

**BỘ GIÁO DỤC  
VÀ ĐÀO TẠO**

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

**HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ**

---



**Lê Phương Chi**

**NGHIÊN CỨU HOẠT TÍNH KHÁNG SINH, GÂY ĐỘC TẾ  
BÀO UNG THƯ VÀ ĐẶC ĐIỂM DI TRUYỀN HỆ GEN CỦA  
XẠ KHUẨN *Streptomyces* sp. VCCM 22255**

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

**Hà Nội – 2023**

**BỘ GIÁO DỤC  
VÀ ĐÀO TẠO**

**HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ**

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



**Lê Phương Chi**

**NGHIÊN CỨU HOẠT TÍNH KHÁNG SINH, GÂY ĐỘC TẾ  
BÀO UNG THƯ VÀ ĐẶC ĐIỂM DI TRUYỀN HỆ GEN CỦA  
XẠ KHUẨN *Streptomyces* sp. VCCM 22255**

Chuyên ngành: Sinh học thực nghiệm

Mã số: 8420114

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

NGƯỜI HƯỚNG DẪN KHOA HỌC:

1. PGS. TS. Phí Quyết Tiên
2. TS. Quách Ngọc Tùng

**Hà Nội – 2023**

## LỜI CAM ĐOAN

Tôi xin cam đoan đề tài nghiên cứu: “**Nghiên cứu hoạt tính kháng sinh, gây độc tế bào ung thư và đặc điểm di truyền hệ gen của xạ khuẩn *Streptomyces* sp. VCCM 22255**” là công trình nghiên cứu của tôi dựa trên những tài liệu, số liệu do chính tôi tự tìm hiểu và nghiên cứu. Chính vì vậy, các kết quả nghiên cứu đảm bảo trung thực, khách quan nhất và chưa được công bố trong công trình nghiên cứu khác. Mọi thông tin nội dung tham khảo trong luận văn đều được trích dẫn rõ ràng tên tác giả, tên công trình, thời gian, địa điểm và nguồn gốc. Các số liệu, kết quả nêu trong luận văn là trung thực nếu sai tôi hoàn chịu trách nhiệm trước pháp luật.

*Hà Nội, ngày 14 tháng 04 năm 2023*

Tác giả luận văn

Lê Phương Chi

## LỜI CẢM ƠN

Lời đầu tiên, tôi xin bày tỏ lòng biết ơn sâu sắc tới hai thầy hướng dẫn là PGS.TS Phí Quyết Tiên và TS. Quách Ngọc Tùng (Trung tâm Giống và Bảo tồn nguồn gen Vi sinh vật, Viện Công nghệ sinh học) bởi những kiến thức quý báu và sự tận tình trực tiếp hướng dẫn trong suốt thời gian thực hiện luận văn.

Tôi xin gửi lời cảm ơn chân thành tới Trung tâm Giống và Bảo tồn nguồn gen Vi sinh vật và các cán bộ thuộc Trung tâm đã tạo điều kiện thuận lợi nhất để tôi hoàn thành nghiên cứu này. Đặc biệt, tôi xin bày tỏ lòng biết ơn tới TS. Vũ Thị Hạnh Nguyên, ThS. Nguyễn Thị Thu An, ThS. Bùi Thị Liên đã hỗ trợ nhiệt tình và truyền đạt nhiều kinh nghiệm thực tiễn quý báu.

Luận văn được thực hiện nhờ sự hỗ trợ kinh phí của Đề tài: “Giải trình tự hệ gen và xác định đặc tính di truyền liên quan đến hoạt tính kháng sinh, gây độc tế bào ung thư của xạ khuẩn nội sinh trên cây màng tang (*Litsea cubeba* (Lour.) Pers.)” được cấp bởi Viện Hàn lâm Khoa học và Công nghệ Việt Nam với mã số ĐLTE00.03/21-22. Tôi cũng xin trân trọng cảm ơn nguồn động viên to lớn từ Chương trình học bổng thạc sĩ, tiến sĩ trong nước của Quỹ Đổi mới sáng tạo Vingroup (VINIF), Viện Nghiên cứu Dữ liệu lớn, mã số VINIF.2021.ThS.90.

Thứ tư, tôi xin chân thành cảm ơn Ban Lãnh đạo, phòng Đào tạo, các phòng chức năng của Học viện Khoa học và Công nghệ, đặc biệt là các thầy, cô giảng viên tại Khoa Công nghệ sinh học đã truyền đạt kiến thức trong suốt quá trình tôi học tập bậc cao học này.

Cuối cùng, tôi xin gửi lời cảm ơn trang trọng nhất tới gia đình và bạn bè đã ủng hộ và khích lệ tôi hết mình, là chỗ dựa tinh thần vững chắc của tôi trong suốt thời gian qua.

Hà Nội, ngày 14 tháng 04 năm 2023

Học viên

Lê Phương Chi

## DANH MỤC CHỮ VIẾT TẮT

<b>Chữ viết tắt</b>	<b>Tên đầy đủ</b>
26-L5	Ung thư biểu mô ruột kết
A427	Ung thư biểu mô tuyến
A549	Ung thư phổi ở người
AntiSMASH	antibiotics & Secondary Metabolite Analysis Shell
bp	Base pair
BT20	Ung thư vú
BV-BRC	Cơ sở dữ liệu Bacterial and Viral Bioinformatics Resource Center
COG	Nhóm các cụm gen orthologous (Cluster of orthologous group)
CRISPR	Các đoạn gen ngắn lặp lại xen kẽ (Clustered Regularly Interspaced Short Palindromic Repeats)
DNA	Deoxyribonucleic acid
DNA	Deoxyribonucleic acid
GO	Gene Ontology
HