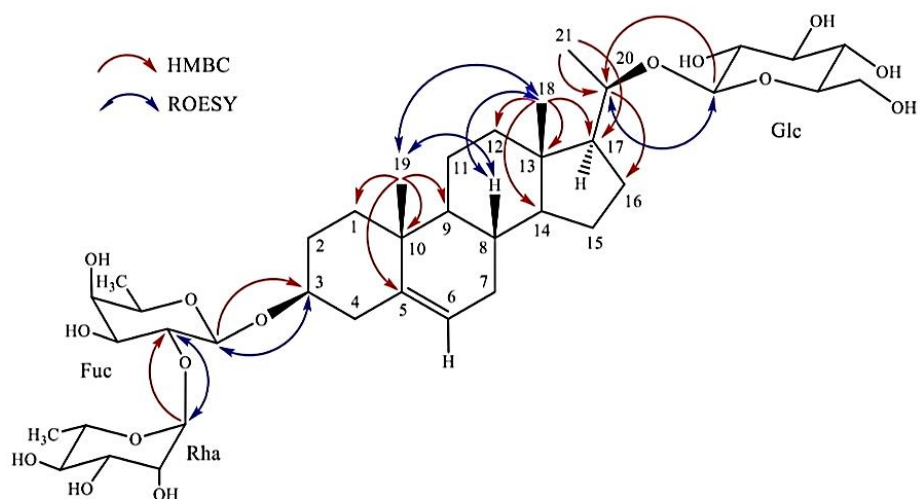
Hợp chất **1** hiển thị đỉnh ion giả phân tử ở m/z 649,3563 $[M+Na]^+$ (calcd. 649,3558) trong phổ HR-ESI-MS (phương pháp quang phổ khối ion hóa phun điện tử độ phân giải cao để đo tỷ lệ oxygen và carbon trung bình trong các sol khí hữu cơ thứ cấp High-Resolution Electrospray Ionization Mass Spectrometry), do đó công thức phân tử được ấn định là $C_{33}H_{54}O_{11}$ với trọng lượng phân tử là 626. Phổ 1H NMR của **1** cho thấy tín hiệu của 3 nhóm methyl ở δH 0,85 (s, H-18), 1,28 (s, H-19) và 1,12 (d, $J = 6,0$ Hz, H-21), 1 proton olefinic ba thế ở δH 5,67 (br s, H-6) và 2 proton dị thường ở δH 4,78 (d, $J = 7,6$ Hz) và 4,90 (d, $J = 7,6$ Hz). Phổ ^{13}C NMR của **1** hiển thị cộng hưởng từ của 33 carbon được phân loại thành 4 methyl ở δC 16,9 (Me-18), 16,0 (Me-19), 23,5 (Me-21) và 17,0 (C'-6), 9 methylene (37,5; 31,3; 39,3; 32,1; 21,0; 39,1; 28,0; 26,4; 62,4), 17 methine (78,1; 125,0; 31,8; 50,3; 56,7; 62,4; 81,3; 99,8; 74,2; 76,6; 72,9; 70,8; 103,4; 74,7; 78,0; 72,2; 78,1) và 3 nguyên tử carbon bậc bốn (140,9; 36,9; 40,5). Những tín hiệu này phù hợp tốt với những tín hiệu của diglycoside mới như được mô tả trong dữ liệu. Các đỉnh chéo ROESY của H-8 với H3-19 và H3-18 cho thấy các nhóm này là đồng mặt và được xác định là hướng β . Sự vắng mặt của các mối tương quan ROESY giữa H-8 và H-9/H-14/H-17 đã thiết lập định hướng α của các proton này. Cấu hình của C-20 không thể xác định được trên phổ ROESY do liên kết C-17 – C-20 quay tự do.

Trong vùng đường, tín hiệu của hai đỉnh chéo trong phổ HSQC ở $\delta H/\delta C$ 4,78 (d, $J = 7,6$ Hz)/99,8 và 4,90 (d, $J = 7,6$ Hz)/103,4, cho thấy sự hiện diện của hai đơn vị đường tương ứng là Fuc và Glc (Bảng 3.10). Đơn vị Fuc được xác định là có liên kết với vị trí C-3 của aglycone bằng tương quan HMBC ở $\delta H/\delta C$ 4,78 (d, $J = 7,6$ Hz, Fuc H-1)/78,1 (C-3), cùng với đỉnh chéo ROESY giữa δH 4,78 (d, $J = 7,6$ Hz, Fuc H-1) và δH 3,93 (m, H-3). Sự liên kết của đơn vị Glc với vị trí C-20 của aglycone được chứng minh bằng tương quan HMBC ở $\delta H/\delta C$ 4,90 (d, $J = 7,6$ Hz, Glc H-1)/81,3 (C-20) và tương quan không gian ROESY ở $\delta H/\delta H$ 4,90 (d, $J = 7,6$ Hz, Glc H-1)/3,98 (m, H-20). Do đó, cấu trúc của **1** được xác định là 3 β ,20-dihydroxy-pregn-5-ene-3-O- β -D-fucopyranoside 20-O- β -D-glucopyranoside và được đặt tên là parasiticoside A (Hình 3.14).



Hình 3.14. Cấu trúc liên kết (HMBC) và cấu hình không gian (ROESY) của hợp chất **1**

Phổ HR-ESI-MS của hợp chất **2** cho đỉnh ion giả phân tử $[M+Na]^+$ ở m/z 795,4142 (calcd. 795,4137), do đó công thức phân tử được xác định là $C_{39}H_{64}O_{15}$ với trọng lượng phân tử là 772. Sự khác biệt giữa trọng lượng phân tử của **2** và **1** là 146 amu, có thể tương ứng với nhóm 6-deoxyhexosyl bổ sung. Khi so sánh tín hiệu ^{13}C NMR của **2** (Bảng 3.10), một tập hợp các tín hiệu bổ sung, tương ứng với đơn vị α -L-rhamnopyranosyl cuối cùng đã xuất hiện. Liên kết glycoside giữa đơn vị Fuc và vị trí C-3 của aglycone được xác định bằng tương quan HMBC ở $\delta H/\delta C$ 4,78 (d, $J = 7,6$ Hz, Fuc H-1)/78,1 (C-3) và tương quan ROESY ở $\delta H/\delta H$ 4,80 (d, $J = 7,6$ Hz, Fuc H-1)/3,97 (m, H-3). Đơn vị Rha đã được chứng minh là gắn ở vị trí C-2 của đơn vị Fuc bằng tương quan HMBC giữa δH 6,35 (br s, Rha H-1) và δC 74,4 (Fuc C-2) và tương quan ROESY giữa δH 6,35 (br s, Rha H-1) và δH 4,53 (dd, $J = 9,4, 7,6$ Hz, Fuc H-2). Glc đã được chứng minh là có liên kết với vị trí C-20 của aglycone bằng tương quan liên kết HMBC ở $\delta H/\delta C$ 4,88 (d, $J = 7,6$ Hz, Glc H-1)/81,2 (C-20) và tương quan không gian ROESY ở $\delta H/\delta H$ 4,88 (d, $J = 7,6$ Hz, Glc H-1)/4,01 (m, H-20). Tất cả các bằng chứng trên đã chứng tỏ **2** là 3 β ,20-dihydroxy-pregn-5 ene-3-O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-fucopyranoside-20-O- β -D-glucopyranoside, được đặt tên là parasiticoside B (Hình 3.15).



Hình 3.15. Cấu trúc liên kết (HMBC) và cấu hình không gian (ROESY) của hợp chất 2

Kết quả phân tích sắc kí và quang phổ cộng hưởng từ hạt nhân xác định hợp chất 1 là 3 β ,20-dihydroxy-pregn-5-ene-3-O- β -D-fucopyranoside 20-O- β -D-glucopyranoside và được đặt tên là parasiticoside A, hợp chất 2 là 3 β ,20-dihydroxy-pregn-5 ene-3-O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-fucopyranoside-20-O- β -D-glucopyranoside, được đặt tên là parasiticoside B. Kết quả tìm kiếm trên thư viện dữ liệu cho thấy đây là 2 hợp chất mới. Việc xác định cấu trúc phân tử của 2 chất mới là cơ sở để xây dựng mô hình cấu trúc phân tử sử dụng trong các mô hình thực nghiệm.

3.3.5. Nghiên cứu tương tác phân tử và động học phân tử *in silico* của saponin

3.3.5.1. Nghiên cứu tương tác phân tử *in silico*

Các nghiên cứu về các hợp chất hữu cơ thực vật trên một số loài thuộc chi *Hoya* đã cho thấy sự hiện diện của triterpenoid và saponin steroid có hoạt tính sinh học như chất kháng oxy hóa, gây độc tế bào ung thư. Là một tác nhân chống khối u, saponin steroid có nhiều cơ chế phân tử khác nhau, saponin steroid có thể gây apoptosis thông qua việc ức chế protein chống apoptosis Bcl-2. Các nghiên cứu *in silico*, bao gồm mô phỏng quá trình tương tác phân tử được thực hiện trên saponin steroid được phân lập từ *H. verticillata* var. *verticillata* như chất ức chế protein chống apoptosis Bcl-2 trong ung thư đại tràng để xác định cơ chế về hiệu lực chống ung thư của chúng trước khi tiến hành các thực nghiệm trong phòng thí nghiệm. Nghiên cứu mô phỏng quá trình gắn kết phân tử của 5 hợp chất tách chiết từ *H. verticillata* var. *verticillata* gồm Hợp chất 1

(Parasiticoside A), Hợp chất 2 (Parasiticoside B), Hợp chất 3 (3 β ,4 α -dihydroxy-5 β -spirost-(25)27-en-1 β -yl O- α -L-arabinopyranoside), Hợp chất 4 (3 β -hydroxy-5 β -spirost-25(27)-en-1 β -yl O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-xylopyranoside), Hợp chất 5 ((23S,24S,25S)-3 β ,4 α ,23,24-tetrahydroxy-5 β -spirostan-1 β -yl O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-fucopyranoside) và chất đối chứng là Navitoclax.

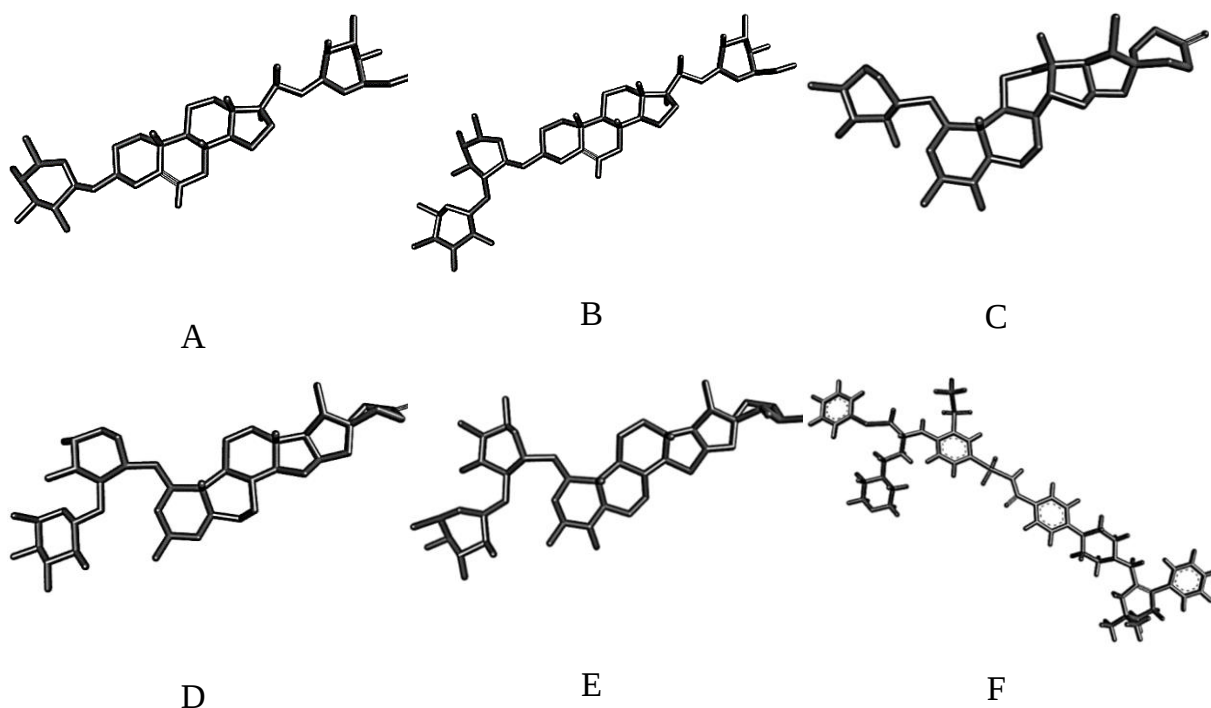
Quá trình tương tác phân tử được thực hiện bằng cách ghép nối mô hình phân tử saponin phân lập từ *H. verticillata* var. *verticillata* (trong các nghiên cứu trước đây và của nghiên cứu này) với mô hình của thụ thể 6GL8 (Hình 3.16).

Mỗi phức hợp tương tác được xác định dựa vào các giá trị năng lượng liên kết tối thiểu (kcal/mol). Mô phỏng tương tác của các phân tử với 6GL8 để hình thành phức hợp liên kết được thể hiện trong Bảng 3.11. Các gốc amino acid và vị trí tham gia vào các tương tác của chúng trong phức hợp đã được xác định. Liên kết hydrogen, tương tác Van der Waals, tương tác kỵ nước giữa protein và các phân tử đã được chứng minh thông qua tương tác phân tử. Mỗi tương tác phân tử cụ thể với các amino acid của 6GL8 đã được chứng minh trong Hình 3.17.

Bảng 3.11. Tương tác phân tử giữa saponin thử nghiệm với protein 6GL8

Hợp chất	Năng lượng liên kết (kcal/mol)	Liên kết hydrogen	Tương tác Van der Waals		Tương tác kỵ nước
1	-8.28	Ala149, Arg107, Leu137, Tyr108	Glu152, Phe112, Phe153, Val156	Gly145, Phe150, Tyr202,	Arg146, Met115, Phe104, Val133, Val148
2	-8.87	Arg107, Phe104, Tyr202, Val133	Ala100, Asp111, Gly145, Met115, Phe153, Val148	Arg146, Glu136, Leu137, Phe115,	Ala149, Tyr108
3	-9.29	Ala149, Arg146, Tyr108	Arg129, Glu152, Leu137, Phe112	Gln118, Leu119, Phe104,	His120, Met115, Phe153, Val133

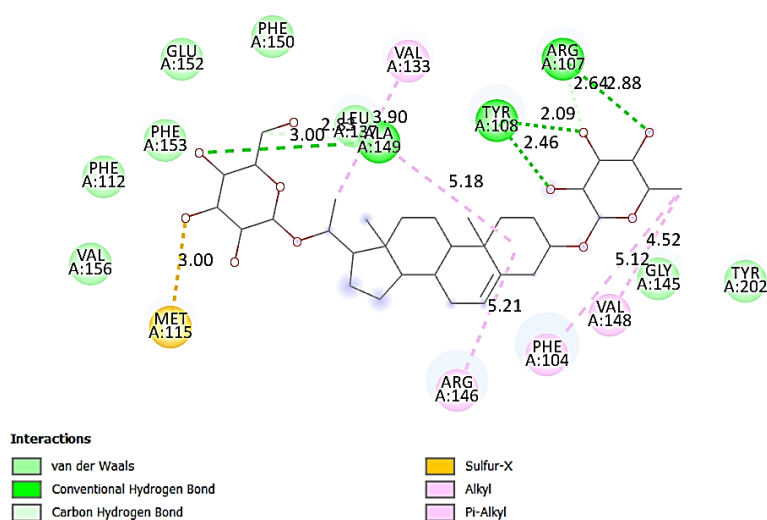
4	-11.09	Ala149, Gln118	Arg129, Asp111, Glu152, Phe112, Phe153	Arg146, Glu136, Phe104,	His120, eu119, Leu137, et115, Val133
5	-11.16	Arg146, Asp140, Gln118, Glu136, Tyr108	Ala149, Asp111, Met115, Phe112, Val133	Arg139, Glu152, Phe104, Phe152,	Leu137
Navitoclax	-8.34	Gly203, Trp144, Tyr202	Ala100, Arg146, Gln118, Gly145, Phe153, Tyr108	Arg107, Asn143, Glu136, Leu201,	Ala149, Asp103, Leu137, Phe104, Val133, Val148



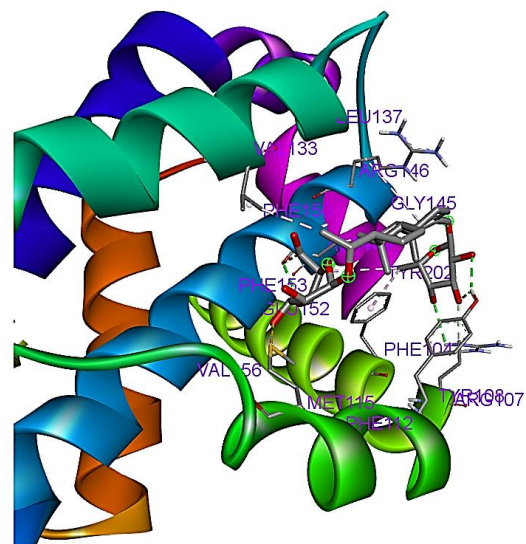
Hình 3.16. Cấu trúc 3D của các phân tử thử nghiệm. Các phân tử thử nghiệm gồm 05 saponin: (A) Hợp chất 1; (B) Hợp chất 2; (C) 3 β ,4 α -dihydroxy-5 β -spirost-(25)27-en-1 β -yl O- α -L-arabinopyranoside (3); (D) 3 β -hydroxy-5 β -spirost-25(27)-en-1 β -yl O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-xylopyranoside (4); (E) (23S,24S,25S)-3 β ,4 α ,23,24-

tetrahydroxy-5 β -spirostan-1 β -yl O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-fucopyranoside (5); và chất đối chứng là (F) Navitoclax.

Hợp chất **1**, **2**, **3**, **4**, **5** và Navitoclax thể hiện tương tác phân tử yếu không giới hạn như tương tác Van der Waals. Các liên kết của Hợp chất **1** gồm tương tác Van der Waals với các gốc amino acid Glu152, Gly145, Phe112, Phe150, Phe153, Tyr202 và Val156 của 6GL8. Đối với Hợp chất **2** là Ala100, Arg146, Asp111, Glu136, Gly145, Leu137, Met115, Phe115, Phe153 và Val148. Hợp chất **3** cho thấy tương tác Van der Waals với các gốc amino acid Arg129, Gln118, Glu152, Leu119, Leu137, Phe104 và Phe112. Hợp chất **4** tương tác với 6GL8 thông qua tương tác Van der Waals với các gốc amino acid Arg129, Arg146, Asp111, Glu136, Glu152, Phe104, Phe112, Phe153. Hợp chất **5** là Ala149, Arg139, Asp111, Glu152, Met115, Phe104, Phe112, Phe152 và Val133. Navitoclax là phân tử so sánh, thể hiện tương tác Van der Waals đối với các gốc amino acid là Ala100, Arg107, Arg146, Asn143, Gln118, Glu136, Gly145, Leu201, Phe153 và Tyr108. Các tương tác Van der Waals là loại liên kết đóng vai trò quan trọng trong việc duy trì cấu trúc của các phân tử sinh học và đảm bảo chức năng của các phân tử này, góp phần phân biệt các cấu trúc tinh thể của hợp chất, xác định sự được phức hợp được hình thành và mức độ ổn định của phức hợp đó [22], [116].

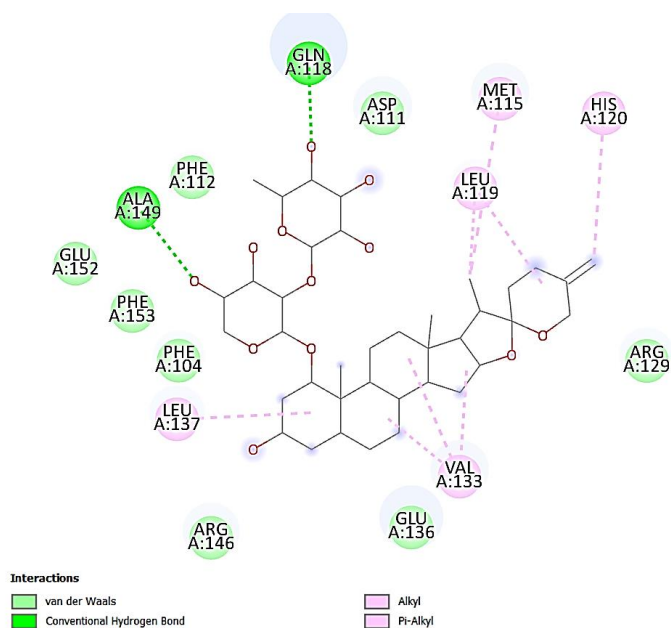


A-1

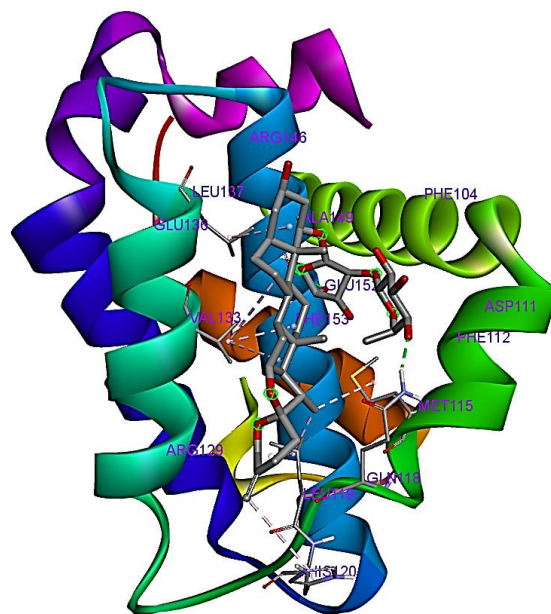


A-2

A: Mô hình tương tác phân tử 2D và 3D của Hợp chất **1** với thụ thể 6CL8

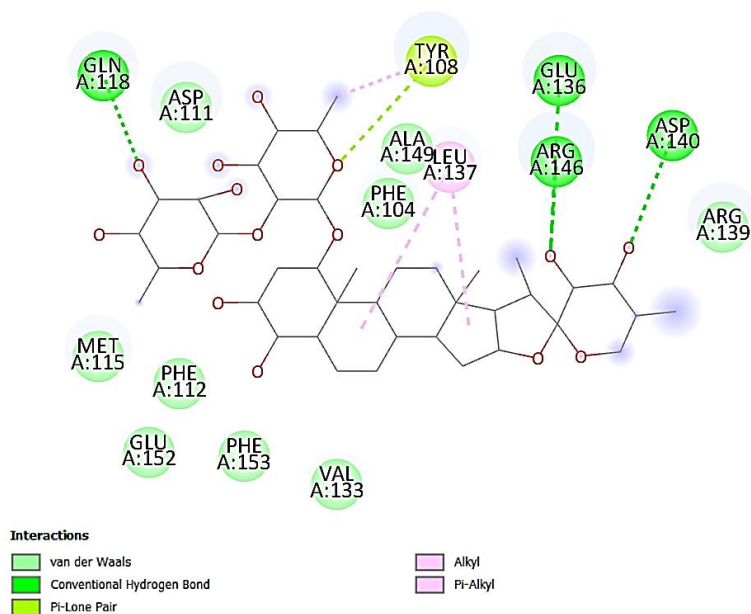


D-1

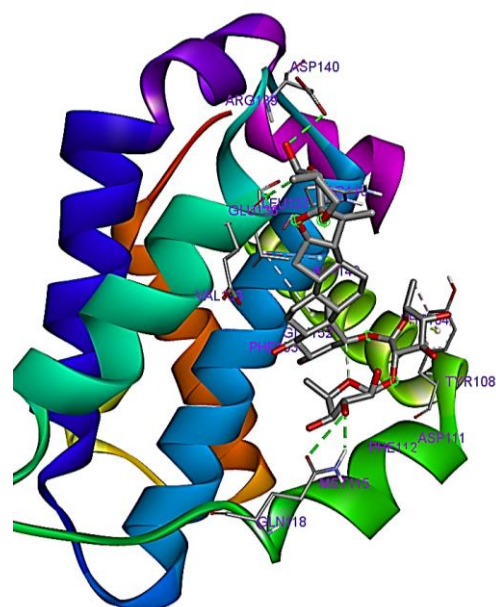


D-2

D: Mô hình tương tác phân tử 2D và 3D của Hợp chất **4** với thụ thể 6CL8

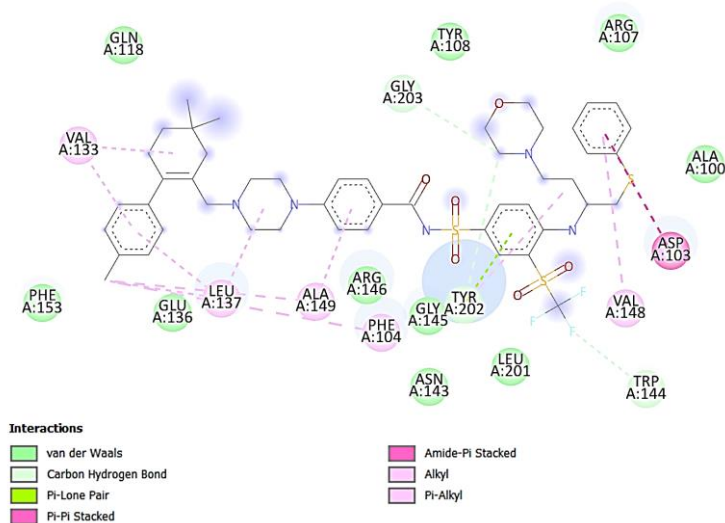


E-1

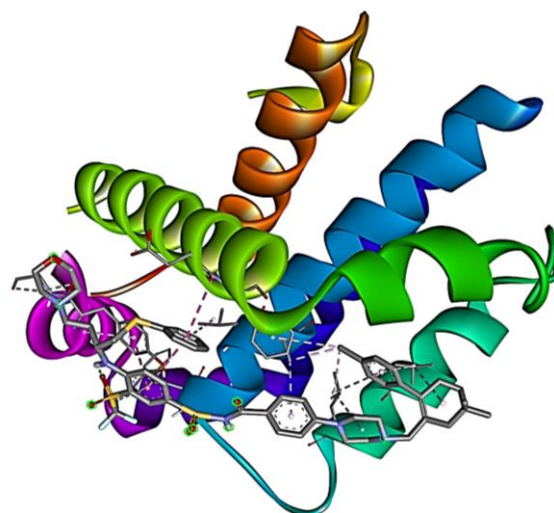


E-2

E: Mô hình tương tác phân tử 2D và 3D của Hợp chất **5** với thụ thể 6CL8



F-1



F-2

F: Mô hình tương tác phân tử 2D và 3D của Navitoclax với thụ thể 6CL8

Hình 3.17. Mô hình tương tác phân tử 2D 3D của Hợp chất 1 (A-1, A-2), Hợp chất 2 (B-1, B-2), Hợp chất 3 (C-1, C-2), Hợp chất 4 (D-1, D-2), Hợp chất 5 (E-1, E-2) và Navitoclax (F-1, F-2) với thụ thể 6GL8.

Bên cạnh lực Van der Waals, các tương tác phân tử có liên kết mạnh như liên kết hydrogen và tương tác kỵ nước cũng được thấy trong mỗi phức hợp hình thành. Các liên kết hydrogen được tìm thấy trong mô phỏng tương tác của Hợp chất 1 với các gốc amino acid Ala149, Arg107, Leu137, Tyr108 (Hình 3.17A); Hợp chất 2 với các amino acid Arg107, Phe104, Tyr202 và Val133 (Hình 3.17B). Hợp chất 3 thể hiện các liên kết hydrogen với các amino acid Ala149, Arg146, Tyr108 (Hình 3.17C); Hợp chất 4 với các amino acid Ala149, Gln118 (Hình 3.17D); Hợp chất 5 với các amino acid Arg146, Asp140, Gln118, Glu136, Tyr108 (Hình 3.17E). Navitoclax biểu hiện liên kết hydrogen với các amino acid Gly203, Trp144 và Tyr202 (Hình 3.17F). Ngoài ra, Hợp chất 1 hình thành tương tác kỵ nước với các amino acid Arg146, Met115, Phe104, Val133 và Val148; Hợp chất 2 với các amino acid Ala149, Tyr108. Hợp chất 3 biểu hiện tương tác với các amino acid His120, Met115, Phe153, Val133; Hợp chất 4 với các amino acid His120, Leu119, Leu137, Met115, Val133; Hợp chất 5 với các amino acid Leu137,

Tyr108. Navitoclax thể hiện tương tác kỵ nước với các amino acid Ala149, Asp103, Leu137, Phe104 và Val133.

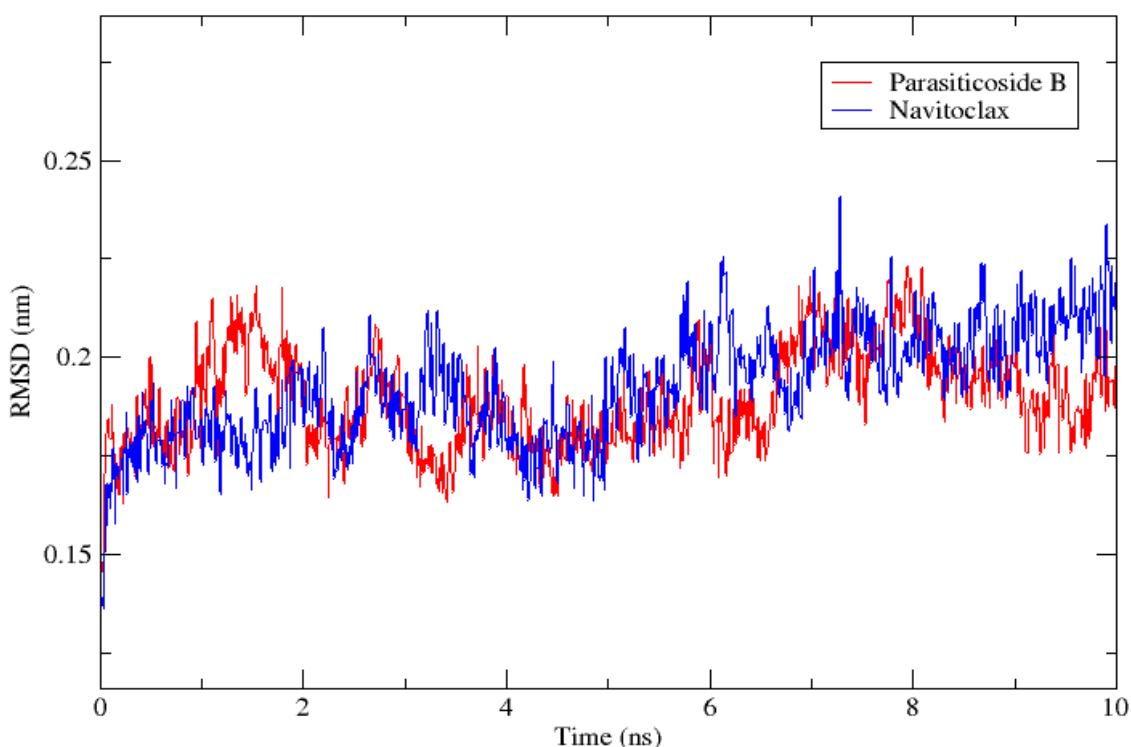
3.3.5.2. Nghiên cứu động học phân tử *in silico*

Hầu hết các tương tác được hình thành trong quá trình nghiên cứu động học giữa protein gây bệnh với các hợp chất nghiên cứu (Hợp chất 1, Hợp chất 2, Hợp chất 3, Hợp chất 4, Hợp chất 5 và Navitoclax) là các liên kết hydrogen; các liên kết này đóng vai trò quan trọng trong việc ổn định phức hợp hợp chất - protein. Hợp chất 2 thể hiện 4 liên kết hydrogen với các amino acid Arg107, Phe104, Tyr202 và Val133 tại các vị trí hoạt động của 6GL8 (Hình 3.17B). Các liên kết hydrogen tạo nên độ ổn định của protein-phối tử, mang lại cho Hợp chất 2 độ ổn định liên kết tốt là -8,87 kcal/mol. Trong trường hợp của Hợp chất 1, 4 liên kết hydrogen được hình thành với Ala149, Arg107, Leu137 và Tyr108 (Hình 3.17B); Ala149 là một amino acid tự nhiên không phân cực giống như Val133, có ý nghĩa trong việc làm cho Hợp chất 1 có độ ổn định liên kết tốt là -8,28 kcal/mol (Bảng 3.11). Các amino acid cũng được xác định có hình thành liên kết với Hợp chất 3 và Hợp chất 4, dẫn đến độ ổn định liên kết tốt với thụ thể protein lần lượt là -9,29 và -11,09 kcal/mol (Hình 3.17 C-2 và D-2). Mặt khác, Hợp chất 5 thể hiện 5 liên kết hydrogen với các amino acid Arg146, Asp140, Gln118, Glu136, Tyr108 và chỉ có tương tác pi-alkyl yếu với Leu137 dẫn đến độ ổn định liên kết tốt của 5 là -11,16 kcal/mol (Hình 3.17E). Navitoclax thể hiện 3 liên kết hydrogen khi tương tác với các amino acid Gly203, Trp144 và Tyr202, 6 tương tác kỵ nước với các amino acid Ala149, Asp103, Leu137, Phe104, Val133, Val148 và có độ ổn định liên kết tốt là -8,34 kcal/mol (Hình 3.17F).

Terefe và cs (2022) đã chứng minh rằng ái lực liên kết giữa thụ thể và phối tử càng cao thì ái lực liên kết càng âm [151]. Vì vậy, chất 3 tạo ra các liên kết ổn định hơn các saponin khác vì hợp chất này có tương tác phân tử mạnh hơn và có độ ổn định liên kết cao nhất. Tuy đã có một số nghiên cứu *in silico* về saponin, nhưng nghiên cứu của chúng tôi quan tâm đến các hợp chất được phân lập từ *H. verticillata* var. *verticillata*. Vì vậy, các nghiên cứu động học phân tử đã được thực hiện để hiểu rõ hơn về tương tác của các phối tử trong liên kết với protein mục tiêu. Các thí nghiệm tập trung vào đánh giá động học phân tử của Hợp chất 2 với 6GL8 trong đó nghiên cứu sự hình thành tương

tác của cấu trúc phức hợp hợp chất-protein, sử dụng mô phỏng động học phân tử của phức hợp giữa Navitoclax và 6GL8 để so sánh, đánh giá.

Tổng năng lượng và giá trị thế năng thu được trong mô phỏng động học phân tử đối với Hợp chất 2 được tính lần lượt là -280.721 và -225.612 kJ/mol. So với Navitoclax tham chiếu, tổng năng lượng và giá trị thế năng thu được trong mô phỏng động học phân tử đối với Navitoclax được tính lần lượt là -225.955 và -280.962 kJ/mol. Hệ thống đã được cân bằng ở nhiệt độ phòng (27°C), độ ổn định của khung protein đã được kiểm tra bằng cách đánh giá giá trị RMSD [158]. Có thể quan sát thấy rằng giá trị RMSD đối với phức hợp Hợp chất 2-6GL8 tăng dần từ 0,15 đến 0,22 nm trong các giai đoạn đầu của mô phỏng động học phân tử cho đến 2 ns hiệu chuẩn và sau đó ổn định, duy trì RMSD trung bình là 0,2 nm trong suốt phần còn lại của quỹ đạo động học phân tử. Trong trường hợp phức hợp Navitoclax-6GL8, giá trị RMSD khá giống với giá trị RMSD của phức hợp Hợp chất 2 với 6GL8 (Hình 3.18).

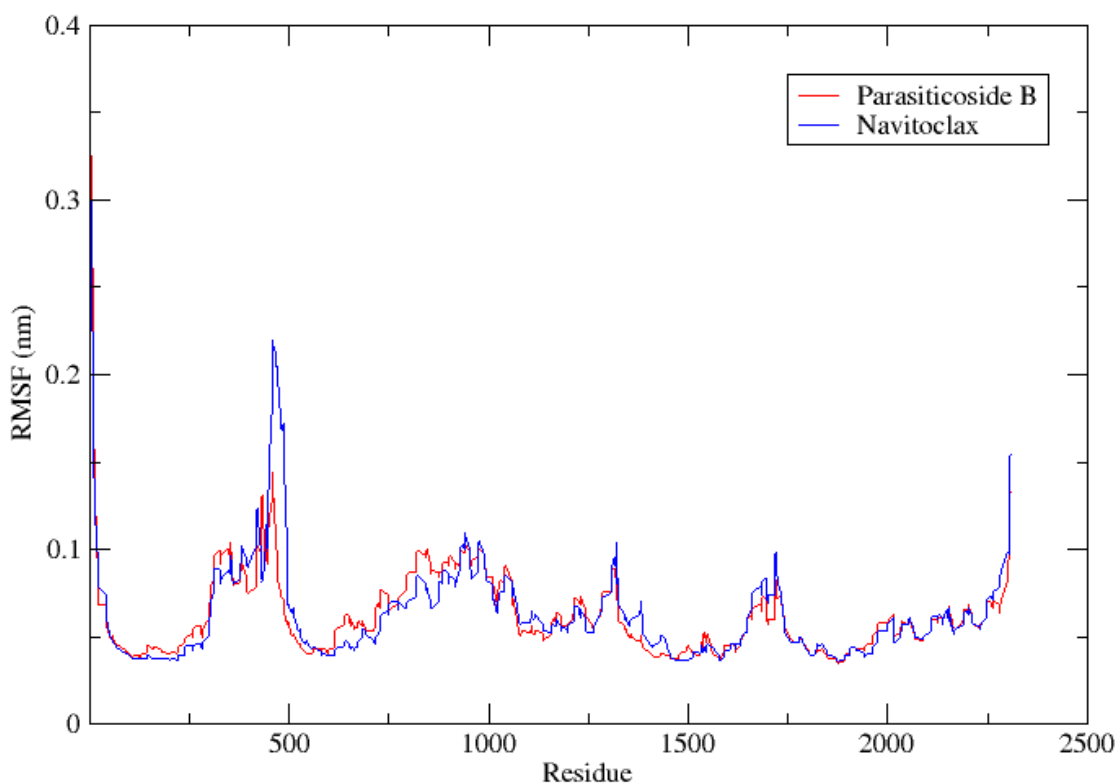


Hình 3.18. Biểu đồ so sánh độ ổn định cấu trúc (RMSD) của phức hợp Hợp chất 2 (Parasiticoside B) và Navitoclax trong 10 ns mô phỏng động học phân tử.

Trục tung là độ lệch bình phương trung bình căn bậc hai (RMSD) tính bằng nanomet (nm) dao động từ 0 đến 0,25 nm; trục hoành là thời gian tính bằng nano giây

(ns) dao động từ 0 đến 10 ns. Đường dữ liệu thể hiện Hợp chất 2 màu đỏ, Navitoclax hiển thị màu xanh. Navitoclax (đường màu xanh) cho thấy giá trị RMSD cao hơn (trong khoảng từ 0,15 nm đến 0,2 nm) và dao động lớn hơn trong suốt quá trình mô phỏng, cho thấy có những thay đổi về cấu hình đáng kể và kém ổn định trong môi trường mô phỏng. Hợp chất 2 (đường màu đỏ) duy trì giá trị RMSD thấp hơn đáng kể và ổn định hơn nhiều (khoảng 0,05-0,1 nm) trong toàn bộ quá trình mô phỏng. Điều này cho thấy Hợp chất 2 liên kết ổn định và chịu độ lệch cấu trúc tối thiểu, do đó có thể tạo thành một phức hợp mạnh hơn và ổn định hơn. Dựa trên phân tích giá trị RMSD, Hợp chất 2 thể hiện độ ổn định cấu trúc vượt trội so với Navitoclax, đây có thể là một chỉ số quan trọng về hiệu quả tiềm năng, là cơ sở để đề xuất thực nghiệm trong điều chế dược phẩm.

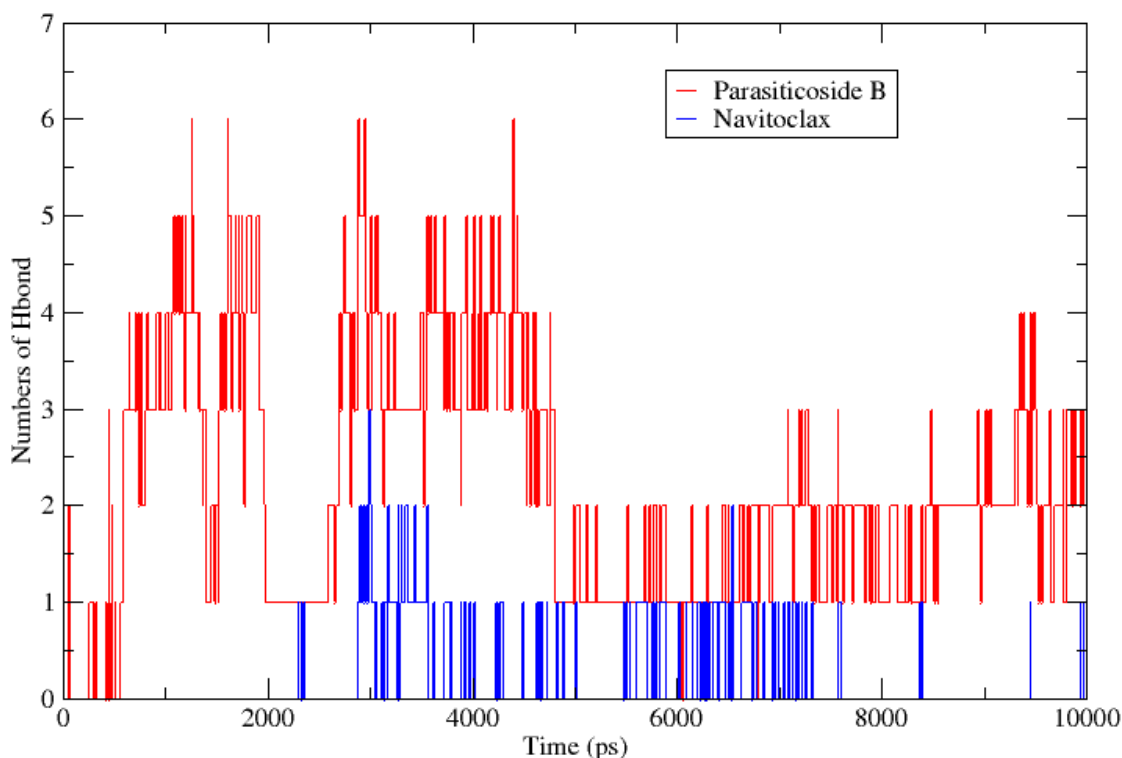
Giá trị RMSF được phân tích để xác định các biến động tiềm ẩn của các gốc protein trong khung protein biểu hiện là 0,15 nm trong phức hợp Hợp chất 2 với 6GL8 và 0,22 nm trong phức hợp Navitoclax-6GL8 (Hình 3.19).



Hình 3.19. Biểu đồ so sánh dao động cấu trúc (RMSF) của phức hợp Hợp chất 2 (Parasiticoside B) và Navitoclax trong mô phỏng động học phân tử.

Trục tung là biến động bình phương trung bình căn bậc hai tính bằng nano mét (nm), trục hoành là số đơn vị liên kết. Hợp chất 2 (đường màu đỏ), Navitoclax (đường màu xanh). Đồ thị động học phân tử của Hợp chất 2 và Navitoclax thể hiện sự tương đồng trong số liên kết được hình thành và giá trị biến động; đồ thị của Hợp chất 2 thể hiện sự ổn định trong cấu trúc phức hợp, dẫn đến độ linh hoạt thấp hơn (giá trị RMSF thấp hơn) trên hầu hết toàn bộ quá trình; Navitoclax tạo ra độ linh hoạt cao hơn trên hầu hết các protein, vì giá trị RMSF của nó luôn cao hơn giá trị của hệ thống liên kết Navitoclax.

Liên kết hydrogen được quan sát thấy trong suốt quá trình mô phỏng, cho thấy 1-6 liên kết hydrogen được hình thành liên tục trong phức hợp Hợp chất 2-6GL8 và 1-3 liên kết hydrogen được hình thành liên tục trong phức hợp Navitoclax-6GL8. Những kết quả này chỉ ra rằng parasiticoside B vẫn nằm trong phức hợp liên kết trong toàn bộ quá trình mô phỏng và mạnh hơn Navitoclax (Hình 3.20).



Hình 3.20. Biểu đồ so sánh sự thay đổi số liên kết hydrogen theo thời gian giữa Hợp chất 2 (Parasiticoside B) và Navitoclax trong mô phỏng động học phân tử.

Trục tung là số liên kết hydrogen được hình thành, trục hoành là thời gian tính bằng pico giây (ps). Tương tác của Hợp chất 2 với các liên kết của 6GL8 trước và sau khi mô phỏng động học phân tử (Hình 3.18, 3.19, 3.20). Sau 10000 ps (=10ns) mô phỏng động học phân tử, liên kết hydrogen và tương tác alkyl đã được thể hiện giữa Hợp chất 2 và 6GL8, bao gồm Glu77, Gly86, Tyr143 và Met56, Phe53, Tyr49. Ngoài ra, tương tác Van der Waals đã được thể hiện giữa Hợp chất 2 và các amino acid Ala41, Ala90, Arg80, Arg87, Asn84, Asp52, Glu85, Leu78, Phe45, Phe94, Val74, Val89 (Hình 3.18B). Navitoclax tham gia vào các liên kết hydrogen trong vị trí liên kết của 6GL8 với các amino acid Arg87 và Gly86. Nó cũng tham gia vào các tương tác kỵ nước với các amino acid Ala41, Phe45, Phe53, Phe94, Phe139, Tyr49 và Val89. Chất nghiên cứu được bao quanh bởi các tương tác Van der Waals với các amino acid Ala90, Arg48, Asn84, Asp44, Asp52, Glu77, Leu78, Leu142, Met56, Thr37, Trp85, Tyr143 (Hình 3.17A và 3.17B).

Tất cả các hợp chất (saponin steroid) đều cho thấy có sự hình thành liên kết hydrogen, tương tác Van der Waals và tương tác kỵ nước trong quá trình tương tác phân tử. Mô hình tương tác gồm 05 hợp chất (02 hợp chất phân lập trong nghiên cứu này là 1, 2 và 03 hợp chất phân lập được trong các nghiên cứu trước đây là 3, 4, 5) chất đối chứng là navitoclax (chất đã được kiểm chứng chức năng ức chế Bcl-2 gây ung thư ở người thông qua đáp ứng thụ thể 6GL8 của protein Bcl-2). Kết quả cho thấy độ ổn định liên kết lần lượt là -8,28, -8,87, -9,29, -11,09, -11,16 và -8,34 kcal/mol. Trên cơ sở nghiên cứu mô hình tương tác phân tử giúp hiểu rõ hơn về khả năng hình thành liên kết giữa các hợp chất cần nghiên cứu với protein mục tiêu. Mô phỏng động học phân tử cho thấy liên kết của Hợp chất 2 với thụ thể 6GL8 là ổn định trong suốt quá trình mô phỏng.

Dựa trên các bằng chứng trên, có thể kết luận rằng kết quả mô phỏng tương tác phân tử cho thấy chế độ ức chế liên tục và tương tác liên kết nhất quán trong tất cả các mô phỏng [110]. Nghiên cứu này chỉ ra rằng sự liên kết của Hợp chất 2 với thụ thể 6GL8 là ổn định trong suốt quá trình mô phỏng. Những kết quả này hỗ trợ độ tin cậy và xác thực việc sử dụng phương pháp mô phỏng động học phân tử trong nghiên cứu này. Hợp chất 2 có thể là ứng cử viên tốt chống lại protein ung thư Bcl-2.

Mô hình mô phỏng phân tử đã thể hiện sự tương tác giữa các hợp chất với protein gây bệnh ung thư ở người, tương tác trong cấu trúc ổn định trong quá trình mô phỏng là cơ sở quan trọng trong đề xuất các bước thực nghiệm để đánh giá hoạt tính của chất trong điều kiện phòng thí nghiệm. Những phát hiện này chỉ ra rằng sự liên kết của Hợp chất 2 với thụ thể 6GL8 là ổn định trong suốt quá trình mô phỏng, đây chính là cơ sở để đề xuất Hợp chất 2 là ứng viên tiềm năng cho nghiên cứu thực nghiệm hoạt tính chống lại protein Bcl-2 trong ung thư trực tràng.

KẾT LUẬN VÀ ĐỀ NGHỊ

1. Kết luận

1.1. Hệ gene lục lập hoàn chỉnh của Lan tai cáo (*H. verticillata* var. *verticillata*) đã được giải trình tự với kích thước 176.429 bp, có cấu trúc 4 vùng điển hình của phân tử DNA dạng vòng, bao gồm vùng sao chép đơn lớn-LSC (91.020 bp), vùng sao chép đơn nhỏ-SSC (2.297 bp) và hai vùng lặp lại đảo ngược IR, mỗi vùng IR có kích thước 41.556 bp, trong đó hàm lượng GC chiếm 37,1%. Hệ gene lục lập của Lan tai cáo có 144 gene, trong đó 98 gene mã hóa protein, 38 gene tRNA và 8 gene rRNA.

1.2. Phân tích sự phát sinh chủng loại cho thấy Lan tai cáo (*H. verticillata* var. *verticillata*) có mối quan hệ di truyền gần nhất với loài *H. carnosa*, chi *Hoya* và có mối quan hệ di truyền gần với các loài của chi khác trong cùng tông Marsdenieae. Vùng gene *ndhF* và các vùng đệm *trnK-rps16*, *rps16-trnQ*, *psbI-atpA*, *ndhC-trnV*, *rbcL-accD* được đề xuất là ứng viên mã vạch DNA tiềm năng có thể sử dụng hỗ trợ nhận diện loài thuộc chi *Hoya*.

1.3. Các phân đoạn chứa saponin tách từ cao chiết tổng số của lá Lan tai cáo *H. verticillata* var. *verticillata*) có hoạt tính kháng oxy hoá, gây độc tế bào ung thư và ức chế α -glucosidase. Hai saponin mới được phân lập từ phân đoạn B2 và B3 là 3 β ,20-dihydroxy-pregn-5-ene-3-O- β -D-fucopyranoside 20-O- β -D-glucopyranoside và 3 β ,20-dihydroxy-pregn-5-ene-3-O- α -L-rhamnopyranosyl-(1 $\rightarrow$ 2)- β -D-fucopyranoside-20-O- β -D-glucopyranoside. Phân tích mô hình tương tác và động học phân tử *in silico* của các saponin phân lập từ Lan tai cáo đã xác định chất Hợp chất 2 có thể là ứng cử viên tiềm năng chống protein Bcl-2 trong ung thư.

2. Đề nghị

2.1. Tiếp tục phân tích tiến hóa phân tử dựa trên trình tự DNA lục lập của các loài *Hoya* từ dữ liệu của GenBank để củng cố hoặc bổ sung cho hệ thống các loài thuộc họ Apocynaceae (Asclepiadaceae).

2.2. Tiếp tục thực nghiệm để chứng minh vùng gene *ndhF* và các vùng đệm *trnK-rps16*, *rps16-trnQ*, *psbI-atpA*, *ndhC-trnV*, *rbcL-accD* làm mã vạch DNA lục lạp hỗ trợ định danh chi *Hoya* và phân biệt loài thuộc chi này.

2.3. Tiếp tục nghiên cứu xây dựng mô hình cấu trúc saponin tối ưu có hoạt tính kháng các dòng tế bào ung dựa trên kết quả phân tích mô hình tương tác và động học phân tử *in silico*.

CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ LIÊN QUAN ĐẾN LUẬN ÁN

Các bài báo

1. Hoang, C. V., Tu, T. Q., Nguyen, L. T. N., Nguyen, H. D., Nguyen, Q. H., & Chu, M. H. (2023). Two New C21 Steroidal Glycosides from the Leaves of *Hoya parasitica*. *Records of Natural Products*, 17(6), pp. 1046-1051. <http://doi.org/10.25135/rnp.419.2307.2831> (SCI-E).
2. Cuong Viet Hoang, Tan Quang Tu, Thu Thi Mai Lo, Mau Hoang Chu (2024), “Dataset on intergenic spacer regions of chloroplast genome and potential DNA barcode of *H. verticillata* var. *verticillata* in Vietnam”, *Data in Brief*, 54, 110471. <https://doi.org/10.1016/j.dib.2024.110471> (Scopus).
3. Hoàng Việt Cường, Phạm Thị Thu Hiền, Từ Quang Tân, Chu Hoàng Mậu (2024), “Phân tích sự phát sinh chủng loại dựa trên trình tự *psbz-rps14* và *rbcl-accd* nhằm hỗ trợ định danh loài lan tai cáo (*Hoya verticillata* var. *verticillata*)”, *Tạp chí Khoa học và Công nghệ Đại học Thái Nguyên*, 230(01), tr. 177-184. <https://doi.org/10.34238/tnu-jst.10996> (ACI).
4. Cuong Viet Hoang, Tan Quang Tu, Hung Duc Nguyen, Mau Hoang Chu (2025), “*In silico* studies of saponins from *Hoya verticillata* var. *verticillata* with important apoptosis potency”, *Letters in Organic Chemistry*, 22 (11), pp. 846 - 854. <https://doi.org/10.2174/0115701786363779250528013022> (SCI-E).

Trình tự hệ gene lục lập đăng kí trên GenBank

1. Jo,S., Mau,C.H., Tan,T.Q., Cuong, H.V., Lee,C., Quan,N.H., Thuong,S.D. and Lan,N.T.N. (2023), “*Hoya verticillata* chloroplast, complete genome”, GenBank: OR475244.1; <https://www.ncbi.nlm.nih.gov/nuccore/OR475244.1>

TÀI LIỆU THAM KHẢO

Tài liệu tiếng Việt

- 1 Nguyễn Kim Phi Phụng (2007), *Phương pháp cô lập hợp chất hữu cơ*, Nxb Đại học Quốc gia TP. Hồ Chí Minh.
- 2 Viện hàn lâm Khoa học và Công nghệ Việt Nam (2017), *Thực vật chí Việt Nam*, tập 15. Họ Thiên lý – Asclepiadaceae R. Br. Trần Thế Bách. Nxb Khoa học Tự nhiên và Công nghệ, Hà Nội. tr. 7–33.

Tài liệu tiếng Anh

- 3 Abe, F., Fujishima, H., Iwase, Y., Yamauchi, T., Kinjo, K., & Yaga, S. (1999). Pregnanes and pregnane glycosides from *Hoya carnosa*. *Chemical and pharmaceutical bulletin*, 47(8), pp. 1128-1133.
- 4 Abramović, H., Grobin, B., Poklar Ulrih, N., & Cigić, B. (2018). Relevance and Standardization of *In Vitro* Antioxidant Assays: ABTS, DPPH, and Folin–Ciocalteu. *Journal of Chemistry*, 2018, pp. 1–9
- 5 Addisu, S., & Assefa, A. (2016). Role of plant containing saponin on livestock production; a review. *Advances in Biological Research*, 10(5), pp. 309-314.
- 6 Afrose, S., Hossain, M. S., Maki, T., & Tsujii, H. (2009). Karaya root saponin exerts a hypocholesterolemic response in rats fed a high-cholesterol diet. *Nutrition research*, 29(5), pp. 350-354.
- 7 Agarwal, S., & Mehrotra, R. J. J. C. (2016). An overview of molecular docking. *Jai Swaminarayan Multichem*, 4(2), pp. 1024-1028.
- 8 Al-Snafi, A. E. (2018). Pharmacological and therapeutic activities of *Hedera helix*- A review. *Journal of Pharmacy*, 8(5), pp. 41-53.
- 9 Amiryousefi, A., Hyvönen, J., & Poczar, P. (2018). IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics*, 34(17), pp. 3030-3031.
- 10 Arabski, M., Węgierek-Ciuk, A., Czerwonka, G., Lankoff, A., & Kaca, W. (2012). Effects of saponins against clinical *E. coli* strains and eukaryotic cell line. *BioMed Research International*, 2012(1), pp. 286216.

- 11 Aumsuwan, P., Khan, S. I., Khan, I. A., Ali, Z., Avula, B., Walker, L. A., ... & Dasmahapatra, A. K. (2016). The anticancer potential of steroidal saponin, dioscin, isolated from wild yam (*Dioscorea villosa*) root extract in invasive human breast cancer cell line MDA-MB-231 *in vitro*. *Archives of Biochemistry and Biophysics*, 591, pp. 98-110.
- 12 Baas, W. J. (1983). Dihydronyctanthic acid methyl ester and other 3, 4-seco-pentacyclic triterpenoids from *Hoya lacunosa*. *Phytochemistry*, 22(12), pp. 2809-2812.
- 13 Baas, W. J., & Niemann, G. J. (1979). Investigations on *Hoya* species. I. *Planta Medica*, 35(04), pp. 348-353.
- 14 Baas, W. J., & Van Berkel, I. E. (1991). 3, 4-Seco-triterpenoid acids and other constituents of the leaf wax of *Hoya naumanii*. *Phytochemistry*, 30(5), pp. 1625-1628.
- 15 Baas, W. J., Van Berkel, I. E., Versluis, C., Heerma, W., & Kreyenbroek, M. N. (1992). Ring-A fissioned 3, 4-seco-3-nor-triterpene-2-aldehydes and related pentacyclic triterpenoids from the leaf wax of *Hoya australis*. *Phytochemistry*, 31(6), pp. 2073-2078.
- 16 Baas, W. J., Warnaar, F., & Niemann, G. J. (1981). Investigations on *Hoya* species. VI. Latex composition and leaf phenolics and their taxonomic significance. *Acta Botanica Neerlandica*, 30(4), pp. 257-263.
- 17 Bach, T. T., Kim, J. H., Kim, D. K., Lee, J., Ha, B. T., Juhonewe, N. S., & Rodda, M. (2011). *Hoya ignorata* (Apocynaceae, Asclepiadoideae): an overlooked species widely distributed across southeast Asia. *Novon: A Journal for Botanical Nomenclature*, 21(4), pp. 508-514.
- 18 Basir, S., Akbar, M. A., Talip, N., Baharum, S. N., & Bunawan, H. (2022). An integrative volatile terpenoid profiling and transcriptomics analysis in *Hoya cagayanensis*, *Hoya lacunosa* and *Hoya coriacea* (Apocynaceae, Marsdenieae). *Horticulturae*, 8(3), pp. 224.
- 19 Battistuzzi, F. U., & Kumar, S. Molecular Phylogeny Reconstruction. *Encyclopedia of Life Sciences*, 1, pp. 558-564.

- 20 Beier, S., Thiel, T., Münch, T., Scholz, U., & Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics*, 33(16), pp. 2583-2585.
- 21 Berlowska, J., Dudkiewicz, M., Kregiel, D., Czyzowska, A., & Witonska, I. (2015). Cell lysis induced by membrane-damaging detergent saponins from *Quillaja saponaria*. *Enzyme and microbial technology*, 75, pp. 44-48.
- 22 Bitencourt-Ferreira, G., Veit-Acosta, M., & de Azevedo Jr, W. F. (2019). Van der Waals potential in protein complexes. In *Docking Screens for Drug Discovery*, pp. 79-91.
- 23 Borlagdan, M., Aurigue, F. B., Shen, C. C., & Ragasa, C. Y. (2018). Chemical constituents of *Hoya madulidii*. *Chemistry of Natural Compounds*, 54, pp. 198-199.
- 24 Brown, J. S., Amend, S. R., Austin, R. H., Gatenby, R. A., Hammarlund, E. U., & Pienta, K. J. (2023). Updating the definition of cancer. *Molecular Cancer Research*, 21(11), pp. 1142-1147.
- 25 Chaachouay, N., & Zidane, L. (2024). Plant-derived natural products: a source for drug discovery and development. *Drugs and Drug Candidates*, 3(1), pp. 184-207.
- 26 Chen, H., Li, T., Chen, X., Qu, T., Zheng, X., Luo, J., ... & Fu, Z. (2024). Insights into comparative genomics, structural features, and phylogenetic relationship of species from Eurasian *Aster* and its related genera (Asteraceae: Astereae) based on complete chloroplast genome. *Frontiers in Plant Science*, 15, pp. 1367132.
- 27 Chen, K., Wang, N., Zhang, X., Wang, M., Liu, Y., & Shi, Y. (2023). Potentials of saponins-based adjuvants for nasal vaccines. *Frontiers in Immunology*, 14, pp. 1153042.
- 28 Chopade, A. R., Sayyad, F. J., & Pore, Y. V. (2014). Molecular docking studies of phytocompounds from the *phyllanthus* species as potential chronic pain modulators. *Scientia pharmaceutica*, 83(2), pp. 243.
- 29 Clegg, M. T., & Durbin, M. L. (1990). Molecular approaches to the study of plant biosystematics. *Australian Systematic Botany*, 3(1), pp. 1-8.
- 30 Cui, M., Song, F., Zhou, Y., Liu, Z., & Liu, S. (2000). Rapid identification of saponins in plant extracts by electrospray ionization multi-stage tandem mass spectrometry and liquid chromatography/tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 14(14), pp. 1280-1286.

- 31 Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature methods*, 9(8), pp. 772-772.
- 32 de Paula Barbosa, A. (2014). Saponins as immunoadjuvant agent: a review. *Afr J Pharm Pharmacol*, 8, pp. 1049-1057.
- 33 Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic acids research*, 45(4), e18-e18.
- 34 Dighe, A. S., & Tajamulhaq, A. E. (2024). An Overview of Molecular Docking. *Asian Journal of Pharmaceutical Research*, 14(3), pp. 336-340.
- 35 Ebajo Jr, V. D., Shen, C. C., & Ragasa, C. Y. (2014). Triterpenes and sterol from *Hoya mindorensis*. *Der Pharma Chemica*, 6(4), pp. 321-325.
- 36 Ebajo Jr, V. D., Shen, C. C., & Ragasa, C. Y. (2015). Terpenoids and sterols from *Hoya multiflora* Blume. *Journal of applied pharmaceutical science*, 5(4), pp. 033-039.
- 37 Ebajo Jr, V. D., Urban, S., & Ragasa, C. Y. (2015). Chemical constituents of *Hoya buotii* Kloppenb. *Journal of Applied Pharmaceutical Science*, 5(11), pp. 069-072.
- 38 El Aziz, M. M. A., Ashour, A. S., & Melad, A. S. G. (2019). A review on saponins from medicinal plants: chemistry, isolation, and determination. *J. Nanomed. Res*, 8(1), pp. 282-288.
- 39 Elekofehinti, O. O., Iwaloye, O., Olawale, F., & Ariyo, E. O. (2021). Saponins in cancer treatment: Current progress and future prospects. *Pathophysiology*, 28(2), pp. 250-272.
- 40 Elhag, R., Almalki, A., Mazzio, E., Darling-Reed, S., & Soliman, K. F. (2022). Immune-stimulating properties of Wild Yam (*Dioscorea villosa*) root extract in triple-negative breast cancer cells. *The FASEB Journal*, pp. 36.
- 41 Elliott, M., & Chithan, K. (2017). The impact of plant flavonoids on mammalian biology: implications for immunity, inflammation and cancer. In *The flavonoids advances in research since 1986*, pp. 619-652.
- 42 Ellis, R. J. (1977). The synthesis of chloroplast proteins. In *Nucleic acids and protein synthesis in plants*, pp. 195-212.

- 43 Endress, M. E., Meve, U., Middleton, D. J., & Liede-Schumann, S. (2018). Apocynaceae: Periplocaceae Schltr.(1905). *Flowering Plants. Eudicots: Apiales, Gentianales (except Rubiaceae)*, pp. 207-411.
- 44 Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *evolution*, 39(4), pp. 783-791.
- 45 Filipe, H. A., & Loura, L. M. (2022). Molecular dynamics simulations: advances and applications. *Molecules*, 27(7), pp. 2105.
- 46 Flores-Perez, U., Sauret-Gueto, S., Gas, E., Jarvis, P., & Rodríguez-Concepción, M. (2008). A mutant impaired in the production of plastome-encoded proteins uncovers a mechanism for the homeostasis of isoprenoid biosynthetic enzymes in Arabidopsis plastids. *The Plant Cell*, 20(5), pp. 1303-1315.
- 47 Gao, C., Li, T., Zhao, X., Wu, C., Zhang, Q., Zhao, X., ... & Li, Z. (2023). Comparative analysis of the chloroplast genomes of *Rosa* species and RNA editing analysis. *BMC Plant Biology*, 23(1), pp. 318.
- 48 Gilabert-Oriol, R., Mergel, K., Thakur, M., von Mallinckrodt, B., Melzig, M. F., Fuchs, H., & Weng, A. (2013). Real-time analysis of membrane permeabilizing effects of oleanane saponins. *Bioorganic & medicinal chemistry*, 21(8), pp. 2387-2395.
- 49 Gillham, N. W. (1978). Organelle heredity, *Raven Press*, pp. 347-380.
- 50 Gou, W., Jia, S. B., Price, M., Guo, X. L., Zhou, S. D., & He, X. J. (2020). Complete plastid genome sequencing of eight species from Hansenia, Haplosphaera and Sinodielsia (Apiaceae): Comparative analyses and phylogenetic implications. *Plants*, 9(11), pp. 1523.
- 51 Graf, B. L., Raskin, I., Cefalu, W. T., & Ribnicky, D. M. (2010). Plant-derived therapeutics for the treatment of metabolic syndrome. *Current opinion in investigational drugs (London, England: 2000)*, 11(10), pp. 1107.
- 52 Greiner, S., Lehwark, P., & Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3. 1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic acids research*, 47(W1), pp. 59-64.

- 53 Grosso, C., Santos, M., & Barroso, M. F. (2023). From plants to psycho-neurology: unravelling the therapeutic benefits of bioactive compounds in brain disorders. *Antioxidants*, 12(8), pp. 1603.
- 54 Gu, X., Li, L., Li, S., Shi, W., Zhong, X., Su, Y., & Wang, T. (2023). Adaptive evolution and co-evolution of chloroplast genomes in Pteridaceae species occupying different habitats: overlapping residues are always highly mutated. *BMC Plant Biology*, 23(1), pp. 511.
- 55 Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *electrophoresis*, 18(15), pp. 2714-2723.
- 56 Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics*, 4, pp. 1-17.
- 57 He, S., Wang, Y., Volis, S., Li, D., & Yi, T. (2012). Genetic diversity and population structure: implications for conservation of wild soybean (*Glycine soja* Sieb. et Zucc) based on nuclear and chloroplast microsatellite variation. *International Journal of Molecular Sciences*, 13(10), pp. 12608-12628.
- 58 Hostettmann, K.A. and Marston, A. (1995) Saponins. *Cambridge University Press, Cambridge*, pp. 548.
- 59 Hu, C. C., Lin, J. T., Liu, S. C., & Yang, D. J. (2007). A spirostanol glycoside from wild yam (*Dioscorea villosa*) extract and its cytostatic activity on three cancer cells. *Journal of Food and Drug Analysis*, 15(3), pp. 3.
- 60 Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. *British journal of pharmacology*, 162(6), pp. 1239-1249.
- 61 Ibrahim, N. M., & Kadhim, E. J. (2015). Phytochemical investigation and antioxidant activity of Iraqi *Tribulus terrestris*. *Iraqi Journal of Pharmaceutical Sciences* (P-ISSN 1683-3597 E-ISSN 2521-3512), 24(1), pp. 68-73.
- 62 Jabborova, D., Davranov, K., & Egamberdieva, D. (2019). Antibacterial, antifungal, and antiviral properties of medical plants. *Medically Important Plant Biomes: Source of Secondary Metabolites*, pp. 51-65.

- 63 Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., DePamphilis, C. W., Yi, T. S., & Li, D. Z. (2020). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome biology*, 21, pp. 1-31.
- 64 Jolly, A., Hour, Y., & Lee, Y. C. (2024). An outlook on the versatility of plant saponins: a review. *Fitoterapia*, pp. 105858.
- 65 Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4), pp. 772-780.
- 66 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), pp. 1647-1649.
- 67 Khakimov, B., Tseng, L. H., Godejohann, M., Bak, S., & Engelsens, S. B. (2016). Screening for triterpenoid saponins in plants using hyphenated analytical platforms. *Molecules*, 21(12), pp. 1614.
- 68 Khan, A. W., Farooq, M., Haseeb, M., & Choi, S. (2022). Role of plant-derived active constituents in cancer treatment and their mechanisms of action. *Cells*, 11(8), p. 1326.
- 69 Khoddami, A., Wilkes, M. A., & Roberts, T. H. (2013). Techniques for analysis of plant phenolic compounds. *Molecules*, 18(2), pp. 2328-2375.
- 70 Kim, K. H., Choi, I., Lee, Y. W., Cho, C. K., Yoo, H. S., Lee, S. B., ... & Jang, J. H. (2014). Target genes involved in antiproliferative effect of modified ginseng extracts in lung cancer A549 cells. *Acta Biochim Biophys Sin*, 46(6), pp. 441-449.
- 71 Kim, S., Kim, N., Jeong, J., Lee, S., Kim, W., Ko, S. G., & Kim, B. (2021). Anti-cancer effect of Panax ginseng and its metabolites: from traditional medicine to modern drug discovery. *Processes*, 9(8), pp. 1344.
- 72 Kimura, M. (1991). The neutral theory of molecular evolution: a review of recent evidence. *The Japanese Journal of Genetics*, 66(4), pp. 367-386.
- 73 Kirk, J.T.O.; Tilney-Bassett, R.A.E. (1978). The Plastids: Their chemistry, structure, growth and inheritance (2nd ed.). *Amsterdam: Elsevier/North Holland*.

- 74 Kolekar, S. M., Jain, B. U., & Kondawarkar, M. S. (2019). A review on steroids and terpenoids (stereochemistry, structural elucidation, isolation of steroids and terpenoids). *Research Journal of Pharmaceutical Dosage Forms and Technology*, 11(2), pp. 126-130.
- 75 Kruakaew, S., Singha, S., Seeka, C., Mongkolvisut, W., & Sutthivaiyakit, S. (2018). Triterpenoid cinnamate, acetate and palmitate from *Hoya kerrii* leaves. *Phytochemistry Letters*, 27, pp. 119-122.
- 76 Kung, S. D., Lee, C. I., Wood, D. D., & Moscarello, M. A. (1977). Evolutionary conservation of chloroplast genes coding for the large subunits of fraction 1 protein. *Plant Physiology*, 60(1), pp. 89-94.
- 77 Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., & Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic acids research*, 29(22), pp. 4633-4642.
- 78 Kuwada, K., Kawase, S., Nakata, K., Shinya, N., Narukawa, Y., Fuchino, H., ... & Kiuchi, F. (2020). LC–MS analysis of saponins of *Achyranthes* root in the Japanese market. *Journal of natural medicines*, 74, pp. 135-141.
- 79 Lee, S. R., Kim, K., Lee, B. Y., & Lim, C. E. (2019). Complete chloroplast genomes of all six *Hosta* species occurring in Korea: molecular structures, comparative, and phylogenetic analyses. *BMC genomics*, 20, pp. 1-13.
- 80 Leese, F., Mayer, C., & Held, C. (2008). Isolation of microsatellites from unknown genomes using known genomes as enrichment templates. *Limnology and Oceanography: Methods*, 6(9), pp. 412-426.
- 81 Li, H. T., Luo, Y., Gan, L., Ma, P. F., Gao, L. M., Yang, J. B., ... & Li, D. Z. (2021). Plastid phylogenomic insights into relationships of all flowering plant families. *BMC biology*, 19, pp. 1-13.
- 82 Li, J., Wang, S., Yu, J., Wang, L., & Zhou, S. (2013). A modified CTAB protocol for plant DNA extraction. *Chinese Bulletin of Botany*, 48(1), pp. 72.

- 83 Li, X. H., Qi, Y., Cai, R. L., Liu, B., Song, Y., & Xie, C. (2010). Studies on the anti-inflammatory mechanism of total saponins of radix *Glycyrrhiza in vitro*. *Chinese Journal Experimental Traditional Medical Formulae*, 5, pp. 039.
- 84 Lorent, J. H., Quetin-Leclercq, J., & Mingeot-Leclercq, M. P. (2014). The amphiphilic nature of saponins and their effects on artificial and biological membranes and potential consequences for red blood and cancer cells. *Organic & biomolecular chemistry*, 12(44), pp. 8803-8822.
- 85 Manganyi, M. C., & Ateba, C. N. (2020). Untapped potentials of endophytic fungi: A review of novel bioactive compounds with biological applications. *Microorganisms*, 8(12), pp. 1934.
- 86 Martel, J., Ojcius, D. M., Ko, Y. F., Chang, C. J., & Young, J. D. (2019). Antiaging effects of bioactive molecules isolated from plants and fungi. *Medicinal Research Reviews*, 39(5), pp. 1515-1552.
- 87 Mayer, C., Leese, F., & Tollrian, R. (2010). Genome-wide analysis of tandem repeats in *Daphnia pulex*-a comparative approach. *BMC genomics*, 11, pp. 1-28.
- 88 Mazzio, E., Almalki, A., Darling-Reed, S. F., & Soliman, K. F. (2021). Effects of wild yam root (*Dioscorea villosa*) extract on the gene expression profile of triple-negative breast cancer cells. *Cancer Genomics & Proteomics*, 18(6), pp. 735-755.
- 89 Mieres-Castro, D., & Mora-Poblete, F. (2023). Saponins: research progress and their potential role in the post-COVID-19 pandemic era. *Pharmaceutics*, 15(2), pp. 348.
- 90 Moghimipour, E., Ameri, A., Handali, S., Ramezani, Z., Azemi, M. E., & Sadaghi-Nejad, B. (2015). In-vitro evaluation of antibacterial activity of *Glycyrrhiza glabra* and *Acanthopyllum squarrusom* total saponins. *CABI Databases*, pp. 764-768.
- 91 Moses, T., Papadopoulou, K. K., & Osbourn, A. (2014). Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. *Critical reviews in biochemistry and molecular biology*, 49(6), pp. 439-462.
- 92 Mukherjee, P. K., Maity, N., Nema, N. K., & Sarkar, B. K. (2011). Bioactive compounds from natural resources against skin aging. *Phytomedicine*, 19(1), pp. 64-73.

- 93 Negi, A. S., Kumar, J. K., Luqman, S., Shanker, K., Gupta, M. M., & Khanuja, S. P. S. (2008). Recent advances in plant hepatoprotectives: a chemical and biological profile of some important leads. *Medicinal research reviews*, 28(5), pp. 746-772.
- 94 Negi, J. S., Singh, P., Pant, G. J. N., & Rawat, M. S. M. (2011). High-performance liquid chromatography analysis of plant saponins: An update 2005-2010. *Pharmacognosy reviews*, 5(10), pp. 155.
- 95 Ng, V. A. S., Malabed, R. S., Aurigue, F. B., & Ragasa, C. Y. (2019). Triterpenes and sterols from leaves of *Hoya meliflua* Merr. *Pharmacognosy Journal*, 11(1), pp. 48-52.
- 96 Nguyen, D. H., Mitaine-Offer, A. C., Maroso, S., Papini, A. M., Paululat, T., Bellaye, P. S., ... & Lacaille-Dubois, M. A. (2019). Cytotoxic glycosides from the roots of *Weigela* x “Bristol Ruby”. *Fitoterapia*, 137, pp. 104242.
- 97 Nguyen, D. H., Mitaine-Offer, A. C., Miyamoto, T., Tanaka, C., Bellaye, P. S., Collin, B., ... & Lacaille-Dubois, M. A. (2020). Phytochemical analysis of two *Weigela florida* cultivars, “Pink Poppet” and “Jean’s Gold”. *Phytochemistry Letters*, 37, pp. 85-89.
- 98 Nguyen, D. H., Mitaine-Offer, A. C., Miyamoto, T., Tanaka, C., Bellaye, P. S., Collin, B., ... & Lacaille-Dubois, M. A. (2021). Steroidal glycosides from the Vietnamese cultivar *Cordyline fruticosa* “Fairchild red”. *Phytochemistry*, 192, pp. 112966.
- 99 Nguyen, H. D. (2023). Two new triterpenoid saponins from the underground parts of *Weigela* x “Bristol Ruby”. *Journal of Asian Natural Products Research*, 25(3), pp. 218-224.
- 100 Nguyen, H. Q., Nguyen, T. N. L., Doan, T. N., Nguyen, T. T. N., Phạm, M. H., Le, T. L., ... & Chu, H. M. (2021). Complete chloroplast genome of novel *Adinandra megaphylla* Hu species: molecular structure, comparative and phylogenetic analysis. *Scientific Reports*, 11(1), pp. 11731.
- 101 Nguyen, N. H., Ha, T. K. Q., Yang, J. L., Pham, H. T. T., & Oh, W. K. (2021). Triterpenoids from the genus *Gynostemma*: Chemistry and pharmacological activities. *Journal of ethnopharmacology*, 268, pp. 113574.

- 102 Nguyen, N. T. T., Pho, H. T. T., Nguyen, Q. H., Doan, N. T., Nguyen, L. T. N., Pham, H. M., ... & Chu, M. H. (2023). Characteristics of the chloroplast genome of *Adinandra bockiana* and comparative analysis with species of Pentaphragmataceae family. *Plant Molecular Biology Reporter*, 41(4), pp. 611-621.
- 103 Niemann, G. J., Baas, W. J., Besson, E., & Chopin, J. (1979). Investigations on *Hoya* Species. III. Leaf Phenolics and Latex Lipids of *Hoya lacunosa* Bl. *Zeitschrift für Naturforschung C*, 34(12), pp. 1125-1128.
- 104 Odago, W. O., Waswa, E. N., Nanjala, C., Mutinda, E. S., Wanga, V. O., Mkala, E. M., ... & Wang, Q. F. (2022). Analysis of the complete plastomes of 31 species of *Hoya* group: insights into their comparative genomics and phylogenetic relationships. *Frontiers in Plant Science*, 12, pp. 814833.
- 105 Ohta, T. (1992). The nearly neutral theory of molecular evolution. *Annual review of ecology and systematics*, pp. 263-286.
- 106 Orton, L. M., Barberá, P., Nissenbaum, M. P., Peterson, P. M., Quintanar, A., Soreng, R. J., & Duvall, M. R. (2021). A 313 plastome phylogenomic analysis of Pooideae: Exploring relationships among the largest subfamily of grasses. *Molecular phylogenetics and evolution*, 159, pp. 107110.
- 107 Pal, A., & Pal, A. (2022). Molecular Evolutionary Study: Phylogenetic Tree. *Protocols in Advanced Genomics and Allied Techniques*, pp. 159-180.
- 108 Pandey, S. C., Singh, S. S., Ghosh, A. C., Desh Deepak, D. D., & Anakshi Khare, A. K. (2004). Chemistry and pharmacological aspects of *Hoya* species-a review. *CABI Databases*, pp. 775-783.
- 109 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13), pp. 1605-1612.
- 110 Phong, N. V., Heo, M. S., Vinh, L. B., Kim, Y. H., & Yang, S. Y. (2024). Investigation of the inhibitory activity of triterpenoids isolated from *Actinidia polygama* stems against β -glucuronidase via enzyme kinetics, molecular docking, and molecular dynamics analyses. *Journal of Molecular Structure*, 1317, pp. 139135.

- 111 Potterat, O., & Hamburger, M. (2008). Drug discovery and development with plant-derived compounds. *Natural Compounds as Drugs Volume I*, pp. 45-118.
- 112 Qu, Z. Y., Wang, H. C., Jin, Y. P., Li, Y. L., & Wang, Y. P. (2022). Isolation, identification, and quantification of triterpene saponins in the fresh fruits of *Panax notoginseng*. *Natural Product Research*, 36(20), pp. 5319-5329.
- 113 Ragasa, C. Y., Borlagdan, M. S., Aurigue, F. B., Brkljača, R., & Urban, S. (2017). Chemical constituents of *Hoya cagayanensis* CM Burton. *Journal of Applied Pharmaceutical Science*, 7(5), pp. 061-065.
- 114 Reza, M. S. H., Mandal, C., Alam, K. A., Salam, A., Rahman, M. A., Amin, M. R., ... & Ahmed, F. (2007). Phytochemical, antibacterial and antinociceptive studies of *Hoya parasitica*. *CABI Databases*, pp. 753-756
- 115 Ribeiro, F. F., Mendonca Junior, F. J., Ghasemi, J. B., Ishiki, H. M., Scotti, M. T., & Scotti, L. (2018). Docking of natural products against neurodegenerative diseases: general concepts. *Combinatorial chemistry & high throughput screening*, 21(3), pp. 152-160.
- 116 Righetti, P. G., & Boschetti, E. (2013). Low-Abundance Protein Access by Combinatorial Peptide Libraries. *Low-Abundance Proteome Discovery; Elsevier: Amsterdam, The Netherlands*, pp. 79-157.
- 117 Rodda, M. (2015). Two new species of *Hoya* R. Br.(Apocynaceae, Asclepiadoideae) from Borneo. *PhytoKeys*, (53), pp. 83.
- 118 Rodda, M., & Niissalo, M. A. (2021). Plastome evolution and organisation in the *Hoya* group (Apocynaceae). *Scientific Reports*, 11(1), pp. 14520.
- 119 Rodda, M., Simonsson, N., Ercole, E., Khew, G., Niissalo, M., Rahayu, S., & Livshultz, T. (2020). Phylogenetic studies in the *Hoya* group (Apocynaceae, Marsdenieae): the position of *Anatropanthus* and *Oreosparte*. *Willdenowia*, 50(1), pp. 119-138.
- 120 Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular biology and evolution*, 34(12), pp. 3299-3302.

- 121 Rumaling, M. K., Fong, S. Y., Rao, P. V., Gisil, J., Sani, M. H. M., & Wan Saudi, W. S. (2025). Pharmacological properties of *Hoya* (Apocynaceae): A systematic review. *Natural Product Research*, 39(4), pp. 905-921.
- 122 Sadhu, S. K., Khatun, A., Ohtsuki, T., & Ishibashi, M. (2008). Constituents from *Hoya parasitica* and their cell growth inhibitory activity. *Planta medica*, 74(07), pp. 760-763.
- 123 Sadowska, B., Budzyńska, A., Więckowska-Szakiel, M., Paszkiewicz, M., Stochmal, A., Moniuszko-Szajwaj, B., ... & Różalska, B. (2014). New pharmacological properties of *Medicago sativa* and *Saponaria officinalis* saponin-rich fractions addressed to *Candida albicans*. *Journal of Medical Microbiology*, 63(8), pp. 1076-1086.
- 124 Sager, R., & Lane, D. (1972). Molecular basis of maternal inheritance. *Proceedings of the National Academy of Sciences*, 69(9), pp. 2410-2413.
- 125 Sánchez, M., González-Burgos, E., Iglesias, I., Lozano, R., & Gómez-Serranillos, M. P. (2020). The pharmacological activity of *Camellia sinensis* (L.) Kuntze on metabolic and endocrine disorders: a systematic review. *Biomolecules*, 10(4), pp. 603.
- 126 Savarino, P., Demeyer, M., Decroo, C., Colson, E., & Gerbaux, P. (2023). Mass spectrometry analysis of saponins. *Mass Spectrometry Reviews*, 42(3), pp. 954-983.
- 127 Savoia, D. (2012). Plant-derived antimicrobial compounds: alternatives to antibiotics. *Future microbiology*, 7(8), pp. 979-990.
- 128 Seeff, L. B., Bonkovsky, H. L., Navarro, V. J., & Wang, G. (2015). Herbal products and the liver: a review of adverse effects and mechanisms. *Gastroenterology*, 148(3), pp. 517-532.
- 129 Shah, A., Varma, C., Patankar, S., & Kadam, V. (2013). Plant Glycosides and Aglycones Displaying Antiproliferative and Antitumour Activities—A Review. *Current Bioactive Compounds*, 9(4), pp. 288-305.
- 130 Sharma, P., Tyagi, A., Bhansali, P., Pareek, S., Singh, V., Ilyas, A., ... & Poddar, N. K. (2021). Saponins: Extraction, bio-medicinal properties and way forward to anti-viral representatives. *Food and Chemical Toxicology*, 150, pp. 112075.
- 131 Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., ... & Sugiura, M. (1986). The complete nucleotide sequence of the tobacco

chloroplast genome: its gene organization and expression. *The EMBO journal*, 5(9), pp. 2043-2049.

- 132 Shokry, A., El-Shiekh, R., Kamel, G., & Ramadan, A. (2022). Phytochemical contents, biological activities and therapeutic applications of heder helix (ivy leaf) extracts: a review. *The Natural Products Journal*, 12(4), pp. 22-32.
- 133 Sindhu, R. K., Goyal, A., Algin Yapar, E., & Cavalu, S. (2021). Bioactive compounds and nanodelivery perspectives for treatment of cardiovascular diseases. *Applied Sciences*, 11(22), pp. 11031.
- 134 Singh, B. P., Kumar, A., Kaur, H., Singh, H., & Nagpal, A. K. (2020). CpGDB: a comprehensive database of chloroplast genomes. *Bioinformatics*, 16(2), pp. 171.
- 135 Singh, V. K., Arora, D., Ansari, M. I., & Sharma, P. K. (2019). Phytochemicals based chemopreventive and chemotherapeutic strategies and modern technologies to overcome limitations for better clinical applications. *Phytotherapy research*, 33(12), pp. 3064-3089.
- 136 Skehan, P., Storeng, R., Scudiero, D., Monks, A., McMahon, J., Vistica, D., ... & Boyd, M. R. (1990). New colorimetric cytotoxicity assay for anticancer-drug screening. *Journal of the National Cancer Institute*, 82(13), pp. 1107-1112.
- 137 Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), pp. 1312-1313.
- 138 Stan, D., Enciu, A. M., Mateescu, A. L., Ion, A. C., Brezeanu, A. C., Stan, D., & Tanase, C. (2021). Natural compounds with antimicrobial and antiviral effect and nanocarriers used for their transportation. *Frontiers in pharmacology*, 12, pp. 723233.
- 139 Su, H., Ding, X., Liao, B., Zhang, D., Huang, J., Bai, J., ... & Huang, Z. (2023). Comparative chloroplast genomes provided insights into the evolution and species identification on the Datureae plants. *Frontiers in Plant Science*, 14, pp. 1270052.
- 140 Subramaniam, S., Selvaduray, K. R., & Radhakrishnan, A. K. (2019). Bioactive compounds: natural defense against cancer?. *Biomolecules*, 9(12), pp. 758.

- 141 Süleyman, H., Mshvildadze, V., Gepdiremen, A., & Elias, R. (2003). Acute and chronic antiinflammatory profile of the ivy plant, *Hedera helix*, in rats. *Phytomedicine*, 10(5), pp. 370-374.
- 142 Sun, H. X., Xie, Y., & Ye, Y. P. (2009). Advances in saponin-based adjuvants. *Vaccine*, 27(12), pp. 1787-1796.
- 143 Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*, 10(3), pp. 512-526.
- 144 Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, 38(7), pp. 3022-3027.
- 145 Tamura, Y., Miyakoshi, M., & Yamamoto, M. (2012). Application of saponin-containing plants in foods and cosmetics. *Alternative medicine*, pp. 85-101.
- 146 Tan, X. H., Wang, J. H., Zhao, K. K., Zhu, Z. X., & Wang, H. F. (2018). Complete plastome sequence of *Hoya pottsii* Traill and *Hoya liangii* Tsiang (Apocynaceae). *Mitochondrial DNA Part B*, 3(2), pp. 1176-1177.
- 147 Tanemoto, R., Okuyama, T., Matsuo, H., Okumura, T., Ikeya, Y., & Nishizawa, M. (2015). The constituents of licorice (*Glycyrrhiza uralensis*) differentially suppress nitric oxide production in interleukin-1 β -treated hepatocytes. *Biochemistry and biophysics reports*, 2, pp. 153-159.
- 148 Tania, U. H., Hassan, M. R., Eshita, N. J., Akhter, R., & Shahriar, M. (2016). Evaluation of *In vitro* Antioxidant and *In vivo* Pharmacological Activity of Leaf Extracts of *Hoya parasitica* (Wall.). *Journal of Applied Pharmaceutical Science*, 6(05), pp. 163-170.
- 149 Tasneem, S., Liu, B., Li, B., Choudhary, M. I., & Wang, W. (2019). Molecular pharmacology of inflammation: Medicinal plants as anti-inflammatory agents. *Pharmacological research*, 139, pp. 126-140.
- 150 Tatia, R., Moldovan, L., Tarcomnicu, I., Raiciu, A. D., Gavrilă, A., Calinescu, I., & Zalaru, C. (2022). Methods of Obtaining Extracts from *Hedera helix* L. Leaves and Evaluation of the Total Saponins Content. *Chemistry Proceedings*, 7(1), pp. 56.

- 151 Terefe, E. M., & Ghosh, A. (2022). Molecular docking, validation, dynamics simulations, and pharmacokinetic prediction of phytochemicals isolated from *Croton dichogamus* against the HIV-1 reverse transcriptase. *Bioinformatics and Biology Insights*, 16, pp. 1-20.
- 152 Tewari, K. K., Wildman S. G. (1970). Control of organelle development. *Cambridge University Press, London*. pp 147-79.
- 153 Tian Xin, Tian Xin, and Li DeZhu Li DeZhu. (2002). "Application of DNA sequences in plant phylogenetic study.". *CABI Databases*. pp. 170-184.
- 154 Tran, H. H. T., Nguyen, M. C., Le, H. T., Nguyen, T. L., Pham, T. B., Chau, V. M., ... & Nguyen, T. D. (2014). Inhibitors of α -glucosidase and α -amylase from *Cyperus rotundus*. *Pharmaceutical Biology*, 52(1), pp. 74-77.
- 155 Tu, T. Q., Nguyen, Y. T. H., Thi Ngoc Nguyen, L., Duc Nguyen, H., & Chu, M. H. (2025). Three new minor steroidal glycosides from the whole plants of *Hoya parasitica* (Wall. ex Hornem.) Wight. *Natural Product Research*, 39(2), pp. 269-274.
- 156 Unavane, S. A., Syed, S., Jain, H. K., & Bansode, A. (2023). Docking and pharmacokinetic studies for screening terpenoids from erythroxyllum species as anticancer agents. *Letters in Drug Design & Discovery*, 20(12), pp. 2025-2033.
- 157 Uncu, A. O., Uncu, A. T., Celik, İ., Doganlar, S., & Frary, A. (2015). A primer to molecular phylogenetic analysis in plants. *Critical Reviews in Plant Sciences*, 34(4), pp. 454-468.
- 158 Vaghasiya, M. D., Mendapara, J. V., Vaghasiya, S. P., Rajani, D. P., Ahmad, I., Patel, H., & Kumari, P. (2023). Identification of novel 4-thiazolidinones as new TcaR inhibitors: Design, synthesis, molecular docking, MD simulation, ADMET and *in vitro* antibacterial evaluation. *Journal of Molecular Structure*, 1294, pp. 136308.
- 159 Vinarova, L., Vinarov, Z., Atanasov, V., Pantcheva, I., Tcholakova, S., Denkov, N., & Stoyanov, S. (2015). Lowering of cholesterol bioaccessibility and serum concentrations by saponins: *in vitro* and *in vivo* studies. *Food & Function*, 6(2), pp. 501-512.
- 160 Wang, J., Kan, S., Liao, X., Zhou, J., Tembrock, L. R., Daniell, H., ... & Wu, Z. (2024). Plant organellar genomes: much done, much more to do. *Trends in Plant Science*, 29(7), pp. 754-769.

- 161 Wang, L., Li, S., Yao, Y., Yin, W., & Ye, T. (2021). The role of natural products in the prevention and treatment of pulmonary fibrosis: a review. *Food & function*, 12(3), pp. 990-1007.
- 162 Wang, Y., Zhang, C. F., Odago, W. O., Jiang, H., Yang, J. X., Hu, G. W., & Wang, Q. F. (2023). Evolution of 101 Apocynaceae plastomes and phylogenetic implications. *Molecular phylogenetics and evolution*, 180, pp. 107688.
- 163 Wanntorp, L. (2009). Phylogenetic systematics of *Hoya* (Apocynaceae). *Blumea-Biodiversity, Evolution and Biogeography of Plants*, 54(1-2), pp. 228-232.
- 164 Wanntorp, L., Grudinski, M., Forster, P. I., Muellner-Riehl, A. N., & Grimm, G. W. (2014). Wax plants (*Hoya*, Apocynaceae) evolution: Epiphytism drives successful radiation. *Taxon*, 63(1), pp. 89-102.
- 165 Wanntorp, L., Kocyan, A., & Renner, S. S. (2006). Wax plants disentangled: A phylogeny of *Hoya* (Marsdenieae, Apocynaceae) inferred from nuclear and chloroplast DNA sequences. *Molecular phylogenetics and evolution*, 39(3), pp. 722-733.
- 166 Wanntorp, L., Kocyan, A., Van Donkelaar, R., & Renner, S. S. (2006). Towards a monophyletic *Hoya* (Marsdenieae, Apocynaceae): Inferences from the chloroplast trnL region and the rbcL-atpB spacer. *Systematic Botany*, 31(3), pp. 586-596.
- 167 Wei, X. F., Zeng, S. J., Zhang, G. Q., Tang, G. D., & Huang, J. X. (2020). Complete plastome sequence of *H. carnosa* (L. f.) R. Br.(Apocynaceae). *Mitochondrial DNA Part B*, 5(1), pp. 522-523.
- 168 Weng, L., Jiang, Y., Wang, Y., Zhang, X., Zhou, P., Wu, M., ... & Chen, S. (2023). Chloroplast genome characteristics and phylogeny of the sinodielsia clade (apiaceae: apiioideae). *BMC Plant Biology*, 23(1), pp. 284.
- 169 Wieczorek, G., & Niedzialek, D. (2020). Molecular Dynamics. *Encyclopedia of Life Sciences*, pp. 1-18.
- 170 Wijesekara, T., Luo, J., & Xu, B. (2024). Critical review on anti-inflammation effects of saponins and their molecular mechanisms. *Phytotherapy Research*, 38(4), pp. 2007-2022.

- 171 Xiao, H., Xue, Q., Zhang, Q., Li, C., Liu, X., Liu, J., ... & Yang, J. (2019). How ginsenosides trigger apoptosis in human lung adenocarcinoma cells. *The American Journal of Chinese Medicine*, 47(08), pp. 1737-1754.
- 172 Yang, T., Liao, X., Yang, L., Liu, Y., Mu, W., Sahu, S. K., ... & Liu, H. (2019). Comparative analyses of 3654 chloroplast genomes unraveled new insights into the evolutionary mechanism of green plants. *bioRxiv*, pp. 655241.
- 173 Yazaki, K., Arimura, G. I., & Ohnishi, T. (2017). Hidden terpenoids in plants: their biosynthesis, localization and ecological roles. *Plant and Cell Physiology*, 58(10), pp. 1615-1621.
- 174 Zhang, G., & Ma, H. (2024). Nuclear phylogenomics of angiosperms and insights into their relationships and evolution. *Journal of Integrative Plant Biology*, 66(3), pp. 546-578.
- 175 Zhang, N., Ramachandran, P., Wen, J., Duke, J. A., Metzman, H., McLaughlin, W., ... & Handy, S. M. (2017). Development of a reference standard library of chloroplast genome sequences, GenomeTrakrCP. *Planta medica*, 83(18), pp. 1420-1430.

PHỤ LỤC

1. Phụ lục PL1. Các trình tự hệ gene lục lạp của họ Apocynaceae tải xuống từ GenBank được sử dụng trong phân tích phát sinh chủng loại

TT	Tên loài	Mã truy cập trên GenBank
1	<i>Cynanchum sibiricum</i>	OQ390041.2
2	<i>Cynanchum wilfordii</i>	KT220733.1
3	<i>Dregea sinensis</i> var. <i>corrugata</i>	OP133575.1
4	<i>Gongronemopsis tenacissima</i>	MW861760.1
5	<i>Gymnema latifolium</i>	OP133579.1
6	<i>Gymnema sylvestre</i>	MN711723.1
7	<i>Gymnema yunnanense</i>	OP133577.1
8	<i>Vincetoxicum shaanxiense</i>	MH210646.1
9	<i>Lygisma inflexum</i>	OP133573.1
10	<i>Stephanotis volubilis</i>	OP133576.1
11	<i>Jasminanthes maingayi</i>	MW719056.1
12	<i>Marsdenia tinctoria</i>	OP133574.1
13	<i>Leichhardtia flavescens</i>	MW719052.1
14	<i>Popuahoya urniflora</i>	MW719062.1
15	<i>Hoya verticillata</i>	OR475244.1
16	<i>Hoya angustifolia</i>	OL754669.1
17	<i>Hoya ariadna</i>	OL754671.1
18	<i>Hoya carnosa</i>	OK539590.1
19	<i>Hoya caudata</i>	OL754672.1
20	<i>Hoya commutata</i>	OL826862.1
21	<i>Hoya dimorpha</i>	OL826863.1
22	<i>Hoya exilis</i>	MW719054.1
23	<i>Hoya griffithii</i>	OL754668.1
24	<i>Hoya ignorata</i>	MW719061.1
25	<i>Hoya kerrii</i>	OL754673.1
26	<i>Hoya lacunosa</i>	OL754667.1
27	<i>Hoya lanceolata</i>	OL826864.1
28	<i>Hoya liangii</i>	OL826865.1
29	<i>Hoya lithophytica</i>	MW719058.1
30	<i>Hoya longifolia</i>	OL754662.1
31	<i>Hoya lyi</i>	MW719055.1
32	<i>Hoya megalaster</i>	MW719063.1
33	<i>Hoya meliflua</i>	OL754674.1
34	<i>Hoya monetteae</i>	MW719053.1
35	<i>Hoya omlorii</i>	MW719060.1
36	<i>Hoya ovalifolia</i>	OL754666.1
37	<i>Hoya pandurata</i>	OL754665.1
38	<i>Hoya pottsii</i>	OL754664.1
39	<i>Hoya pubicalyx</i>	OL754663.1
40	<i>Hoya radicalis</i>	OL826866.1

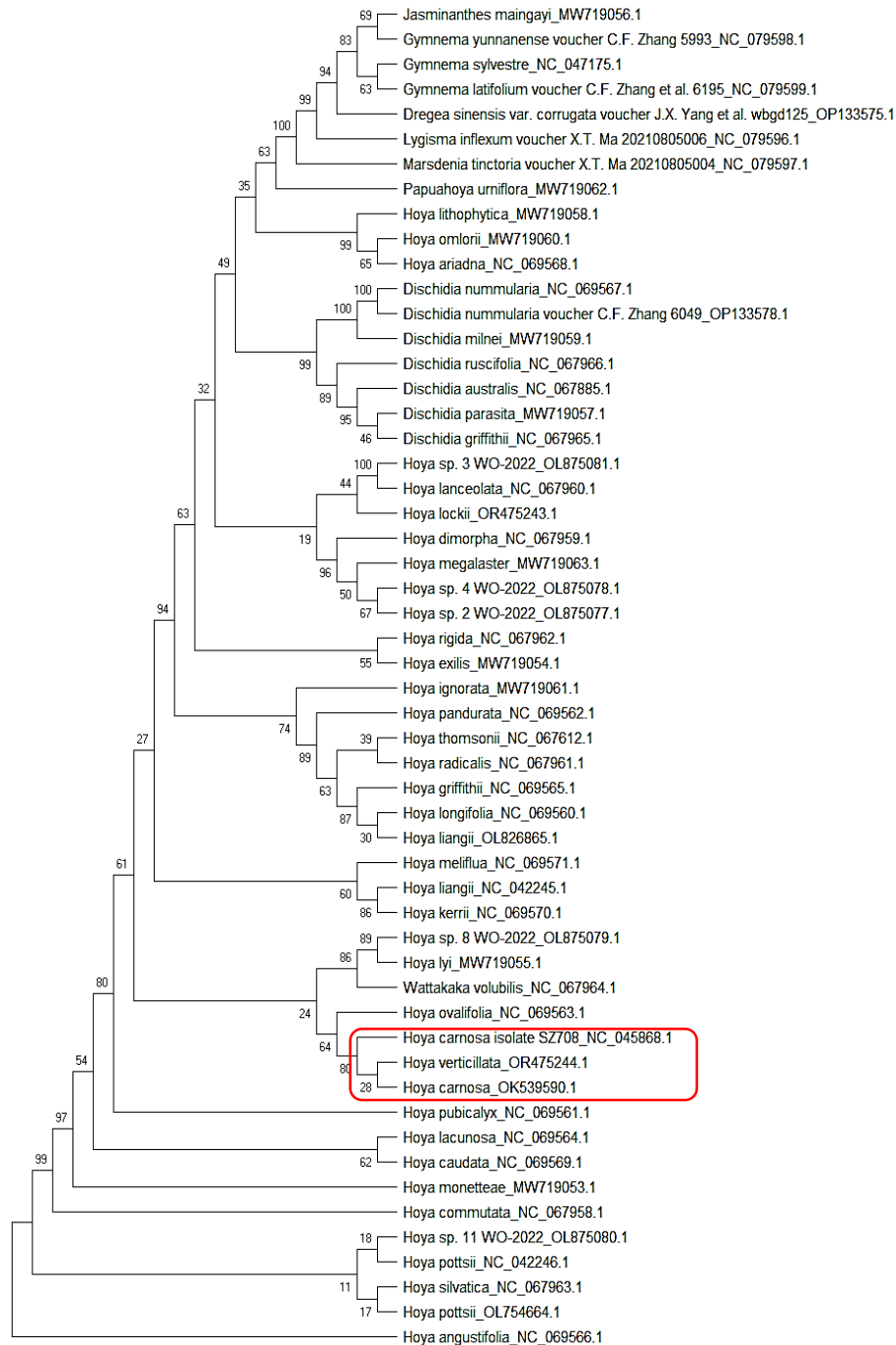
41	<i>Hoya rigida</i>	OL826867.1
42	<i>Hoya silvatica</i>	OL826868.1
43	<i>Hoya thomsonii</i>	OL435944.1
44	<i>Dischidia australis</i>	OL790122.1
45	<i>Dischidia griffithii</i>	OL839301.1
46	<i>Dischidia milnei</i>	MW719059.1
47	<i>Dischidia nummularia 1</i>	OP133578.1
48	<i>Dischidia nummularia 2</i>	OL754670.1
49	<i>Dischidia parasita</i>	MW719057.1
50	<i>Dischidia ruscifolia</i>	OL839302.1

2. Phụ lục PL2. Kích thước và trình tự nucleotide của các gene *ndhF*, *matK* và *rpl16*

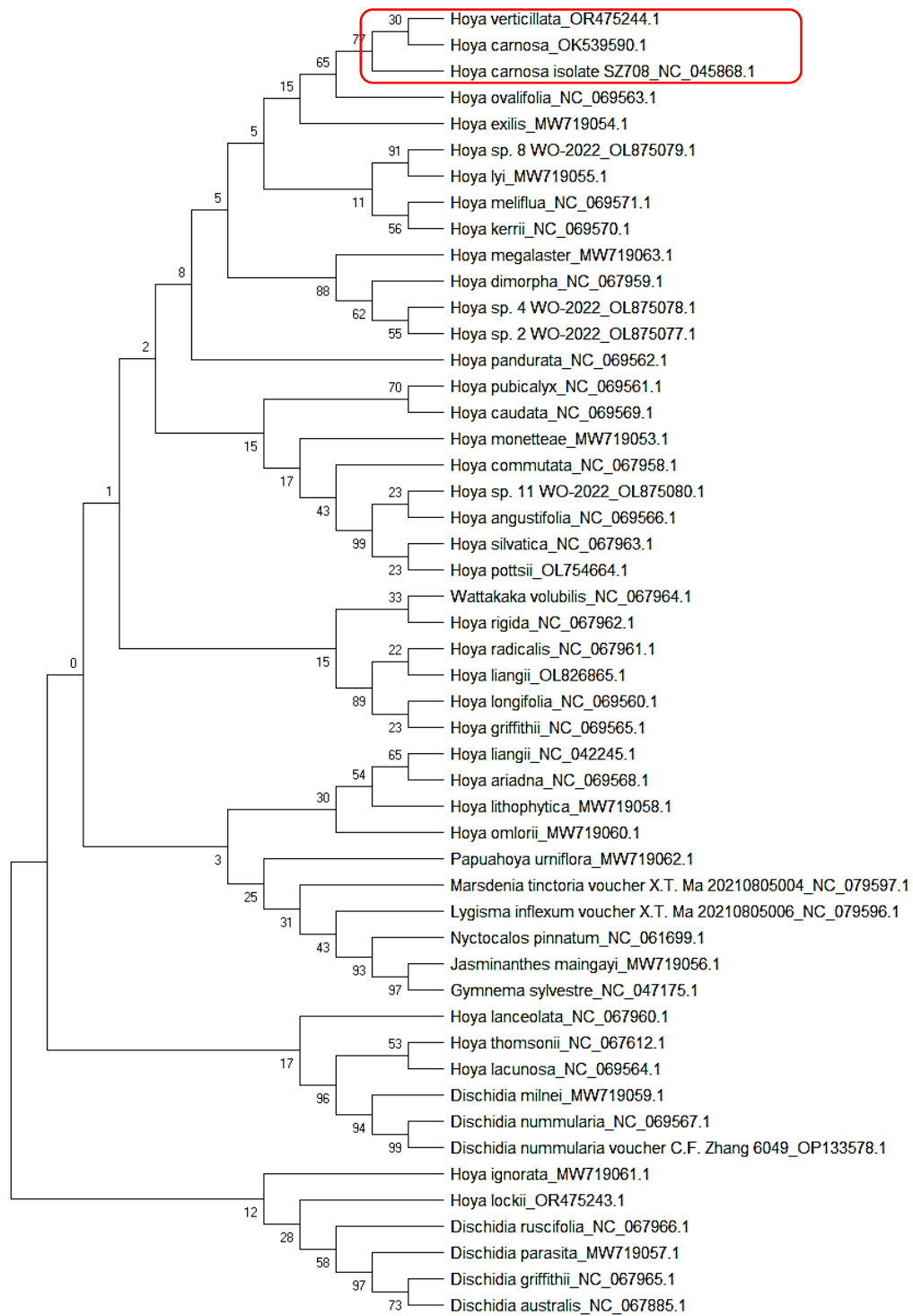
Gene	Trình tự nucleotide (5'-3')
<i>ndhF</i> 2241 bp	TTAAGTATAAGTAAGAAAAGTAATAAAAAAATTAAAAAACAAAAGAATACATACGATAAATAGA AGAAAAGATACGAAGAGATACGTCCGCCCCCTACATATTTTAGAACTTCTCCTGTAAAGAACTA AGAAAACCAACTCCATTCTGAATTCCATCAATTACTCGTCGATCAAAAAAATGAGTTAATTCAGC CAATGCTCTAGTCCCCCAGTTAGGGATCTTTTATAAAAAAGAATCTATATAAGCGCGATTATATG ACCAACAATATATGCCATTTAGCATTTTCTCCGAAAGAGGGCCGCTAGGGAACCCCTTGCCTTTA AGAGTTGAATTTCTTAAATCCAAATTTAATAAAGAGGAATAAGGAGATTTATATAAAAAAGACGC TATAAATATTCCTAAATAACCTATACTGACTGAAAAAAGGAATCTTTGAGAAATTCATACCAAT TGGTCGAATTAGTCAACTTTTTATACAAAAGATTGATCGACGGAGTTAACCATTTAGATAATATA TCAGCATTGACTCCCTCTTGATTAAAAGGAATTCCTATAAACCCAACGAACAGAGTAAATAGTCC CAATACAAGTAGAGGTAATAACATATTATTGTCTGATTCAATAAGGATAAGCATACAGTTTTTTAT TCTCAAAATGAGGAGTAGTAATAAAAGGCCGTGTCTATTTCTACCATTATCATCAATTCGATAT GTTCTTTTTGAAAAAAGGAAAAAATTTTATCATCAGTCATTGCTAATAAACGAACCTTTTTTATT AATTTTTTTTGAAACCCCCCTACCCCATAGAGATATTGAATAGAACGGTATTTTTTGTGTGCCAC TATAATTTGTAAAATAAACATTTAAATGTCCCTCAAAAGTAAGTAAATATATCCGAAACATATAA AATGCGGTTAATCCAGCAGTAACCCAGGCTATTATTGCGAAAATCGTCGAATACGACCAAGTATC ATTAATAATTTTCATCTTTGGACCAAAAACAAGCCAAAGGCGGAATACCACACAGAGAGAGTGTAC CTAATAAAAAAGCATTTTTTGGTAATTGGTACATGTTTTTCTTAAACCTCCCATAAGAATCATATTC TGATTTTTTATACGGAGAATAGCCAACAATCCCTTCCATTGAATGAATAACGGATCCAGATCCCTAA AAATAACAAATGCTTTTTGAATAGGCATGAGTAATCAAATGAAATAAAGCACTTCGGTAAGACCCCA TACCAAGAGCTAACATTATATAACCCAATTGGGACATTGTAGAATAAGCTAAACCTCTCTTAATG TCTTTTTGAGCAAGAGCTAAAGTAGCTCCCAACAAGACAGTTATTATTCTATTAACGAGATTAA ATTCAATTATGGAAGGTATAACTATGAAAAGAGGAAGAAGACGAGCTACAAGAAAAATTCGCGCCG CTACCATAGTAGCAGCGTGTATAAGGGCGGAAATCGGAGTAGGCCCTTCCATAGCATCAGGCAAC CATACATGAAGGGGAAATTGTGCAGATTTAGCAACAGCACCGCCAAATAACAGAGCAGCACACAA AGTAACAAAATAAAAAATTTACCTCGTTATTAGAAATTAAGTCATTGAATATTTTGAATAAATCTT GAAATTCGAAACTACCCGTTATCCAATAAAAACCTAAAATTCCTAATAATAAACCAAAATCCCT ACACGATTTGTTACAAAAGCGTTTTTGGCAAGCATTTGCTGCAAGAGGTCGTGTGGACCAAAATCC TATTAATAAATAGGAACACATTCGACCAATTCCTAAAAAATATAAATTTGTATCAAATTTGAAC TAGTAACCTAATCCTAACATGGAAGTACTGAAAAAATCATATAAGCAAAAAATTTCAAATATCCT TGATCATGAGCCATATAATTATCACTATAAATAAGAACAAAAATTCACACAGTAGTGATTAACAT TGACATAATAGAAGTAATTGGATCTATCAAATATCCGAATTCTAAAGAAAAATCGTTAGTAATGA TCCAAGACCATACATATTGATAGCTATAATTGCTATTTATTTGCTGAATAGACAGATTCATTGAA AAAATCATAACTATACTTAACAATAAAACACTATGAAAAGACCACATGCGACGAAACTTTTTTGT TGCTGTGCGAAAAAGAAGAGCCCAACCCCTAGTAATATAGCAACTGAAAGTGGGATGAAGAGTA TGAGGCACCCGTATTGATATGTCTGTTGCAT

<i>matK</i> 1548 bp	TCATTGCAGATTGGCCAGATCGTTGATACAAGTAATATCCAAATACCAAATCCGACTTCTATAGA CCGCCCAAAAAAGAAAAAACTCTTGGGAAGGTCAATGAAAGGACTTCTTCTTCCGACACGAAG AATTCTTCGAAAAATCCGAACCTAATCTTTTCAAAAAAGCCCGTACAGTACTTTTGTGTTTTTCG AGCTAAAGTTTTAGCACAAGAAAGTCGAAGTATATACTTTATTTCGATACAACTCTTTTTTTTGG AAGATCCGCTATGATAATGAGAAAGGTTTCTGCATATACACCCAAATCTGTCCATAATATCAGAA TCTGCTAAATCAGTCCGAACCGGCTTACTAATGGGATGTCCTAATAGGTTACAAAATTTCTCTTT AGACAATGATCCAATCAGAGGAATAATTGGAACAAGAGTATCTAATTTATTATTCTAATGGCAT TATTAATTAGAAATGAATTTTCTAGCATTTGACTCCGTACCCTTGAAGGGTTTAGTCGCACAATT GAAAGATAACCCAAAAAGTCAAGGGAATTTTGTACAATTGCTTTTATATGGGCCCTTCGCGAGTG AAACCACAGGTCAAAATTACATTGCCAAAACCTTGACAAGGTAAAATTTCCATTTATTCATCAAAA GAGCCGTCCCCCTTGAAAGCAGAATGGATTTTCCTTGATACCTAACATAATGCATGAAAGGATAT TTGACCAAGCACAAGTTCGTCGGAAAATCCTTAGCGAAGACTTCGGCAAGACGTTCAATTTTTTC ATAGAAAATAGATTCTGTTCAAGAAGGGCCCCAAAGGATGTTGGTCGTAAATGAGAAGATTGCTTAT GGAGAAAAGACGAAAATGTATTTCGTATTACATACATAAAAAATTATATAACAAGCAAAATAATCGT TGATTTCTTTTATAGATTTCTTTTTGTAAAAAAGAAAACTGGCTTTCTTTCTGGTTTTCTTTGT AGCACTAGGATTATTGCAATTACAATACTCGCGTAGAAAGAATCGTAATAAGTGCAAAGAAGGGG CATCTTTTACCCAATAGCGAAGGGTTTGAACCAAGATTTCCAGATGAACAGGGTAGGGTATTAGT ATATCTAATATAAAAAATAAAATGTGAAAAATTGTCTCTAAGAAAGGAAAGATTGAATGAATGGA TCGTAAATTATGAGATTTAAATATCCCTTTCCGCTCTTCAGAAGATATTAATCGCATGGAAAACG GAATTTCTACAATAAATGAAAATCCCTCTGATATCTTTTAGAATACAAATTTTTGTTGCGCCCA AAAAATGGATTTTGGCTAGAATCATTAGCAGAAGCAAGAAAATGATTCTGTTGATACATTGCAAT AATTAACCGTTTCGCAATAAGGAACTCGATTTATTGGCATAACCCGGATTTTCCAACAAAATTG CTCGTTTAAACTATGGTCATGGGCAAGTGCATAAATATATTCCTGAAAGATAAGTGGATATAGA AAGTTGGGCTGTTGAGATCTATCAAGCTGTAAATATCTTTGGATTTCTTCCAT
<i>rpl16</i> 992 bp	AGAATCGTGTCAATTTTTTGGAGACTTCATATAATCTAATCTATTAGTAATCTATAGATCTTGTTTAAAT AGACAATTAAAGATTTTAGTTTCTATTTTCGAAATCATATTGTTATTTTAAATTGTTATTTTGAAATA AAAATAGAACATATTTTTTTTTCTTTTATCGTCCAAATTTATCAATTCAAAAAAGAAAGTTTTTCG CGGGCGAATATTTACTCTTCTATTTCAATTTTCGGCTTGCTCCTGACTTCTTACAATAGATGAATTG ACCTCCGTTTCATTTCCGATCCCGACAGTGAATCATTAAAGATTCCTTTTTCACTAAAATCTTTTGC ATTCACAGCTTCCATCGTTCCCATCGCTTCTTACTTAATGCTTAGGTCCAATTTTTACAACGGAGCTC GTAATGAAATTTGTTCTTGAGTCAATCTTCTTAGTCTTTATTGGCTCGAAGCTCTTGATTTTTTTTGT TTTGTCTCTATAATTTTTAAATTTTTTAATTTAAAATTAATTTATGTTATATATTAAATTATATTAAAT TTATATTAAATAAGAAATATATTAAATTTATATTTAAAATATATATTAAATATAAAGATTTAATTATG TTATATAAGAATCAGTATTAATGCTTTATTACACTGCCTTTTATAAGATGACTTATAGACCTTACATA TTACATATTGGAATCTATATCATTGATATTTTTTCTCTCTCTTCTCTCATCCTTCCATTTATATCC ACATCTTTCTTCTCTATTTTGCCTTACAGCTTAAAAATCAGATTTTTTTTTCAGAAACACAAAATTC AGTTGTTACAAAGATATGACCAATATATCATATCTTGACTGGTTCCTTAGATCGAGATAATGCGAAGT AATGAGTTGGTTATTAGTTCTATGGTTAGTTATTAGTTCATACCTCTGGTGTGGGCTGGTCTTTTTTA ATCCTAACTCTAAAAACCAACGAGTCACACACTAAGCAT

3. Phụ lục PL3. Sơ đồ phát sinh chủng loại có độ tin cậy cao nhất (Maximum Likelihood) được xây dựng dựa trên trình tự vùng gene *matK*



4. Phụ lục PL4. Sơ đồ phát sinh chủng loại có độ tin cậy cao nhất (Maximum Likelihood) được xây dựng dựa trên trình tự vùng gene *rpl16*



5. Phụ lục PL5. Kích thước và trình tự các vùng đệm của hệ gene lục lạp Lan tai cáo

Vùng đệm	Trình tự nucleotide (5'-3')
<i>trnK-rps16</i> 973 bp	GAAAAGAAGAAACCAATTATAGAAGATAATCTATCGACTATAAAATATATGTATAATAAAAGT ATAATAAAAAAGAAGTGCATAAATACAATCGGAATTATAAAAAATTAAACAATAAATTAACCA AAGAAATTAGACGAAGCAATTAAACCATTAAAGCTAACAAATCGAAAAAATGGATTTTCTAATG AATTCGGAACATAAAAAAATATATAGATCGGATGACAATATAAGAAATTCTCAGATAAAATTA TATTTAGATAAAGCAAAACAAAAAGAAAAAAAAAAAAATGAAAATAGAGATTCAAAATTTATGA ACGGATCCTTTTGTTATCCCCTCTTTTCAATCAACAAAAAAGCTCGTGTATCAAAGAACCTG TCGACGTAAAGTAAAAATCAAACAAACCAATAACTAGATCTAACTTCACTTATCATGATGA ATTATATTTGTTCAATACACTGTTGTCAATATAAATGTTACGAAAAAAAAAGAATAAAAAAA AAAAAGAAATTCAATTGAATTTCTTTTTTTTTTATTCTTTTTTGTAGTTAAATTCAAAATAA AATGAAAACCATCAAGTTTAAAGTTGTGCGGAATGCCTAAATTTTCTAGGATCAAGGAAATA GGGTTTTGTGCTTATAGAACACAGGATTGGAAGAAAACAATGATAAATAGATACGAATAGAAC GGGAAAAGGAAAAAAGGGGGGTAATAAAAAAGAGTAGAGAAAGGTTATCCAAAGTTATATA CAAATAACTGCCCCCTTTTTTGTATTTCCTTAATCC
<i>rps16-trnQ</i> 1328 bp	AACATCCCTCTAATTTGGAACCAAGTATGGAATTGATTCAATTATGGAATCATGAATAGTCATT GGTTTGTAGACATCATAATCTATATCCCTTTGTTCTATAATGTTCTTTAGAATATAGAAGAGA GATTTTATATATAGCTATAGATCGGTAAATAAAGAATATAGCTACAAATCGGTAAAGAAGTTT CTGAATAAGTTTTAGCAAGTCCTCTTAATTTAAAAATGAAAAAGAATTTCTATTTAATAATATA TTTAATAATAGATTAAAGTTAGGGTTAAATATTTGTTAAATATTTTCATTTTAAATTCAAA ATCGAGGGGGTCAAAAACCTTTTTTCACAAAAATAGAATAGAAGGATACATATACGTTAAACAT TTCCTCGTTTTTTTTTATTTTGTAAAGCTGTGTAGGTGAGCAAGATTTTCGCTAATCGAATCT TTCCATACATAATATAATATCCATACATAATAGAATAGAAAGTGAATAAAAGAGAATAAAAGA TATAAACATAAAACACCCCCTTTTAAAGTGTACATAAATTTACAATTATTTATTATTTATTA GGGCCAAACACGAAATACTTATTCAAAGAAAATACATCGGATTTCGGATAGATATATTCGGATA GATATATAATTACATATATAATTTTATAATTTGATTTTTGTTAATGCAATTCAGGCAGACACA GAAATTTTAATATCTTCGGTTAATGTATTGCCCCCGGGTACTACAGGACTAATGGGACATA AGAGAAAAAAATGGGAATTATGCTCGGGGATGCGGACTGACTAGGTACAAATCCCAACAAGT CCAAATTAAGTTGCTCCTTTTTTTTTTATTTTAGTCTAACCCTTAGGAGAATTAATCCTATC GGCTTTGAATGAATTTTCATTAGTACCAAAATAGTTAGAATTAATAAATCTATATAAAAAATAAA AATTCCTAAAAATATAGTTTCTAGGATACGAAATAATAGATAGGATCGTTGAAAATCAAAATA TTGATAAAATATAAAATCAATATGGGACGGAATTTTTTTCTAAATAACTAACTTCCCGAAACA AGATGGGATGGAGCATAGGATGAAAAGACCCCCTTTTTTAAATTCCAACTCTGGGAACCGAT GTTCCCAATTTACCTGGATCAGAAATAGAGTCAATTACATTTGCGTGTCAATTTGTACGCGATT CAATTCAGTTTGTGTAAAAAGATATGATTCATGCATAAAAGATAAAAAAAAAAGAAATAAAA TCCAGCTAACATTAGACGTTACCCCATTCACAAAATCTTTATTCTATTCTAAGATAATGAAT ATGCT
<i>ndhC-trnV</i> 1492 bp	AGACACTCCCATGAACGTGGAAATATACCGAATTTCTCGATTGCAATTGTCAAGTCATCC ATAACTGTTTAGTCAAAATCAAAACAAGAATTCATTTTTATCGAATTTACTAGTTTCTTTTGT TTACTTATCCCATTTCAAGACAAAATCCTATTTAGATTATACTTATTTCTTTTTTTTTTTCAGTTA GTATAACTTTAGATCTATATCTTATTCTTATACAAATATAAAATAGAAAAATCTCTTATTTT CACCTAGAATTCTCTGTAAAGAAGAAAAGGAATTCAAAAAATATGGAATTCCAAGGGGAATT TGGAACCTCTTTATTTATTTTTTTTTTAATTAATAATTTTTTATTACTATCTATTATATATTTA TCTATTATATATTTATATTTTTTTGAAATTGATGAAATTGAATAAAATCAATTGAAATAAATGA AAAAAAAATATAGATTATTTCTTAATTTCTATAATATTTTCTATAATATTATAGCTATATTA TTATATTATAGATTCTTTATTATTTTTTTTCATTATTATTATTAATAATATTATATATTCT AGATTATTTATTATTATTTTTATTTCTATTTTTTTTTTATTAAAATTTGAATATTAATTAATAA TTTTTTATTATTAAATACAATCTAATAATACAATTTAAATTTCTATTTTCATTTGAAATTTGAAA TGTTTTAAATTGAATTGCTAATTTTTTTTTTATTTATTGAAATTCAATTCCTTTTTCTTTATAGA TTTCGACTTTTATGAATAGAATCGACAGGTCTTTGGTTTTTTTCTTAGTGATTTAGAATAAG AACAAGCAAGAGACTGGCTAGAAATTTCCATTTTCAAGATTTAGAATATCTTGGTGTGTATAT TCTTTTTATTAATATCCTAGCGGAGGACCTTCTCTTGATTGAATATAAAAAAGAAGACAGAC TACCCCTCGTGCTTCGCTAGGTCTAGGTAAGGTATATGGCGAGCCTATTTTACATTTTACATT GTTAATAATGAGACTTACCAAGATATTACATTTTTTAACTTGTACACACAGACAGAGGTCCT TTTCTAGATCCATATTTTGACTTCAAAAATTCACCAACATTGGCTATACCAAGAAAGGGAGT CTCAAGAACTCCTTGAATATGAAATTCGCTCTACCCCCAAGATTAATGCCCCAAGATTAATA ACAAAAGGTTTCTTTGTTTATCCGCGATTGGAAAATCAATTGGGTCCACTGAAACGCCTT

	TTTTTTTTTATTTTCGGGTTTATGTTCTGTTTTAAAGCTTGCCGTGAGTAAACTTATGGGAAA AATTTAGATTTTCGATTAACCAACCCGACAGTTCCAAGCAACAAACAATAATGAAAAAATTATA CAATTTTCATTTTTTTGTATTTGAGTTTTCTTAATCATTTTGTTTTATATTTTATATTAAAAATA TTTTATATATATATAGTATATAATTATAGTAATAGTAATAGTA
<i>psbI-atpA</i> 1878 bp	GGGTTTTTCTTTTCTATTTTCTTCGGATTTTTTGTTTATATAAAAAATGCCAAATTTGAAAAAA AATAAATAAATAAATAAAGTCATCAAGGAAAGCGGAAAGAGAGGGATTCTGAACCCCTCGGTACG AATAACTCGTACAACGGATTAGCAATCCGCCGCTTTAGTCCACTCAGCCATCTCTCCTAATTG AAAATTGGGTTAGATTACACATAAAAAAATAAGGAGTATTTCTTTCTCTATTATATAGATAT GTACCATTTTTTATCAGTAATTTTTTTTTTCGATAAATAAAGAAATAAAGAATAAATAAAGAGAAG GGCTCGAAAGTGCCAACAATATAAAGAAAACCAAAAGCGACCCCTTCGTATTTTGTTTCGAAAG ACCCTTCTTATTGTTATTGACACGGCCTGGTCTGGTCAGTACCCAGCCGGGCTCTTTTTGTT TTGTTCCAACATAATTATAGAAAAATAATGATAATTTTTCATTTATCTGATTTGAAAAACAAAA TGCTTAGTATTTTTTAGTATTATATATATATATTATAATATATATATATATATATATATAAT ATATAAAAAAGGGAATAGGCAATATTCAAGCACGAACAAAAAAGAAGGACTATTTCCCAT CCGCGAAAACGAAAAAAGAGGGTAAATACGCTAAATTTTTGGATTCTTTGATGATCCCAT CTTATTATACCCATTTTCCGGTTTCGACAAAAGGTCCAGTTGTATACAATAATCGAATTGTAGC GGGTATAGTTTAGTGGTAAAAGTGTGATTTCGTTTTTTTTTAACCCCTTGATAGTTAAAGGGTCTT TGTTTTTCGGTTTGATTTCGTATTCCGACCAAAAACTTTATTTGAAAAAGAAAAGGATTTAATC CTTTACCTCTCAATCACGGATTTCGAGAAAAAATATACATTCTCGTGATTTGTATCCAAGGATC ACTTAGAAAGTGACAAATTGGATTATGAAATTACGAAACATAATTTTGAATTGGATTAATAC TTCCAATTGAATCAGTATGAGTAAAGGATCCATGGATAAAGATAGCAAGTAGGTTTCTAATCG TAACTAAATCTTCAATTTTTGCTTTACAAATACAAAATTGAATAAAAAATAGCTATTAAATGAT GACTCTGGTTTACTAGAGGCATCGACCTGTTTTTTTTTAGCTCGGTGGAACAAAATCCCTTTCC TCAGGACCGTCTCAAATAAAAAATAGAGAACGAAGTAACTAGAAAGATTGTTAGAATTACTCTC TTCTAGAGGGATCATCTAGAAAGAAATAAGTAGTCAGACAAAAGTTGACATAGATGTTATGGG TAGAATTTTTTTTTTGTAAGTTTGTTTACATCCATATCCATAAAGGAGCCGAATGAAACCAAGT TTCATGTTTCGGTTTTGAATTAGAGACGTTCAAAATGCTGAATCGACGTCGACTATAACCCCTA GCCTTCCAAGCTAACGATGCGGGTTTCGATTCCCGCTACCCGCTTTCAATTCTTTTTTTTTATT CGCAATTGAAATTTATATATATTCTAATTCTAGACTTAATGCATCATTTAATATACAGTTTCA AAAATTATCTCACATACAATCCGATTCTTGTTTTTTTTTCGGCGAAAAAGTGGGGAAAGTCAAAA TACGAAAAAAATCGGAATGAAAAGCGTCCATTGTCTAATGGATAGGACAGGGGTCTTCTAAAC CTTTGGTATAGGTTCAAATCCTATTGGACGCAAATTTTTTCCATATATCTATTTTTTTTATATT TTGATACCACGAAAGCCTTTTTGAATAAAGTTGAATTTGAGTGGATTGAGACCCCTAATTACA TTACCATTACCTTTTTTAAGATAAATTTAAGATAAAAGAATCCTTTTTGTGC
<i>psbZ-rps14</i> 885 bp	ACCTATCGTCCCAGATCCAAAACCGAAAGGACCCCCGAATTTTTCTCGGTTGTAAGACACAT TAGAATTTAATATAAGTCCCCCAAAGAGAATAGAAAATGCAAATAAAGAAAAACAAAAAATTG AGGGGGGTCAAACCTTCTTCCAACTTATTAaaaaaATACAATTAATATTTATAATTTAATA AAATAAAAAATAATTGGAATCGAGTCGAGGAGAGTCTTTGGTCTGGGACTGCACAAAAATGATC GAGGTGTATATTTATATATGAGGTATAGAGGTTTATATTTATATATAAATTTTATATATAAAT AAATATATTAATATTGTGTGGACATATTATGTATCAAGAACGAAAAAAGCGGATATGGTTGA ATGGTAAAATTTCTCTTTGCCAAGGAGAAGATGTGGGTTTCGATTCCCGCTATCCGCCCACGAT AAAATTTATTATTAAATATAAAAAAATATTAAATAAAATTAGGTTAGGTTTATTAAATATAA TTTGGTATAGTTGTCTGTGATCGTGTAGTCATTCTATCCCCCTCTTTTTCCCTTTCTCTATAG CCTCCCCCAAAGTGGAATTAATTACTAGTTAACTTAACAAAGTAAATCTAAAAATTTTTTGA CAAACAAGATGTTGCGGAGACGGGATTTGAACCCGTGACCTCAAGGTTATGAGCCTTGCGAGC TACCAAAGTCTCTACTCCGCGATGAAGAGAAGAACTGAAACTAATAGATAACCAGGGGTTG AGTTGCCCCCTCTACCATATCTGTACAAATAGGATAGCCCATTTATACAGAATGGTAAAGGGT TATACAGAATGGCAAGGGGGACCTCTACGATCACCGACCCTAGAATAGAAATCTAGAGTTAAT CC
<i>rbcL-accD</i> 907 bp	TTACTTTTTGTTATCTTAGTTGAATTGTAATTAACTCGGCCCAATCTTTTCTTAAAGGATT GAGCCGAATACAAAGATTCTTTTTTTTTTTTTTATTTAATTATTTTTAAAGTATTTAATATTAA TACATACTTATATAGATATAAGATTTGAAATAAAAAAAAAAATCAAAAAATCTAAGGCTCAAA TCTTTCTATTGTTGTCTTGGACCTACAATTAATCTACGGATCCTTGGGATTGGGCTATTCTTT TATATCCTGCAGTTTTCTGAATCAAGCCAAGTATCACAATTCTTTCTATCCATCCCGTATAT TGTCCTTTTCTTTCCATATTGGTGAAATAGAACCTGAAATTTTTACGTTCTTGGGCGAAATTT TACGAAAAAAGGATTTATAAAATATTTTTTTTTTCGATGCGAATTTGATACGACATAAGAAAAG GCGCTCTTTATCATTGTATTTATAATGACAAGAGATTCCCTGATATCATATTAATATATTGAC AATATTCAAAAAAATGTATAAAAAAATGACAAGGGATTCTCTGATATCATATTAATATATAT TGACATTATTCAAAAAATAATGTATAATGTATAAAACAATGACAAGAGATTCCCTGATATGA

	TATATATATATAGTGTATAGTCATATATATCGTATAGTGAAGTATTACTCCGGATTTTCAAAA AAAGAGCATTTTTTTTCAATAATCACACCTAATACCTTCTTATTGTTTCTTATTGGTAACTTA TTAGTTAATAAAAAATGCTAGCGATTGGTTTAGTCTAATAGGAAAGAAGATATTCAAATAAAT AATTTTTGATCGAATGACTATTCATCTATTGTTTTTATATTTTCTTACAAAACAAAAAACAA AAAAAAGGGGGACAAAATAACTCT
--	---

6. Phụ lục PL6. Danh sách các loài và trình tự các vùng đệm được sử dụng trong phân tích phát sinh chủng loại

PL6.1. Danh sách các loài có trình tự vùng đệm *trnK-rps16* được sử dụng trong phân tích phát sinh chủng loại

TT	Tên loài	Mã truy cập trên GenBank
1	<i>Cynanchum wilfordii</i>	NC_029459.1
2	<i>Dregea sinensis</i> var. <i>corrugata</i>	OP133575.1
3	<i>Gymnema latifolium</i>	NC_079599.1
4	<i>Gymnema sylvestre</i>	NC_047175.1
5	<i>Gymnema yunnanense</i>	NC_079598.1
6	<i>Hoya angustifolia</i>	NC_069566.1
7	<i>Hoya carnosa</i>	NC_045868.1
8	<i>Hoya carnosa</i>	OK539590.1
9	<i>Hoya commutata</i>	NC_067958.1
10	<i>Hoya kerrii</i>	NC_069570.1
11	<i>Hoya liangii</i>	NC_042245.1
12	<i>Hoya lyi</i>	MW719055.1
13	<i>Hoya meliflua</i>	NC_069571.1
14	<i>Hoya ovalifolia</i>	NC_069563.1
15	<i>Hoya pottsii</i>	NC_042246.1
16	<i>Hoya pubicalyx</i>	NC_069561.1
17	<i>Hoya silvatica</i>	NC_067963.1
18	<i>Hoya verticillata</i>	OR475244.1
19	<i>Jasminanthes maingayi</i>	MW719056.1
20	<i>Leichhardtia flavescens</i>	MW719052.1
21	<i>Lygisma inflexum</i>	NC_079596.1
22	<i>Marsdenia tinctoria</i>	NC_079597.1
23	<i>Stephanotis volubilis</i>	OP133576.1
24	<i>Vincetoxicum shaanxiense</i>	NC_046785.1

PL6.2. Danh sách các loài có trình tự vùng đệm *rps16-trnQ* được sử dụng trong phân tích phát sinh chủng loại

TT	Tên loài	Mã truy cập trên GenBank
1	<i>Cynanchum wilfordii</i>	NC_029459.1
2	<i>Dischidia australis</i>	NC_067885.1
3	<i>Dischidia griffithii</i>	NC_067965.1
4	<i>Dischidia milnei</i>	MW719059.1
5	<i>Dischidia nummularia</i>	NC_069567.1
6	<i>Dischidia nummularia</i>	OP133578.1
7	<i>Dischidia parasita</i>	MW719057.1
8	<i>Dischidia ruscifolia</i>	NC_067966.1
9	<i>Dregea sinensis</i> var. <i>corrugata</i>	OP133575.1
10	<i>Gymnema latifolium</i>	NC_079599.1
11	<i>Gymnema sylvestre</i>	NC_047175.1
12	<i>Gymnema yunnanense</i>	NC_079598.1
13	<i>Hoya angustifolia</i>	NC_069566.1
14	<i>Hoya ariadna</i>	NC_069568.1
15	<i>Hoya carnosa</i>	NC_045868.1
16	<i>Hoya caudata</i>	NC_069569.1
17	<i>Hoya commutata</i>	NC_067958.1
18	<i>Hoya dimorpha</i>	NC_067959.1
19	<i>Hoya exilis</i>	MW719054.1
20	<i>Hoya griffithii</i>	NC_069565.1
21	<i>Hoya ignorata</i>	MW719061.1
22	<i>Hoya kerrii</i>	NC_069570.1
23	<i>Hoya lanceolata</i>	NC_067960.1
24	<i>Hoya liangii</i>	NC_042245.1
25	<i>Hoya longifolia</i>	NC_069560.1
26	<i>Hoya lyi</i>	MW719055.1
27	<i>Hoya megalaster</i>	MW719063.1
28	<i>Hoya meliflua</i>	NC_069571.1
29	<i>Hoya monetteae</i>	MW719053.1
30	<i>Hoya omlorii</i>	MW719060.1
31	<i>Hoya ovalifolia</i>	NC_069563.1
32	<i>Hoya pandurata</i>	NC_069562.1
33	<i>Hoya pottsii</i>	NC_042246.1
34	<i>Hoya pubicalyx</i>	NC_069561.1
35	<i>Hoya rigida</i>	NC_067962.1
36	<i>Hoya silvatica</i>	NC_067963.1
37	<i>Hoya thomsonii</i>	NC_067612.1
38	<i>Hoya verticillata</i>	OR475244.1
39	<i>Jasminanthes maingayi</i>	MW719056.1
40	<i>Leichhardtia flavescens</i>	MW719052.1

41	<i>Lygisma inflexum</i>	NC_079596.1
42	<i>Marsdenia tinctoria</i>	NC_079597.1
43	<i>Papuahoya urniflora</i>	MW719062.1
44	<i>Stephanotis volubilis</i>	OP133576.1

PL6.3. Danh sách các loài có trình tự vùng đệm *ndhC-trnV* được sử dụng trong phân tích phát sinh chủng loại

TT	Tên loài	Mã truy cập trên GenBank
1	<i>Cynanchum wilfordii</i>	NC_029459.1
2	<i>Dischidia australis</i>	NC_067885.1
3	<i>Dischidia griffithii</i>	NC_067965.1
4	<i>Dischidia milnei</i>	MW719059.1
5	<i>Dischidia nummularia</i>	OP133578.1
6	<i>Dischidia nummularia</i>	NC_069567.1
7	<i>Dischidia parasita</i>	MW719057.1
8	<i>Dischidia ruscifolia</i>	NC_067966.1
9	<i>Dregea sinensis</i> var. <i>corrugata</i>	OP133575.1
10	<i>Gymnema latifolium</i>	NC_079599.1
11	<i>Gymnema sylvestre</i>	NC_047175.1
12	<i>Gymnema yunnanense</i>	NC_079598.1
13	<i>Hoya angustifolia</i>	NC_069566.1
14	<i>Hoya ariadna</i>	NC_069568.1
15	<i>Hoya carnosa</i>	NC_045868.1
16	<i>Hoya caudata</i>	NC_069569.1
17	<i>Hoya commutata</i>	NC_067958.1
18	<i>Hoya dimorpha</i>	NC_067959.1
19	<i>Hoya exilis</i>	MW719054.1
20	<i>Hoya griffithii</i>	NC_069565.1
21	<i>Hoya ignorata</i>	MW719061.1
22	<i>Hoya kerrii</i>	NC_069570.1
23	<i>Hoya lacunosa</i>	NC_069564.1
24	<i>Hoya lanceolata</i>	NC_067960.1
25	<i>Hoya liangii</i>	NC_042245.1
26	<i>Hoya lithophytica</i>	MW719058.1
27	<i>Hoya longifolia</i>	NC_069560.1
28	<i>Hoya lyi</i>	MW719055.1
29	<i>Hoya megalaster</i>	MW719063.1
30	<i>Hoya meliflua</i>	NC_069571.1
31	<i>Hoya monetteae</i>	MW719053.1
32	<i>Hoya omlorii</i>	MW719060.1
33	<i>Hoya ovalifolia</i>	NC_069563.1
34	<i>Hoya pandurata</i>	NC_069562.1
35	<i>Hoya pottsii</i>	NC_042246.1
36	<i>Hoya pubicalyx</i>	NC_069561.1

37	<i>Hoya radicalis</i>	NC_067961.1
38	<i>Hoya rigida</i>	NC_067962.1
39	<i>Hoya silvatica</i>	NC_067963.1
40	<i>Hoya thomsonii</i>	NC_067612.1
41	<i>Hoya verticillata</i>	OR475244.1
42	<i>Jasminanthes maingayi</i>	MW719056.1
43	<i>Leichhardtia flavescens</i>	MW719052.1
44	<i>Lygisma inflexum</i>	NC_079596.1
45	<i>Marsdenia tinctoria</i>	NC_079597.1
46	<i>Papuahoya urniflora</i>	MW719062.1
47	<i>Stephanotis volubilis</i>	OP133576.1

PL6.4. Danh sách các loài có trình tự vùng đệm *psbI-atpA* được sử dụng trong phân tích phát sinh chủng loại

TT	Tên loài	Mã truy cập trên GenBank
1	<i>Dischidia australis</i>	NC_067885.1
2	<i>Dischidia griffithii</i>	NC_067965.1
3	<i>Dischidia milnei</i>	MW719059.1
4	<i>Dischidia nummularia</i>	NC_069567.1
5	<i>Dischidia nummularia</i>	OP133578.1
6	<i>Dischidia parasita</i>	MW719057.1
7	<i>Dischidia ruscifolia</i>	NC_067966.1
8	<i>Dregea sinensis</i> var. <i>corrugata</i>	OP133575.1
9	<i>Gymnema latifolium</i>	NC_079599.1
10	<i>Gymnema sylvestre</i>	NC_047175.1
11	<i>Gymnema yunnanense</i>	NC_079598.1
12	<i>Hoya angustifolia</i>	NC_069566.1
13	<i>Hoya ariadna</i>	NC_069568.1
14	<i>Hoya carnosa</i>	NC_045868.1
15	<i>Hoya caudata</i>	NC_069569.1
16	<i>Hoya dimorpha</i>	NC_067959.1
17	<i>Hoya exilis</i>	MW719054.1
18	<i>Hoya griffithii</i>	NC_069565.1
19	<i>Hoya ignorata</i>	MW719061.1
20	<i>Hoya kerrii</i>	NC_069570.1
21	<i>Hoya lacunosa</i>	NC_069564.1
22	<i>Hoya lanceolata</i>	NC_067960.1
23	<i>Hoya liangii</i>	OL826865.1
24	<i>Hoya lithophytica</i>	MW719058.1
25	<i>Hoya longifolia</i>	NC_069560.1
26	<i>Hoya lyi</i>	MW719055.1
27	<i>Hoya megalaster</i>	MW719063.1
28	<i>Hoya meliflua</i>	NC_069571.1

29	<i>Hoya monetteae</i>	MW719053.1
30	<i>Hoya omlorii</i>	MW719060.1
31	<i>Hoya ovalifolia</i>	NC_069563.1
32	<i>Hoya pandurata</i>	NC_069562.1
33	<i>Hoya pottsii</i>	OL754664.1
34	<i>Hoya pubicalyx</i>	NC_069561.1
35	<i>Hoya radicalis</i>	NC_067961.1
36	<i>Hoya rigida</i>	NC_067962.1
37	<i>Hoya silvatica</i>	NC_067963.1
38	<i>Hoya thomsonii</i>	NC_067612.1
39	<i>Hoya verticillata</i>	OR475244.1
40	<i>Jasminanthes maingayi</i>	MW719056.1
41	<i>Leichhardtia flavescens</i>	MW719052.1
42	<i>Lygisma inflexum</i>	NC_079596.1
43	<i>Marsdenia tinctoria</i>	NC_079597.1
44	<i>Papuahoya urniflora</i>	MW719062.1
45	<i>Stephanotis volubilis</i>	OP133576.1
46	<i>Vincetoxicum shaanxiense</i>	NC_046785.1

PL6.5. Danh sách các loài có trình tự vùng đệm *psbZ-rps14* được sử dụng trong phân tích phát sinh chủng loại

TT	Tên loài	Mã truy cập trên GenBank
1	<i>Hoya verticillata</i>	NC_085236.1
2	<i>Hoya carnosa</i>	NC_045868.1
3	<i>Hoya lyi</i>	MW719055.1
4	<i>Hoya kerrii</i>	NC_069570.1
5	<i>Hoya ariadna</i>	NC_069568.1
6	<i>Hoya monetteae</i>	MW719053.1
7	<i>Hoya thomsonii</i>	NC_067612.1
8	<i>Hoya ovalifolia</i>	NC_069563.1
9	<i>Hoya meliflua</i>	NC_069571.1
10	<i>Hoya silvatica</i>	NC_067963.1
11	<i>Hoya pottsii</i>	NC_042246.1
12	<i>Hoya angustifolia</i>	NC_069566.1
13	<i>Hoya megalaster</i>	MW719063.1
14	<i>Hoya lacunosa</i>	NC_069564.1
15	<i>Hoya dimorpha</i>	NC_067959.1
16	<i>Hoya lithophytica</i>	MW719058.1
17	<i>Hoya commutata</i>	NC_067958.1
18	<i>Hoya caudata</i>	NC_069569.1
19	<i>Hoya radicalis</i>	NC_067961.1
20	<i>Hoya omlorii</i>	MW719060.1
21	<i>Hoya rigida</i>	NC_067962.1

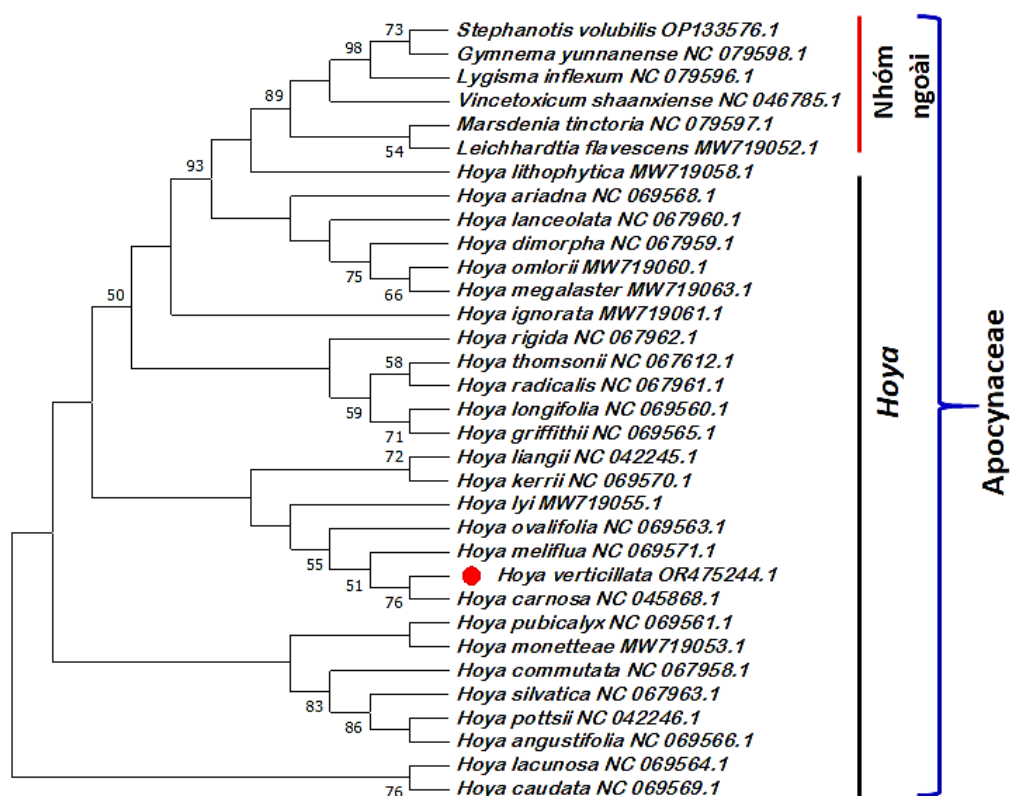
22	<i>Hoya liangii</i>	NC_042245.1
23	<i>Hoya pubicalyx</i>	NC_069561.1
24	<i>Hoya longifolia</i>	NC_069560.1
25	<i>Hoya griffithii</i>	NC_069565.1
26	<i>Hoya ignorata</i>	MW719061
27	<i>Hoya lanceolata</i>	NC_067960.1
28	<i>Leichhardtia flavescens</i>	MW719052.1
29	<i>Marsdenia tinctoria</i>	NC_079597.1
30	<i>Gymnema yunnanense</i>	NC_079598.1
31	<i>Stephanotis volubilis</i>	OP133576.1
32	<i>Lygisma inflexum</i>	NC_079596.1
33	<i>Vincetoxicum shaanxiense</i>	NC_046785.1

PL6.6. Danh sách các loài có trình tự vùng đệm *rbcL-accD* được sử dụng trong phân tích phát sinh chủng loại

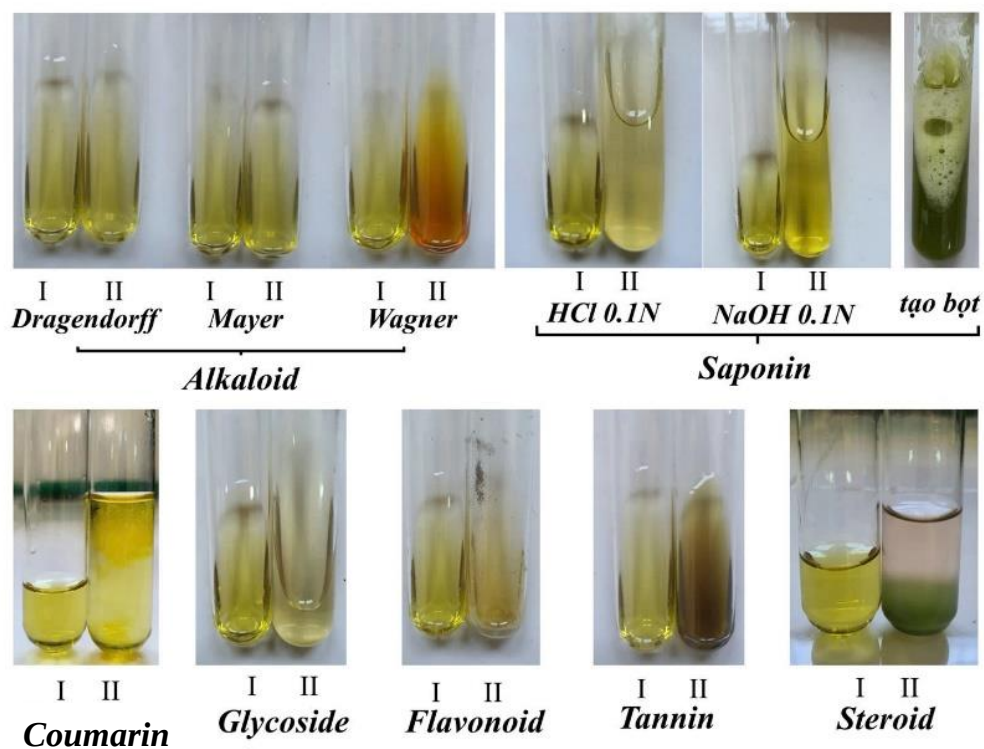
TT	Tên loài	Mã truy cập trên GenBank
1.	<i>Hoya verticilata</i>	NC_085236.1
2.	<i>Hoya monetteae</i>	MW719053.1
3.	<i>Hoya lyi</i>	MW719055.1
4.	<i>Hoya carnosae</i>	OK_539590.1
5.	<i>Hoya omlorii</i>	MW719060.1
6.	<i>Hoya ovalifolia</i>	NC_069563.1
7.	<i>Hoya radicalis</i>	NC_067961.1
8.	<i>Hoya rigida</i>	NC_067962.1
9.	<i>Hoya pubicalyx</i>	NC_069561.1
10.	<i>Hoya griffithii</i>	NC_069565.1
11.	<i>Hoya longifolia</i>	NC_069560.1
12.	<i>Hoya liangii</i>	OL_826865.1
13.	<i>Hoya kerrii</i>	NC_069570.1
14.	<i>Hoya dimorpha</i>	NC_067959.1
15.	<i>Hoya meliflua</i>	NC_069571.1
16.	<i>Hoya angustifolia</i>	NC_069566.1
17.	<i>Hoya silvatica</i>	NC_067963.1
18.	<i>Hoya pottsii</i>	OL_754664.1
19.	<i>Hoya lithophytica</i>	MW719058.1
20.	<i>Hoya caudata</i>	NC_069569.1
21.	<i>Hoya lacunosa</i>	NC_069564.1
22.	<i>Hoya exilis</i>	MW719054.1
23.	<i>Hoya pandurata</i>	NC_069562.1

24.	<i>Hoya commutata</i>	NC_067958.1
25.	<i>Hoya ariadna</i>	NC_069568.1
26.	<i>Hoya megalaster</i>	MW719063.1
27.	<i>Leichhardtia flavescens</i>	MW719052.1
28.	<i>Lygisma inflexum</i>	NC_079596.1
29.	<i>Stephanotis volubilis</i>	OP133576.1
30.	<i>Dregea sinensis</i> var. <i>corrugata</i>	OP133575.1
31.	<i>Marsdenia tinctoria</i>	NC_079597.1
32.	<i>Vincetoxicum shaanxiense</i>	NC_046785.1
33.	<i>Gymnema yunnanense</i>	NC_079598.1
34.	<i>Cynanchum sibiricum</i>	OQ390041.2

7. Phụ lục PL7. Sơ đồ phát sinh chủng loại dựa trên vùng đệm *psbZ-rps14*

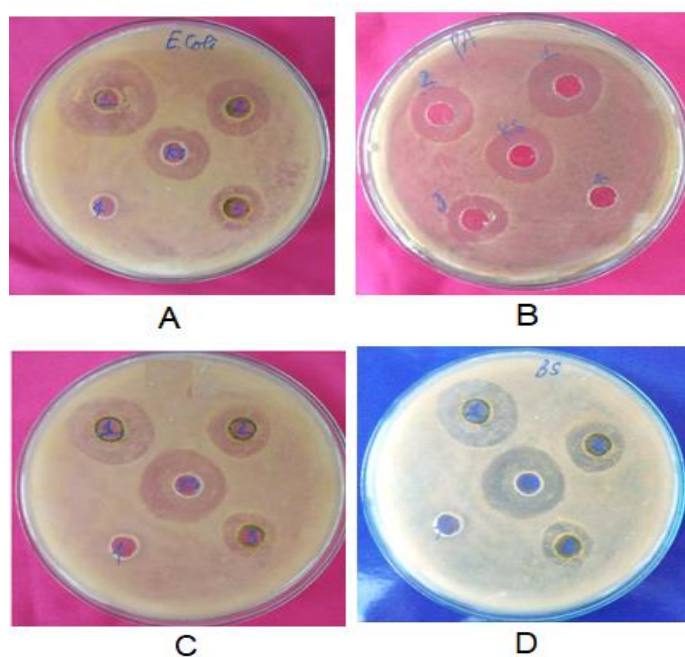


8. Phụ lục PL8. Hình ảnh thử nghiệm định tính các hợp chất có trong lá Lan tai cáo



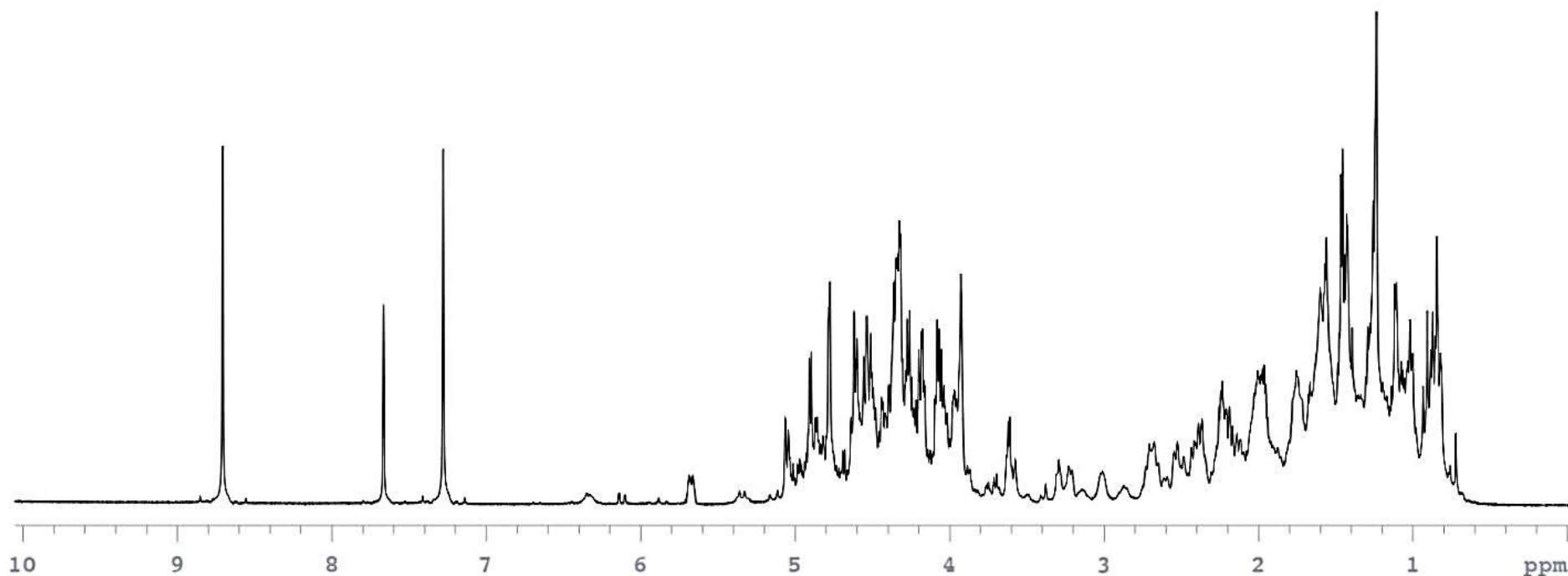
9. Phụ lục PL9. Hình ảnh vòng kháng khuẩn của cao chiết Lan tai cáo

A: *Escherichia coli*; B: *Pseudomonas aeruginosa*; C: *Staphylococcus aureus*; D: *Bacillus subtilis*. 1, 2, 3 là dung dịch cao tổng số với nồng độ lần lượt là 100 mg/mL, 50 mg/mL, 25 mg/mL; 4 là DMSO 1%; KS là kháng sinh Amoxicillin 50 mg/mL



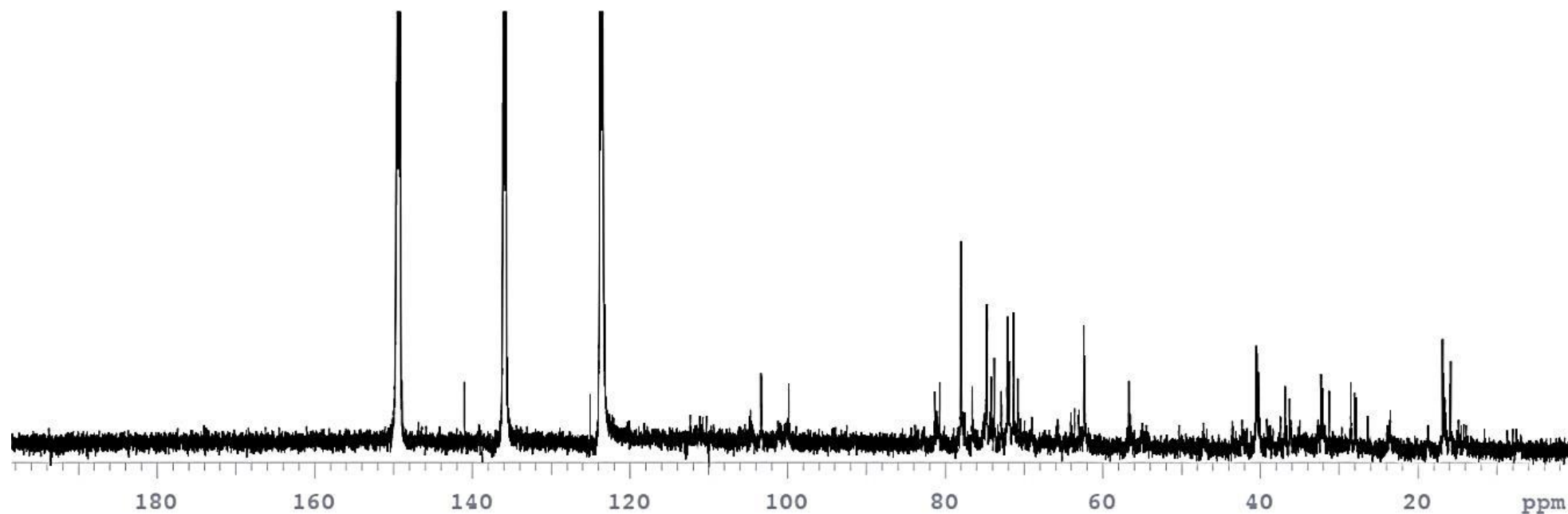
10. Phụ lục PL10. Dữ liệu phân lập, xác định saponin có trong Lan cai cáo.

PL10-1: Phổ cộng hưởng từ hạt nhân ^1H của chất **1** trong dung môi $\text{C}_5\text{D}_5\text{N}$ ở tần số 600 MHz



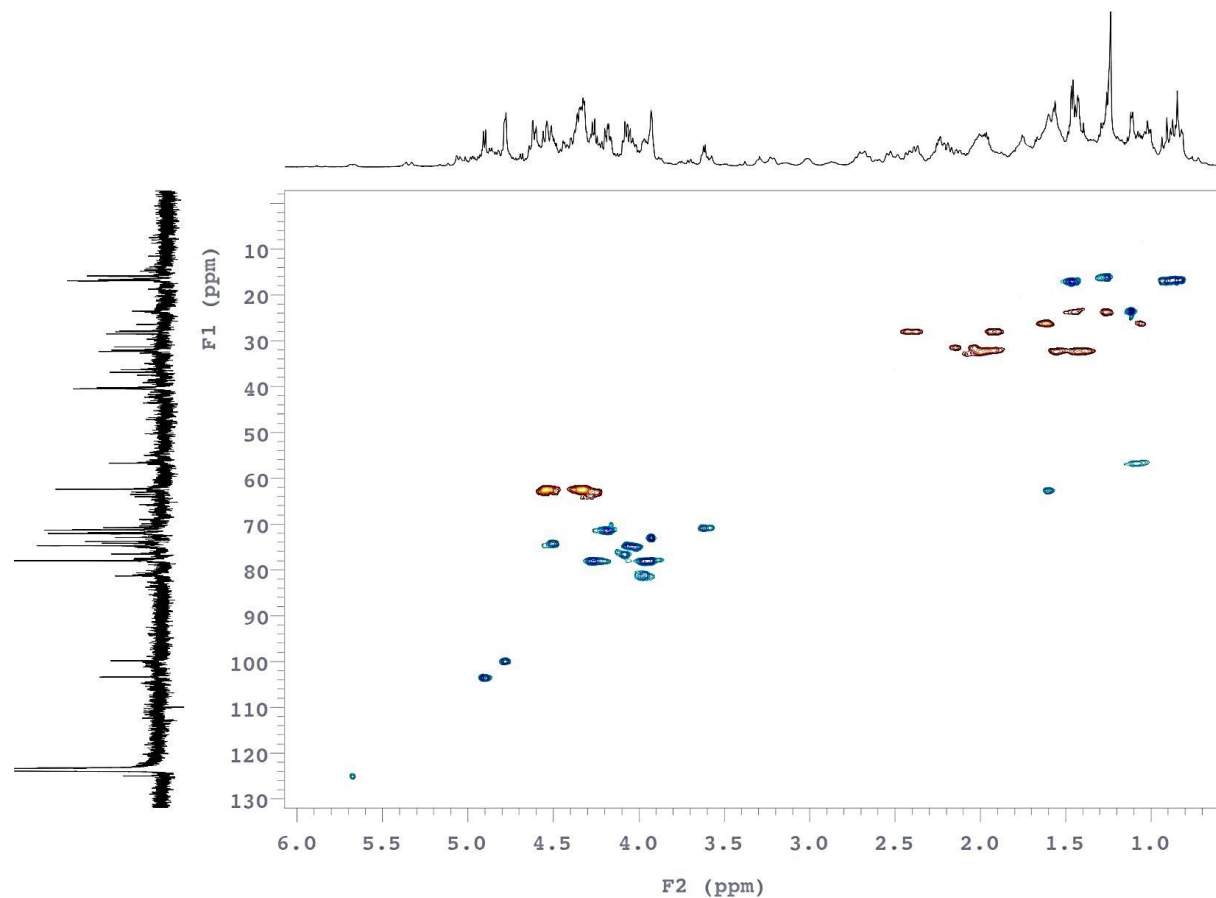
Kết quả phổ cộng hưởng từ hạt nhân proton ^1H cho biết độ dịch chuyển hóa học (ppm) và cường độ của tín hiệu thể hiện qua chiều cao của mỗi tín hiệu. Kết quả có 3 tín hiệu proton (trong khoảng từ 7,2-8,8 ppm) là của dung môi pyridine $\text{C}_5\text{D}_5\text{N}$, còn những tín hiệu khác là của các proton có trong hợp chất **1**.

PL10-2: Phổ cộng hưởng từ hạt nhân ^{13}C của chất **1** trong dung môi $\text{C}_5\text{D}_5\text{N}$ ở tần số 150 MHz



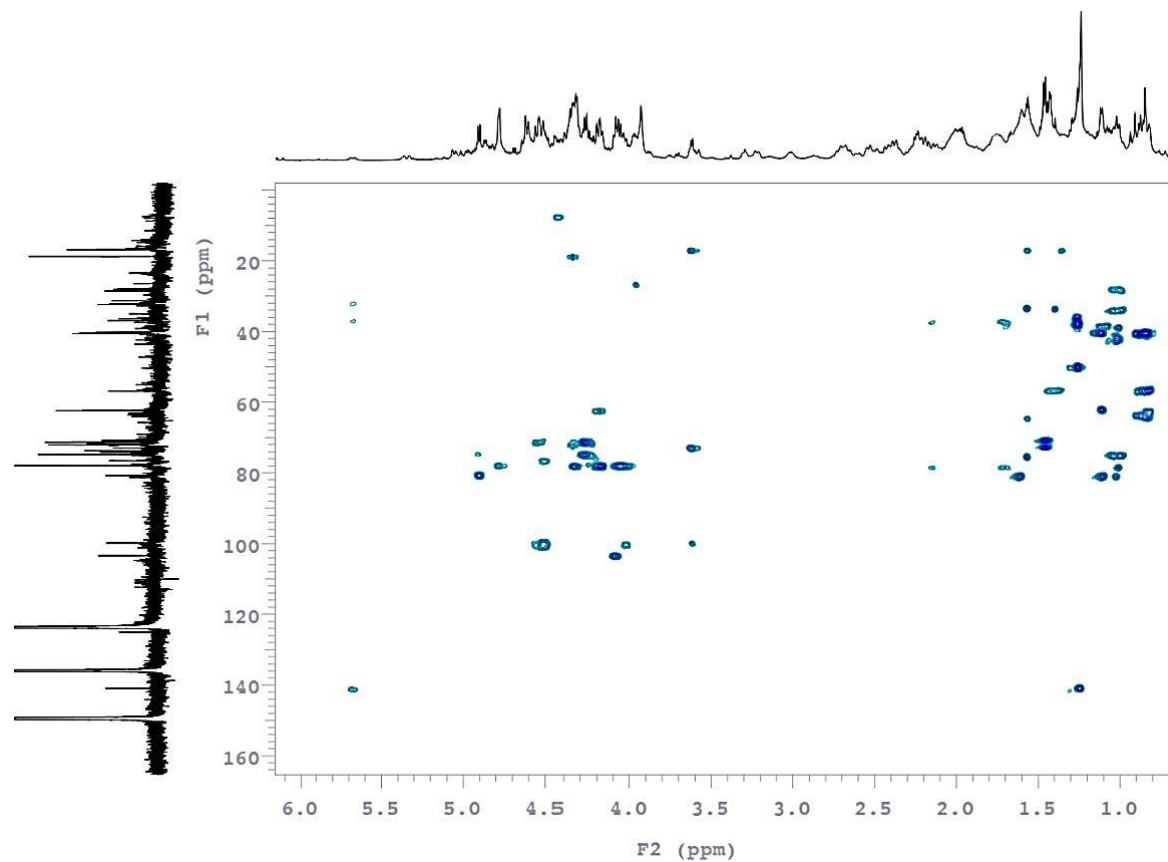
Kết quả cho thấy các tín hiệu của các nhóm carbon ankyl (R-CH_3) trong vùng từ 0–50 ppm; Nhóm CH, CH_2 từ 20–100 ppm; Anken (chứa nối đôi $\text{C}=\text{C}$) từ 100–150 ppm; carbon thơm (Ar-C) từ 110–160 ppm có trong hợp chất **1**. 3 tín hiệu mạnh nhất (trong khoảng từ 120–150 ppm) là của dung môi pyridine $\text{C}_5\text{D}_5\text{N}$.

PL10-3: Phổ tương quan dị hạt nhân qua 1 liên kết (HSQC) của chất **1** trong dung môi $\text{C}_5\text{D}_5\text{N}$ ở tần số 600 MHz



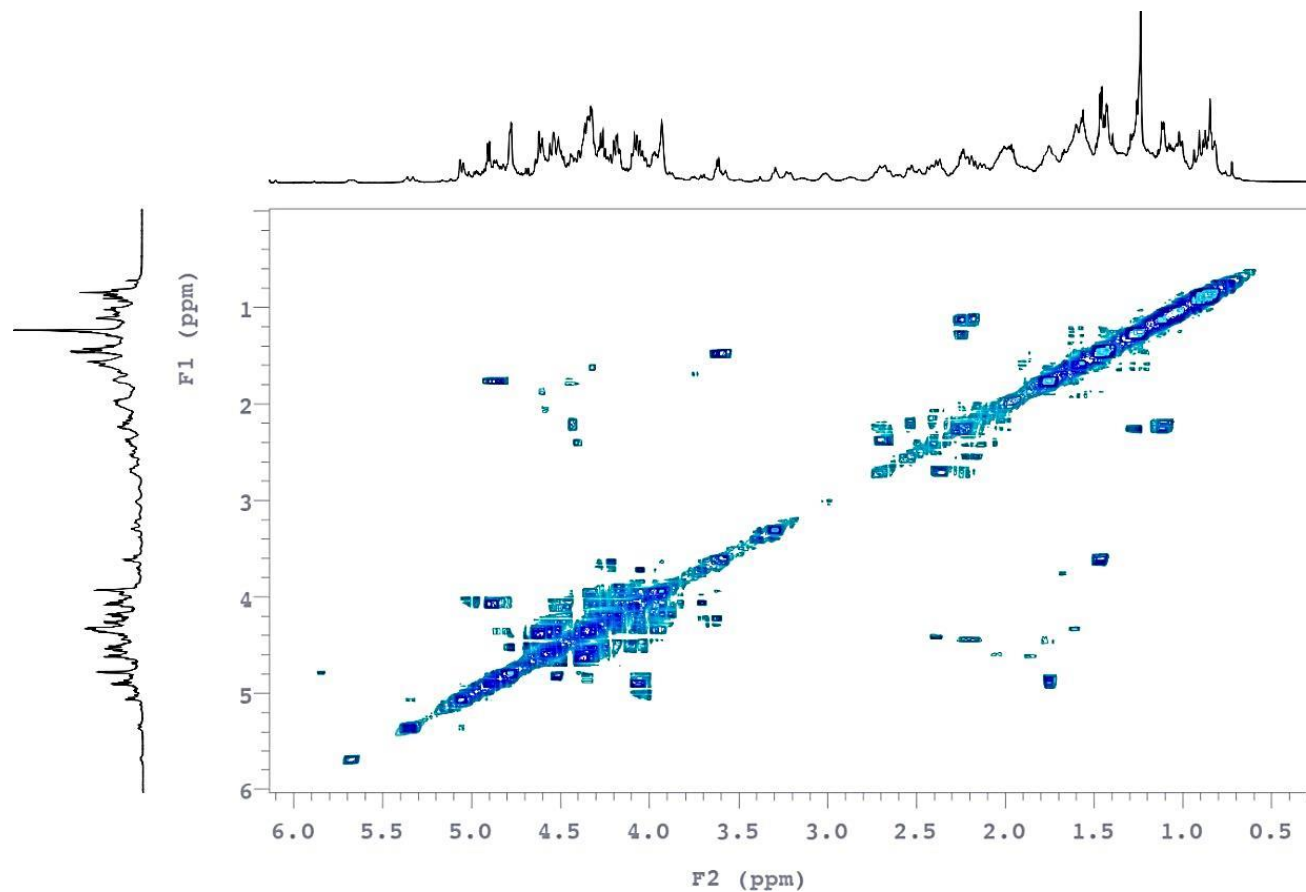
F1 (từ 0-130 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân của carbon (^{13}C); F2 (từ 1,0-6,0 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân của proton (^1H). Phổ khảo sát sự ghép cặp giữa 2 proton ^1H và ^{13}C , hai hạt nhân này nếu liên kết với nhau sẽ để lại tín hiệu trên phổ, đối với carbon bậc 4 không liên kết với proton sẽ không để lại tín hiệu.

PL10-4: Phổ tương quan dị hạt nhân qua nhiều liên kết (HMBC) của chất **1** trong dung môi C_5D_5N ở tần số 600 MHz



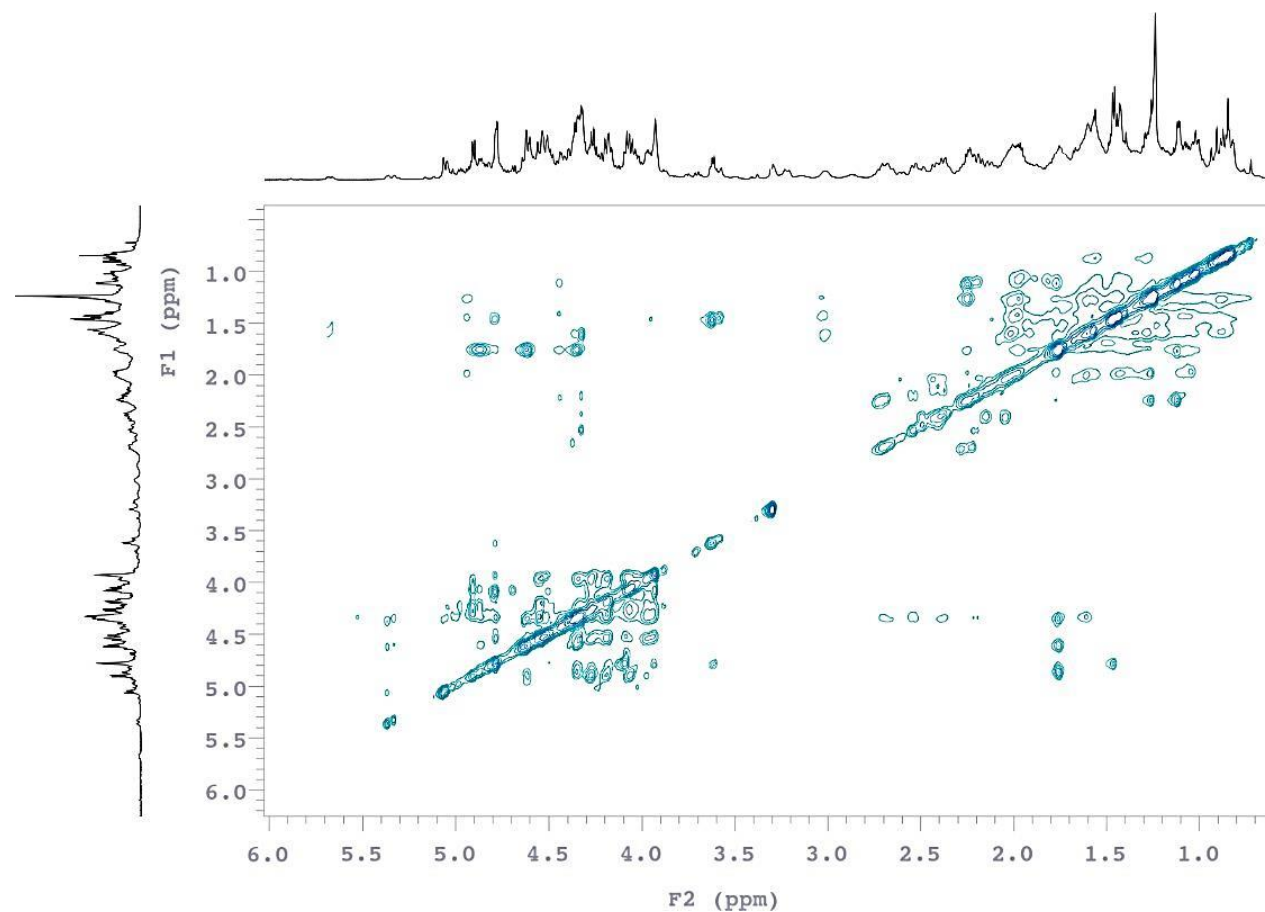
F1 (từ 0-130 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân của carbon (^{13}C); F2 (từ 1,0-6,0 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân của proton (1H). Phổ khảo sát sự ghép cặp giữa 2 proton 1H và ^{13}C qua 2, 3 liên kết (tương tác xa).

PL10-5: Phổ tương quan đồng hạt nhân H-H theo mạch liên kết (COSY) của chất **1** trong dung môi C_5D_5N ở tần số 600 MHz



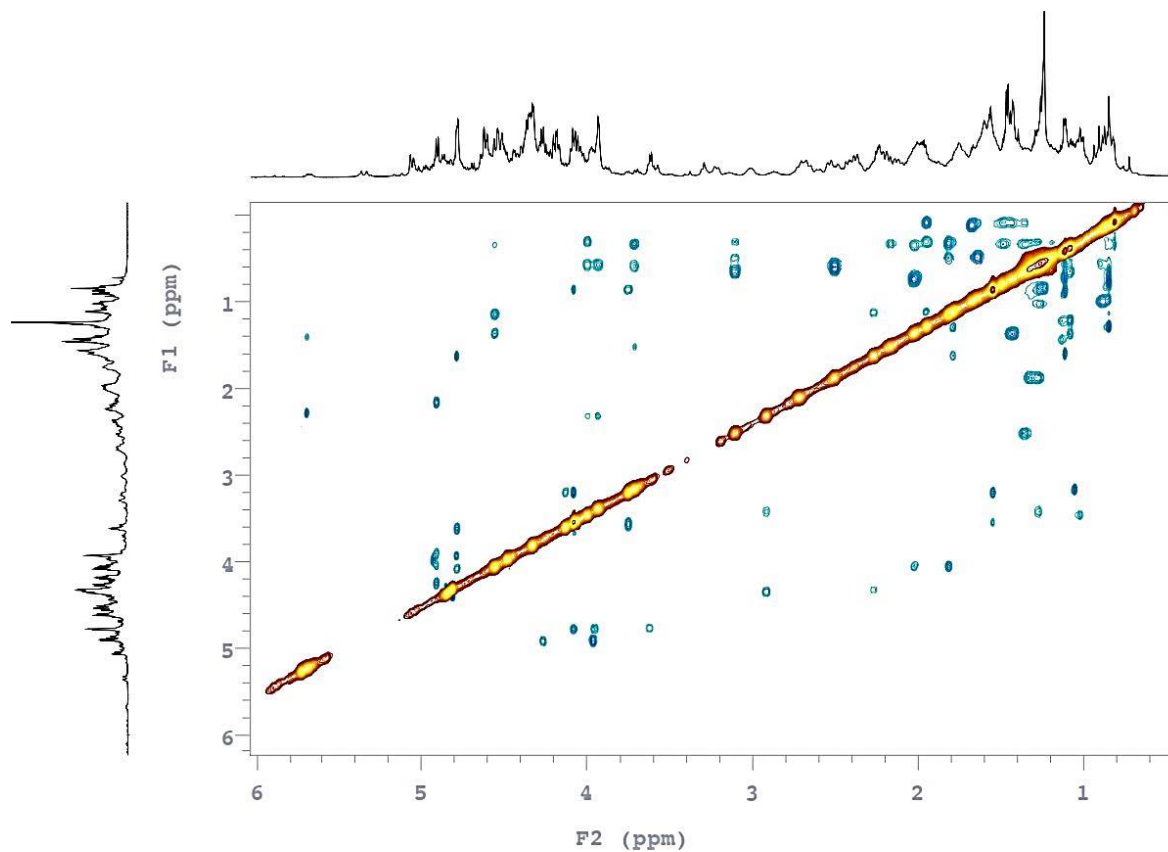
Phổ COSY là phổ tương quan giữa hai hạt nhân proton (1H), phổ cộng hưởng từ hạt nhân proton được biểu diễn trên 2 trục, F1 và F2 với độ dịch chuyển hóa học của các proton (1H) từ 0-6 ppm. Tín hiệu ở trên đường chéo là tín hiệu gốc, tín hiệu ngoài đường chéo cho biết sự ghép cặp của hai hạt nhân cùng loại.

L10-6: Phổ tương quan tổng thể (TOCSY) của chất **1** trong dung môi $\text{C}_5\text{D}_5\text{N}$ ở tần số 600 MHz



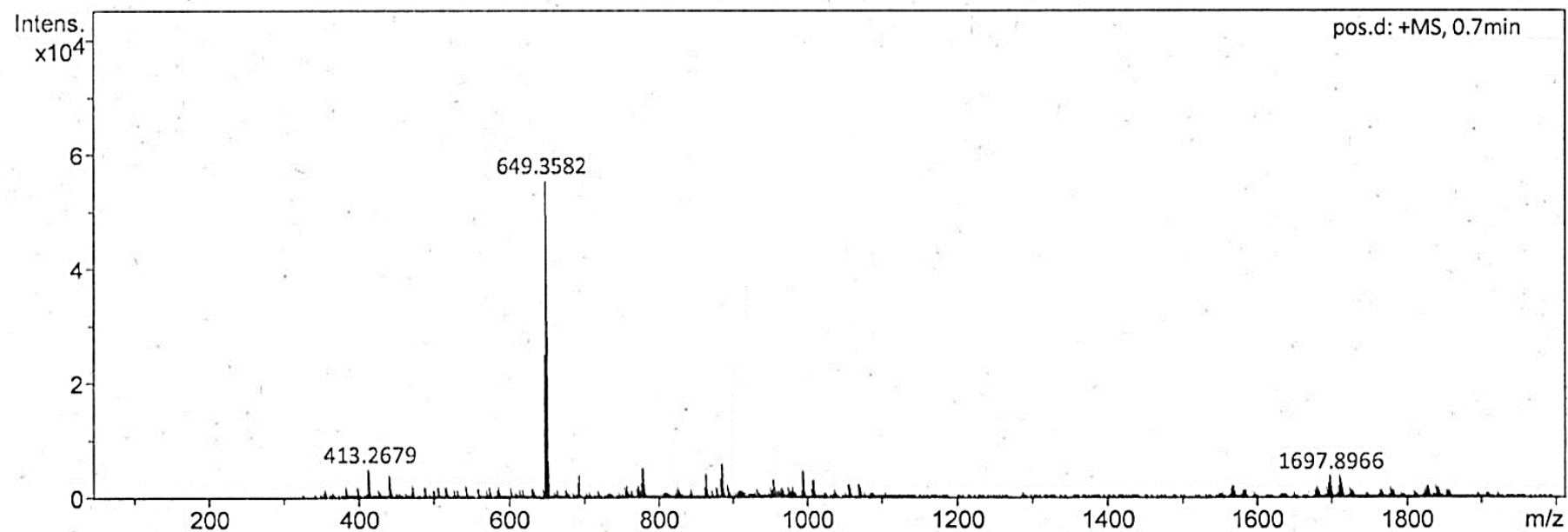
F1 và F2 (từ 0-6 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân proton (^1H). Phổ cho thấy mối tương quan giữa tất cả các proton có trong hợp chất **1**.

PL10-7: Phổ tăng cường Overhauser khung quay (ROESY) của chất **1** trong dung môi C_5D_5N ở tần số 600 MHz



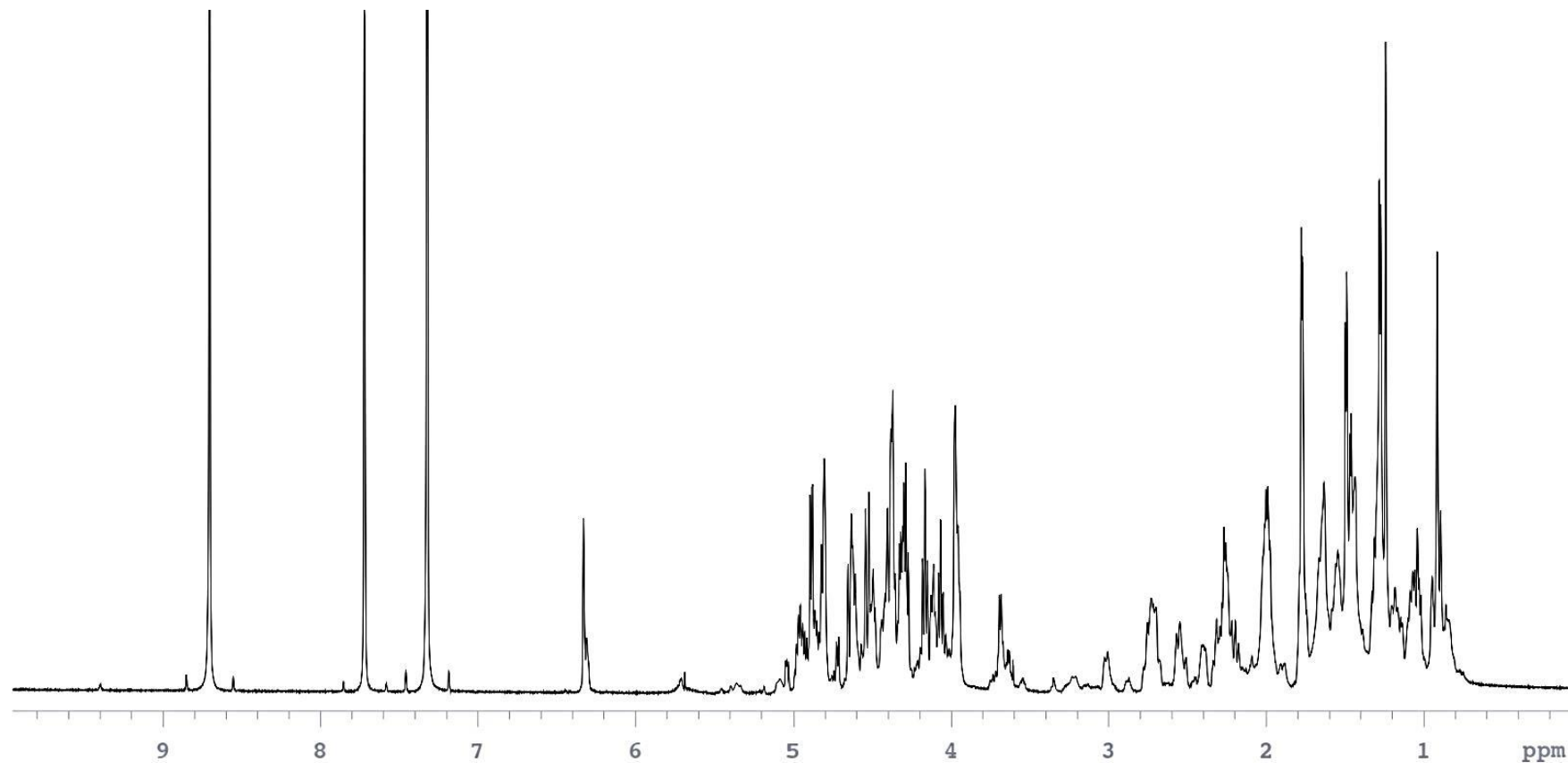
F1 và F2 (từ 0-6 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân proton (1H). Phổ cho thấy mối tương quan giữa các proton trong không gian. Thông qua sử dụng hiệu ứng overhauser để xác định mô hình cấu trúc của chất.

PL10-8: Giá trị phổ khối ion hóa phun điện phân giải cao (HR-ESI-MS) của chất 1



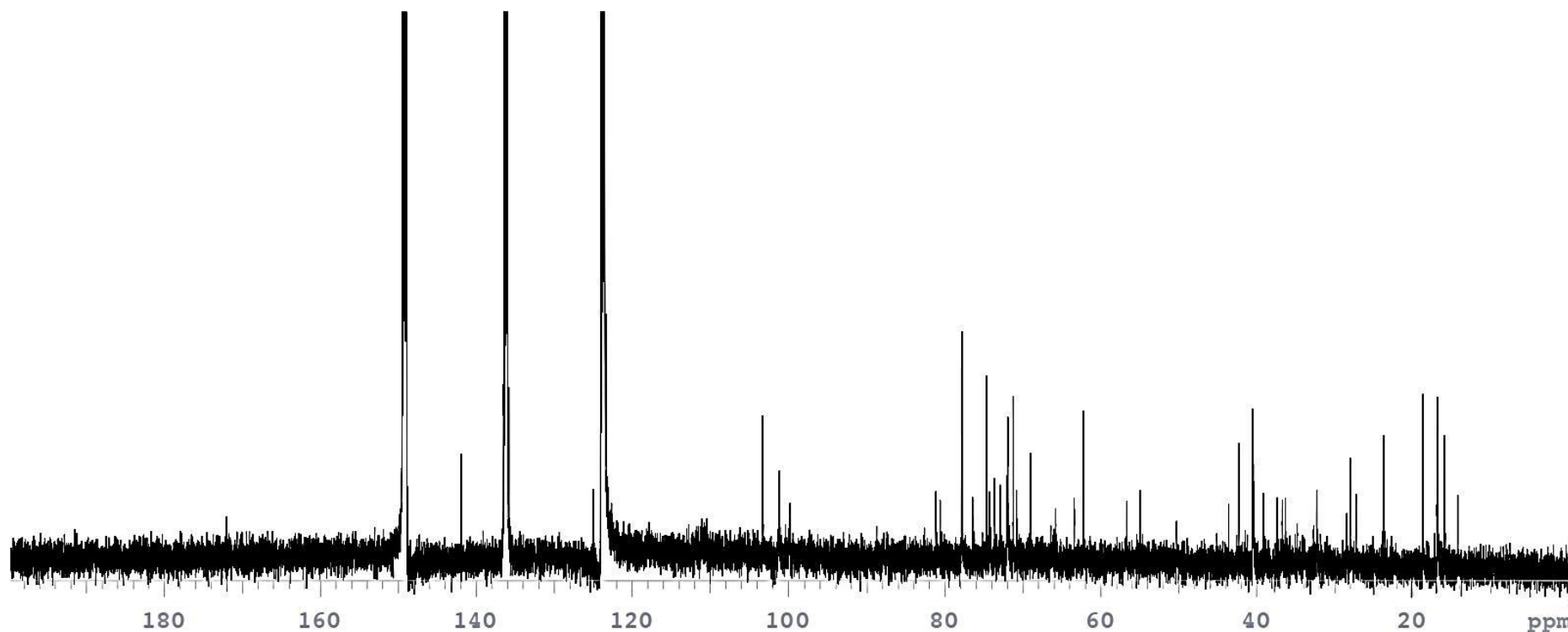
Trục hoành là tỷ số khối lượng phân tử/điện tích (m/z), trục tung là cường độ tín hiệu, kết quả phổ cho thấy tín hiệu có cường độ mạnh nhất (xấp xỉ 60.000) có giá trị m/z là 649,3582 là đỉnh cơ bản (ion mạnh nhất).

PL10-9: Phổ cộng hưởng từ hạt nhân ^1H của chất **2** trong dung môi $\text{C}_5\text{D}_5\text{N}$ ở tần số 150 MHz



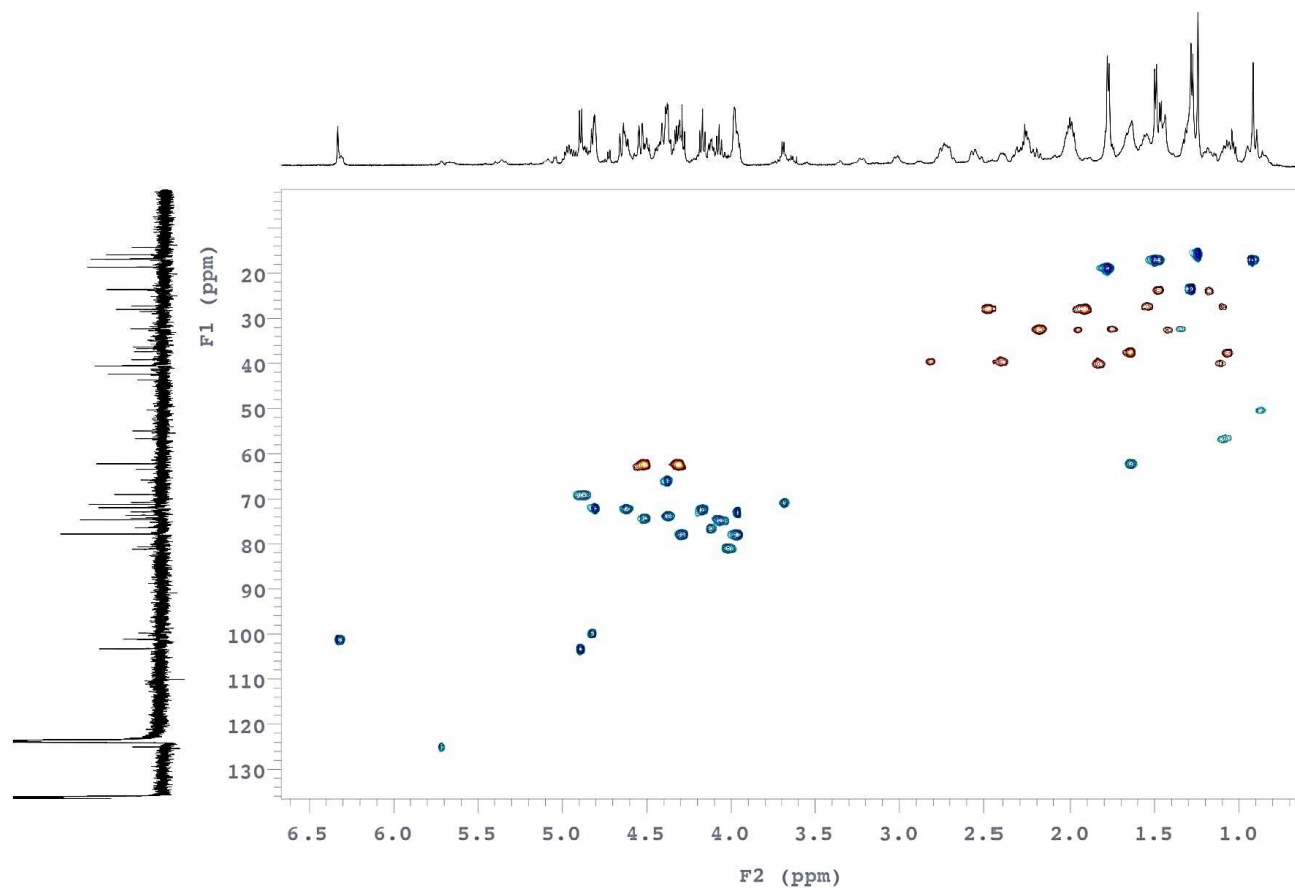
Kết quả phổ cộng hưởng từ hạt nhân proton ^1H cho biết độ dịch chuyển hóa học (ppm) và cường độ của tín hiệu thể hiện qua chiều cao của mỗi tín hiệu. Kết quả có 3 tín hiệu proton (trong khoảng từ 7,2-8,8 ppm) là của dung môi pyridine $\text{C}_5\text{D}_5\text{N}$, còn những tín hiệu khác là của các proton có trong hợp chất **2**.

PL10-10: Phổ cộng hưởng từ hạt nhân ^{13}C của chất **2** trong dung môi $\text{C}_5\text{D}_5\text{N}$ ở tần số 150 MHz



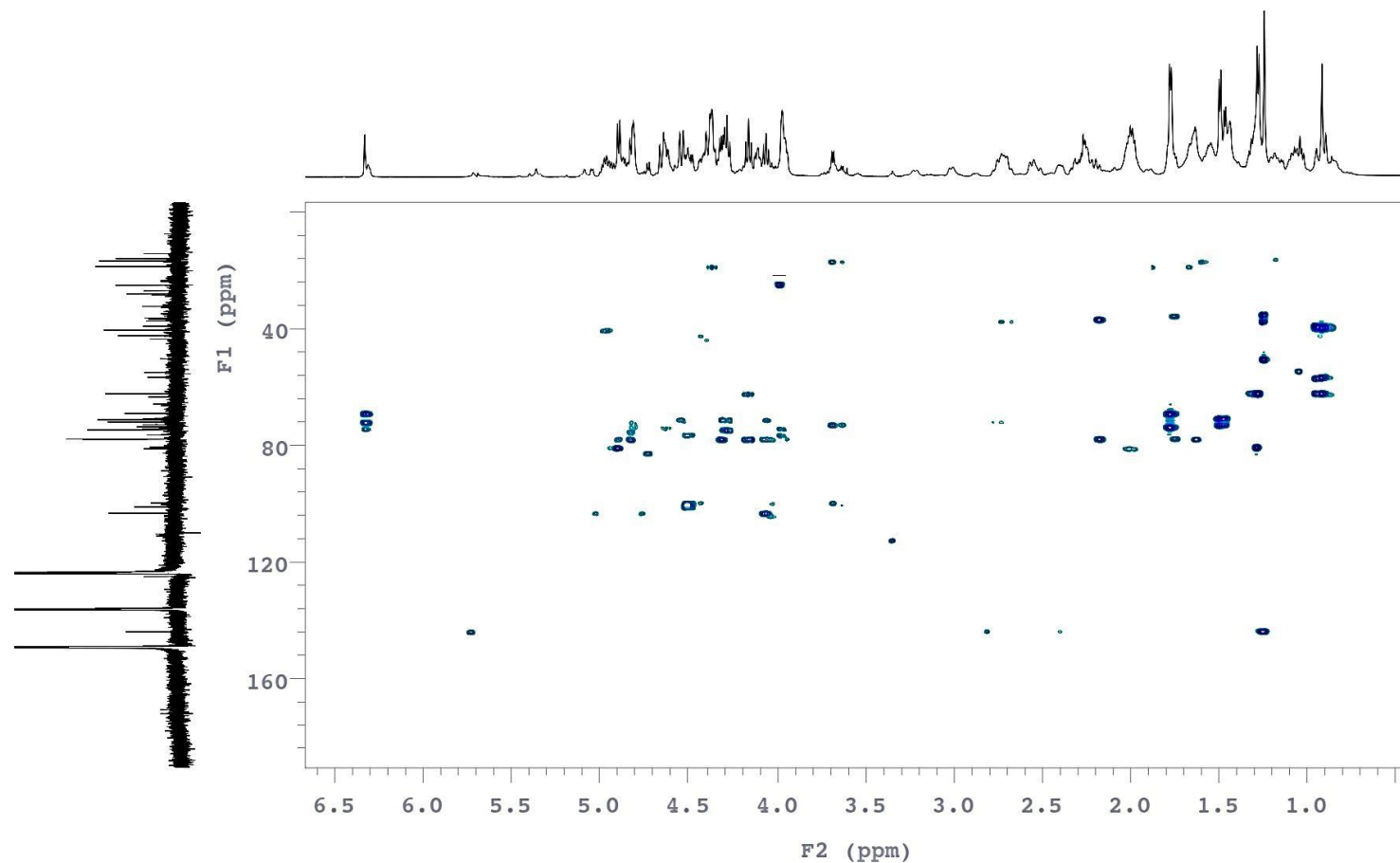
Kết quả cho thấy các tín hiệu của các nhóm carbon ankyl (R-CH_3) trong vùng từ 0–50 ppm; Nhóm CH , CH_2 từ 20–100 ppm; Anken (chứa nối đôi $\text{C}=\text{C}$) từ 100–150 ppm; carbon thơm (Ar-C) từ 110–160 ppm có trong hợp chất **2**. 3 tín hiệu mạnh nhất (trong khoảng từ 120–150 ppm) là của dung môi pyridine $\text{C}_5\text{D}_5\text{N}$.

PL10-11: Phổ tương quan dị hạt nhân qua 1 liên kết (HSQC) của chất **2** trong dung môi C_5D_5N ở tần số 150 MHz



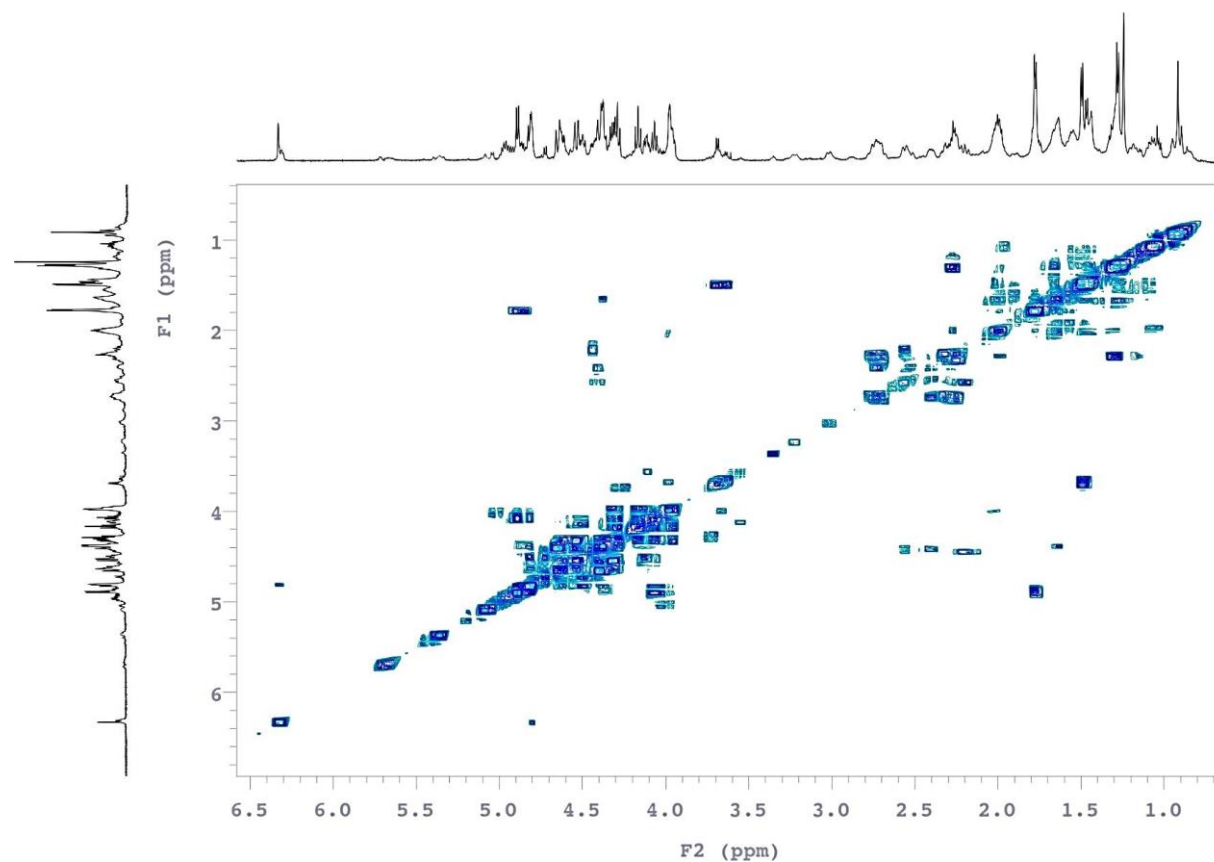
F1 (từ 0-130 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân của carbon (^{13}C); F2 (từ 1,0-6,0 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân của proton (1H). Phổ khảo sát sự ghép cặp giữa 2 proton 1H và ^{13}C , hai hạt nhân này nếu liên kết với nhau sẽ để lại tín hiệu trên phổ, đối với carbon bậc 4 không liên kết với proton sẽ không để lại tín hiệu.

PL10-12: Phổ tương quan dị hạt nhân qua nhiều liên kết (HMBC) của chất **2** trong dung môi C_5D_5N ở tần số 150 MHz



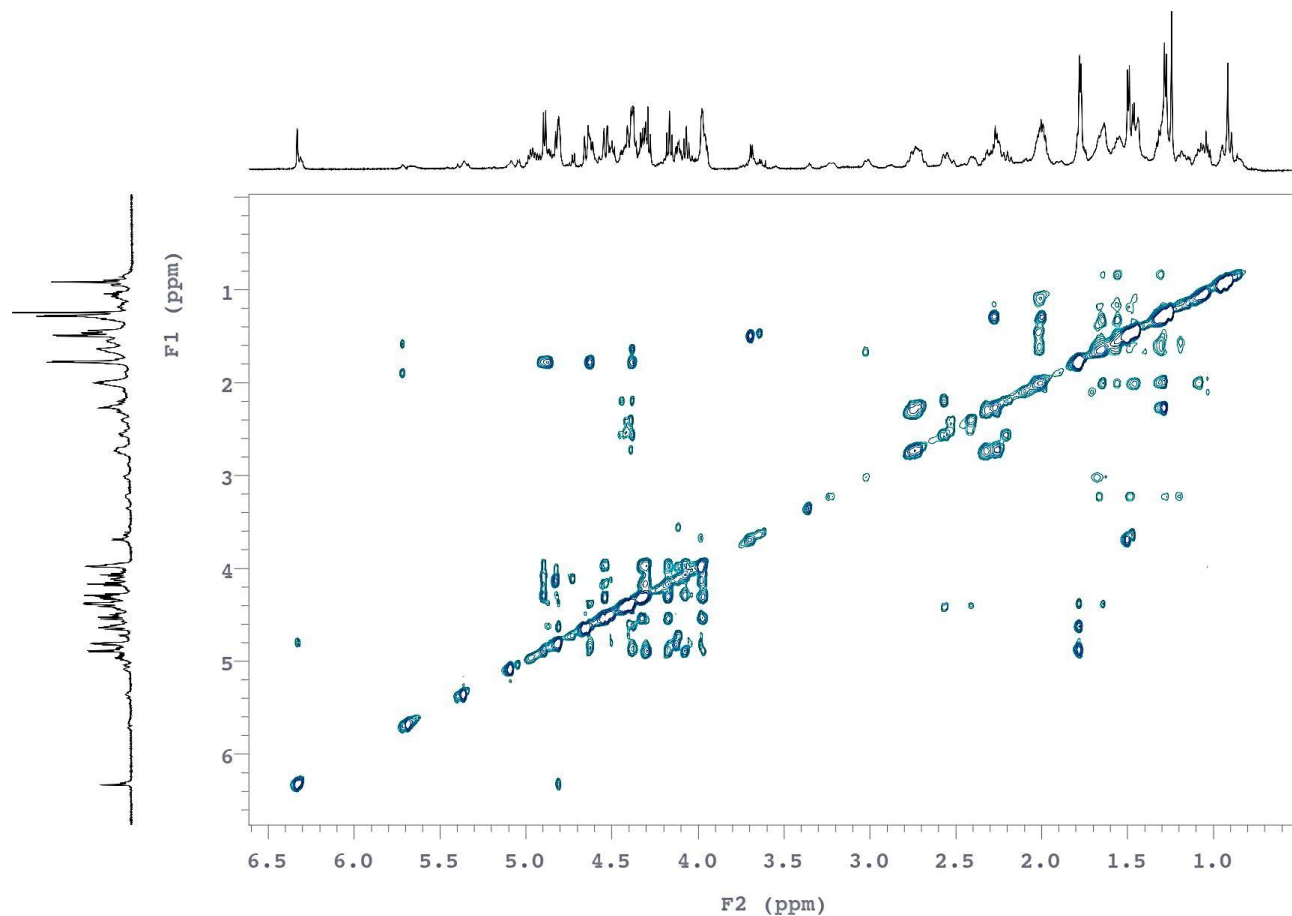
F1 (từ 0-130 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân của carbon (^{13}C); F2 (từ 1,0-6,0 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân của proton (1H). Phổ khảo sát sự ghép cặp giữa 2 proton 1H và ^{13}C qua 2, 3 liên kết (tương tác xa).

PL10-13: Phổ tương quan đồng hạt nhân H-H theo mạch liên kết (COSY) của chất **2** trong dung môi C_5D_5N ở tần số 600 MHz



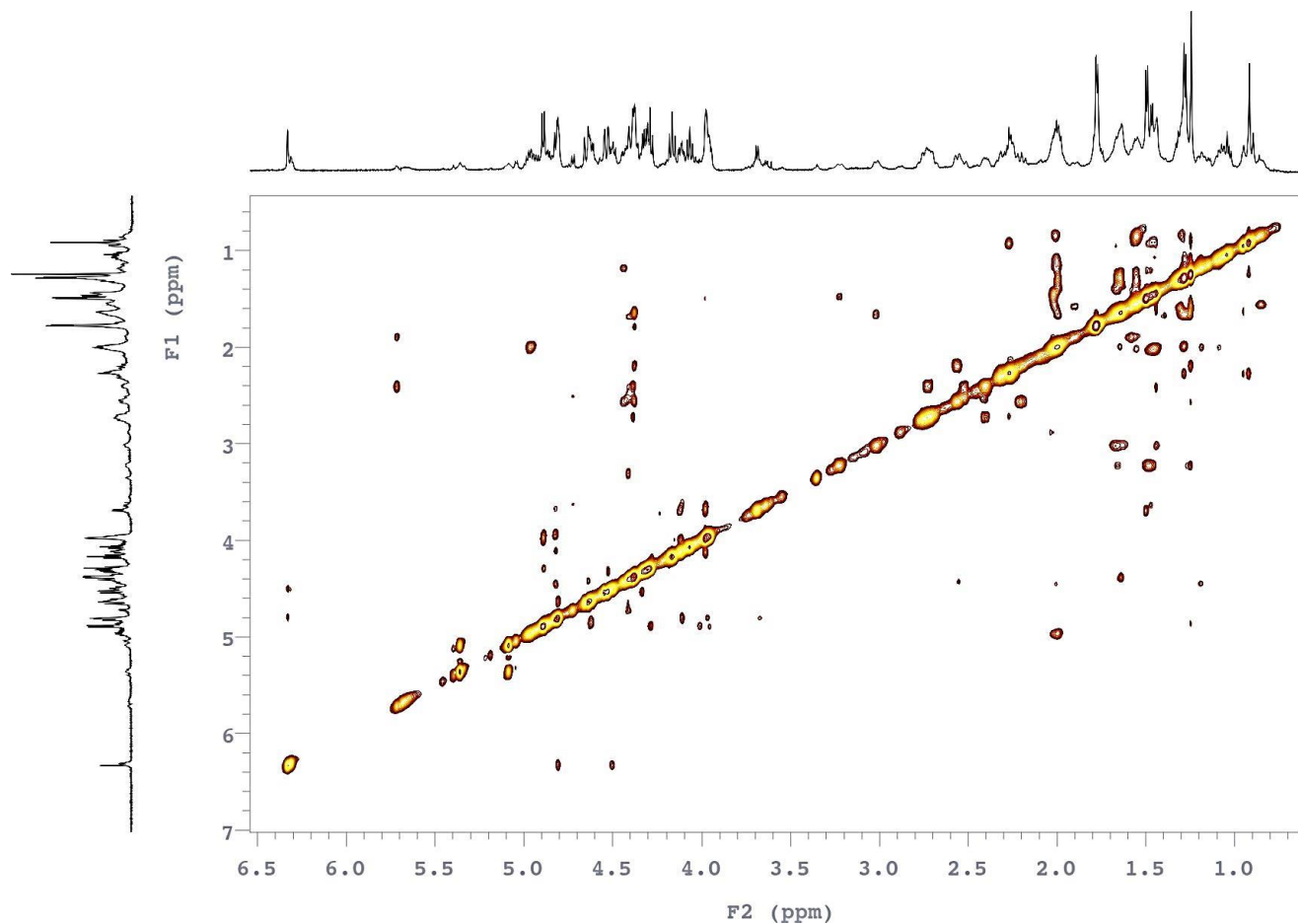
Phổ COSY là phổ tương quan giữa hai hạt nhân proton (1H), phổ cộng hưởng từ hạt nhân proton được biểu diễn trên 2 trục, F1 và F2 với độ dịch chuyển hóa học của các proton (1H) từ 0-6 ppm. Tín hiệu ở trên đường chéo là tín hiệu gốc, tín hiệu ngoài đường chéo cho biết sự ghép cặp của hai hạt nhân cùng loại.

PL10-14: Phổ tương quan tổng thể (TOCSY) của chất **2** trong dung môi $\text{C}_5\text{D}_5\text{N}$ ở tần số 600 MHz



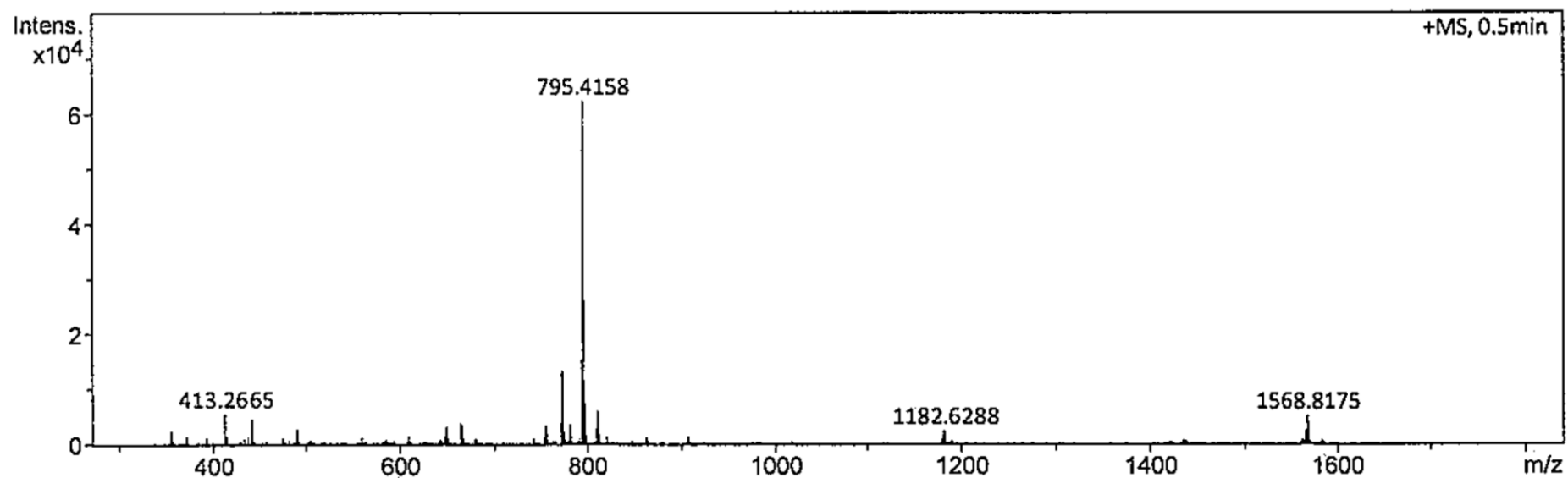
F1 và F2 (từ 0-6 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân proton (^1H). Phổ cho thấy mối tương quan giữa tất cả các proton có trong hợp chất **1**.

PL10-15: Phổ tăng cường Overhauser khung quay (ROESY) của chất **2** trong dung môi C_5D_5N ở tần số 600 MHz



F1 và F2 (từ 0-6 ppm) là độ dịch chuyển hóa học phổ cộng hưởng từ hạt nhân proton (1H). Phổ cho thấy mối tương quan giữa các proton trong không gian. Thông qua sử dụng hiệu ứng overhauser để xác định mô hình cấu trúc của chất.

PL10-16: Giá trị phổ khối ion hóa phun điện phân giải cao (HR-ESI-MS) của chất 2



Trục hoành là tỷ số khối lượng phân tử/điện tích (m/z), trục tung là cường độ tín hiệu, kết quả phổ cho thấy tín hiệu có cường độ mạnh nhất (xấp xỉ 65.000) có giá trị m/z là 795,4158 là đỉnh cơ bản (ion mạnh nhất).

PL10-17a: Kết quả tìm kiếm (Scifinder) cấu trúc chất 1

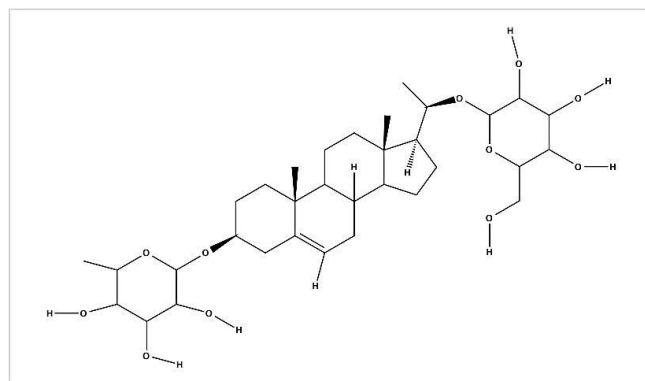
Initiating Search

September 11, 2023, 10:32AM

 Substances:

Filtered By:

Stereochemistry: **Absolute Stereo Match**



Structure Match: **Substructure**

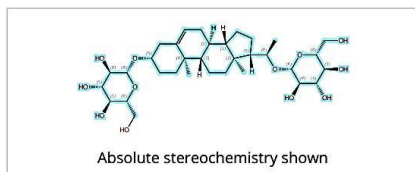
Search Tasks

Task	Search Type	View
Exported: Returned Substance Results + Filters (6)	 Substances	View Results

PL10-17b: Kết quả tìm kiếm (Scifinder) chất giống chất 1

1

113679-33-7



$C_{33}H_{54}O_{12}$

β -D-Glucopyranoside, (3 β ,20 R)-pregn-5-ene-3, 20-diyl bis-



2

References



2

Reactions



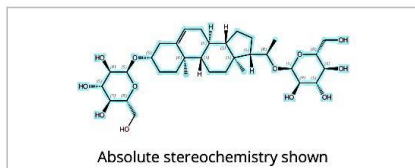
0

Suppliers

Key Physical Properties	Value	Condition
Molecular Weight	642.78	-
Boiling Point (Predicted)	813.3±65.0 °C	Press: 760 Torr
Density (Predicted)	1.39±0.1 g/cm ³	Temp: 20 °C; Press: 760 Torr
pKa (Predicted)	12.91±0.70	Most Acidic Temp: 25 °C

2

123367-47-5



$C_{33}H_{54}O_{12}$

α -D-Glucopyranoside, (3 β ,20 R)-pregn-5-ene-3, 20-diyl bis-



1

Reference



2

Reactions



0

Suppliers

Key Physical Properties	Value	Condition
Molecular Weight	642.78	-
Boiling Point (Predicted)	813.3±65.0 °C	Press: 760 Torr
Density (Predicted)	1.39±0.1 g/cm ³	Temp: 20 °C; Press: 760 Torr
pKa (Predicted)	12.91±0.70	Most Acidic Temp: 25 °C

PL10-18a: Kết quả tìm kiếm (Scifinder) cấu trúc chất 2

Initiating Search

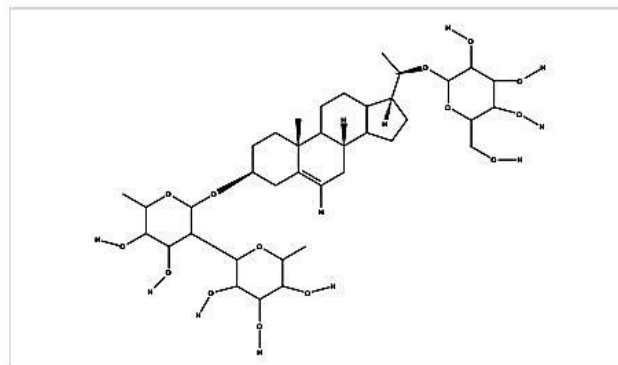
September 11, 2023, 3:25PM

Substances:

Filtered By:

Similarity: 95-98

Number of Components: 1



Structure Match: Similarity

Search Tasks

Task	Search Type	View
Exported: Returned Substance Results + Filters (22)	Substances	View Results

Copyright © 2023 American Chemical Society (ACS). All Rights Reserved.

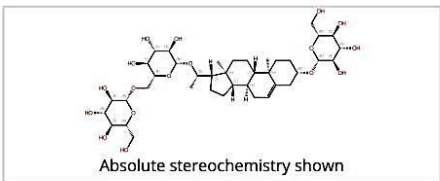
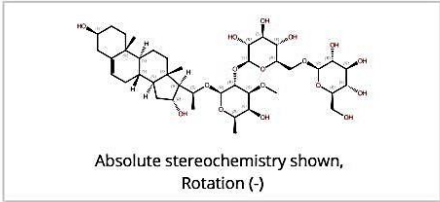
Internal use only. Redistribution is subject to the terms of your SciFinder[®] License Agreement and CAS information Use Policies.

PL10-18b: Kết quả tìm kiếm (Scifinder) chất giống chất 2



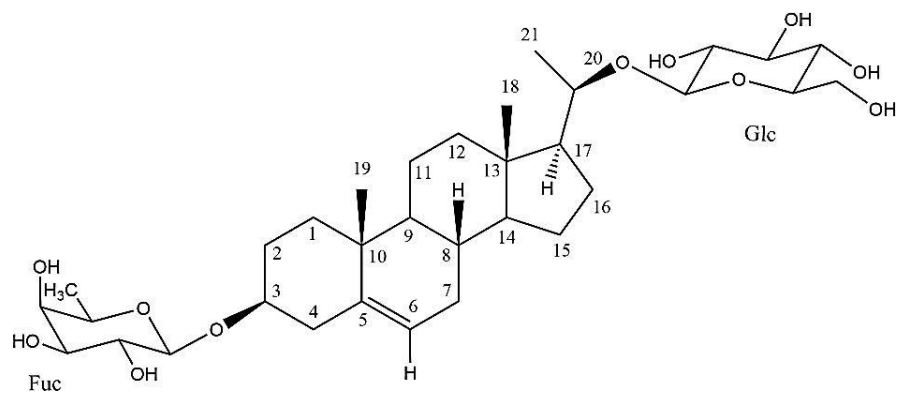
Substances (22)

[View in SciFinderⁿ](#)

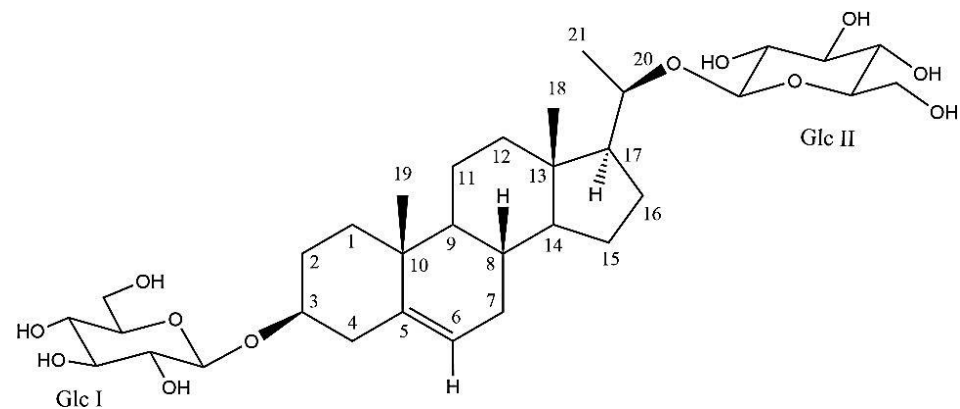
1		Similarity Score: 97															
521970-07-0  Absolute stereochemistry shown C₃₉H₆₄O₁₇ (3β,20S)-3-(β-D-Glucopyranosyloxy)pregn-5-en-20-yl 6-O-β-D-glucopyranosyl-β-D-glucopyranoside 5 References 0 Reactions 2 Suppliers		<table> <thead> <tr> <th>Key Physical Properties</th><th>Value</th><th>Condition</th></tr> </thead> <tbody> <tr> <td>Molecular Weight</td><td>804.92</td><td>-</td></tr> <tr> <td>Boiling Point (Predicted)</td><td>969.8±65.0 °C</td><td>Press: 760 Torr</td></tr> <tr> <td>Density (Predicted)</td><td>1.48±0.1 g/cm³</td><td>Temp: 20 °C; Press: 760 Torr</td></tr> <tr> <td>pKa (Predicted)</td><td>12.89±0.70</td><td>Most Acidic Temp: 25 °C</td></tr> </tbody> </table> Experimental Properties Spectra	Key Physical Properties	Value	Condition	Molecular Weight	804.92	-	Boiling Point (Predicted)	969.8±65.0 °C	Press: 760 Torr	Density (Predicted)	1.48±0.1 g/cm ³	Temp: 20 °C; Press: 760 Torr	pKa (Predicted)	12.89±0.70	Most Acidic Temp: 25 °C
Key Physical Properties	Value	Condition															
Molecular Weight	804.92	-															
Boiling Point (Predicted)	969.8±65.0 °C	Press: 760 Torr															
Density (Predicted)	1.48±0.1 g/cm ³	Temp: 20 °C; Press: 760 Torr															
pKa (Predicted)	12.89±0.70	Most Acidic Temp: 25 °C															
2		Similarity Score: 96															
73532-66-8  Absolute stereochemistry shown, Rotation (-) C₄₀H₆₆O₁₇ (3β,16α,20S)-3,16-Dihydroxypregn-5-en-20-yl O-β-D-glucopyranosyl-(1→6)-O-β-D-glucopyranosyl-(1→2)-6-deoxy-3-O-methyl-β-D-galactopyranoside 4 References 2 Reactions 0 Suppliers		<table> <thead> <tr> <th>Key Physical Properties</th><th>Value</th><th>Condition</th></tr> </thead> <tbody> <tr> <td>Molecular Weight</td><td>818.94</td><td>-</td></tr> </tbody> </table>	Key Physical Properties	Value	Condition	Molecular Weight	818.94	-									
Key Physical Properties	Value	Condition															
Molecular Weight	818.94	-															

PL10-20a: Dữ liệu so sánh công hưởng từ hạt nhân của parasiticoside A (1) và periseoside C

parasiticoside A (1)			periseoside C		
Position	δ_C	δ_H	Position	δ_C	δ_H
1	37.5	1.06 m, 1.62 m	1	37.5	0.95, 1.66
2	31.3	1.70 m, 2.14 m	2	30.2	1.70, 2.09 br d (11.2)
3	78.1	3.93 ov	3	78.1	3.91
4	39.3	2.28 m, 2.72 m	4	39.3	2.43 (br t, 12.8), 2.68 ddd (13.0, 4.5, 2.3)
5	140.9	-	5	140.9	-
6	125.0	5.67 br s	6	121.9	5.29 (m)
7	32.1	1.40 m, 1.95 m	7	32.1	1.44 (m), 1.81
8	31.8	1.30 ov	8	31.8	1.33
9	50.3	0.80 m	9	50.3	0.80
10	36.9	-	10	36.9	-
11	21.0	1.27 m, 1.45 m	11	21.0	1.32, 1.37
12	39.1	1.00 m, 1.82 m	12	39.1	1.00, 1.82
13	40.5	-	13	41.5	-
14	56.7	1.08 ov	14	56.7	0.86
15	28.0	1.91 m, 2.40 m	15	27.2	1.89 m, 2.33 m
16	26.4	1.05 m, 1.53 m	16	24.4	1.05, 1.53
17	62.4	1.60 ov	17	58.3	1.57 m
18	16.9	0.85 s	18	12.4	0.64 s
19	16.0	1.28 s	19	19.4	0.88 s
20	81.3	3.98 ov	20	81.3	3.86 m
21	23.5	1.12 d (6.0)	21	23.3	1.52 d (6.0)
Fuc-1	99.8	4.78 d (7.6)	Glc I-1	102.6	5.02 d (7.7)
2	74.2	4.52 dd (9.4, 7.6)	2	75.4	4.03 pt (7.9)
3	76.6	4.08 dd (9.4, 2.9)	3	78.3	4.27 pt (8.5)
4	72.9	3.93 br d (2.9)	4	71.7	4.23 pt (8.5)
5	70.8	3.62 br q (6.4)	5	78.6	3.95
6	17.0	1.46 d (6.4)	6	62.9	4.39 dd (11.6, 3.0),
Glc-1	103.4	4.90 d (7.6)	Glc II-1	106.2	4.54 br d (11.6) 4.93 d (7.7)
2	74.7	4.05 ov	2	75.8	3.98
3	78.0	4.26 ov	3	78.5	4.25 pt (8.5)
4	72.2	4.18 ov	4	71.8	4.21 pt (8.5)
5	78.1	3.96 m	5	78.7	3.96
6	62.4	4.33 m, 4.54 m	6	63.0	4.38 dd (11.8, 3.1), 4.53 br d, (11.8)



1

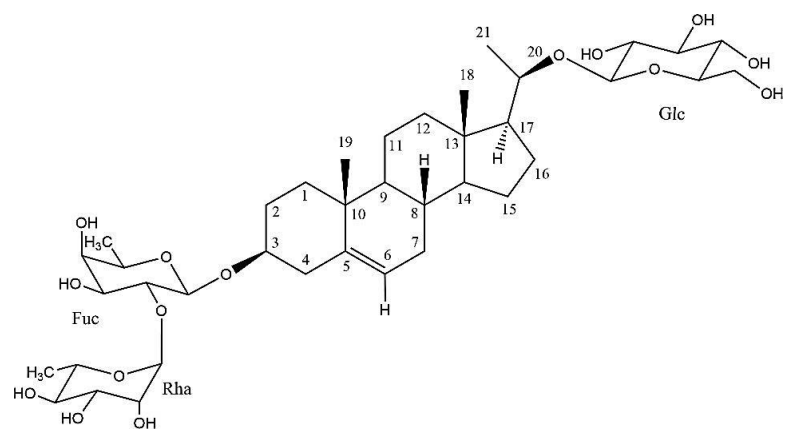


Periseoside C

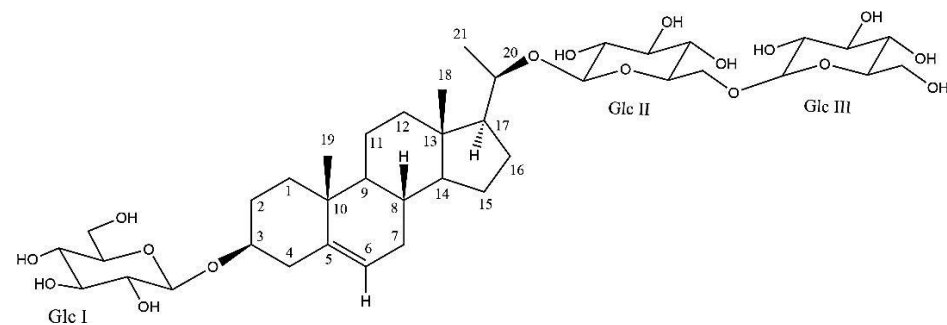
Cấu trúc của chất **1** dựa trên dữ liệu phân tích và so sánh (Chất có đặc điểm tương đồng là periseoside C)

PL10-20b: Dữ liệu so sánh công hưởng từ hạt nhân của parasiticoside B (2) và biondianoside D

Parasiticoside B (2)			Biondianoside D		
Position	δ_C	δ_H	Position	δ_C	δ_H
1	37.4	1.08 m, 1.66 m	1	37.5	
2	32.3	1.75 m, 2.18 m	2	30.2	
3	77.9	3.97 ov	3	78.1	
4	39.2	2.40 m, 2.82 m	4	39.3	
5	140.9	-	5	140.9	5.32 br s
6	125.0	5.72 br s	6	121.9	
7	32.7	1.43 m, 1.97 m	7	32.1	
8	32.2	1.35 ov	8	31.8	
9	50.3	0.87 m	9	50.3	
10	36.3	-	10	36.9	
11	23.9	1.18 m, 1.48 m	11	21.0	
12	39.3	1.11 m, 1.83 m	12	39.0	
13	40.6	-	13	41.5	
14	56.7	1.09 ov	14	56.6	
15	28.0	1.92 m, 2.48 m	15	27.3	
16	26.8	1.09 m, 1.54 m	16	24.4	
17	62.3	1.64 ov	17	58.3	
18	16.9	0.91 s	18	12.4	0.64 s
19	15.9	1.26 s	19	19.4	0.90 s
20	81.2	4.01 ov	20	81.2	3.93 m
21	23.6	1.28 d (6.0)	21	23.6	1.62 d (6.0)
Fuc-1	99.9	4.80 d (7.6)	Glc I-1	105.3	5.03 d (7.6)
2	74.4	4.53 dd (9.4, 7.6)	2	75.5	
3	76.4	4.12 dd (9.4, 2.9)	3	78.4	
4	73.1	3.96 br d (2.9)	4	71.7	
5	70.9	3.68 br q (6.4)	5	77.3	
6	16.8	1.49 d (6.4)	6	70.1	
Glc-1	103.5	4.88 d (7.6)	Glc II-1	106.0	4.89 d (7.6)
2	74.6	4.06 ov	2	75.3	
3	77.9	4.28 ov	3	78.4	
4	72.3	4.18 ov	4	71.7	
5	78.1	3.96 m	5	78.5	
6	62.3	4.30 m, 4.52 m	6	62.8	
Rha-1	101.2	6.35 br s	Glc III-1	102.6	5.19 d (7.7)
2	72.1	4.79 br s	2	75.2	
3	72.2	4.61 dd (9.4, 3.5)	3	78.5	
4	73.8	4.35 dd (9.9, 9.4)	4	71.7	
5	69.0	4.83 dq (9.9, 5.9)	5	78.6	
6	18.7	1.78 d (5.9)	6	62.8	



2



Biondianoside D

Cấu trúc của chất 2 dựa trên dữ liệu phân tích và so sánh (Chất có đặc điểm tương đồng là Biondianoside D)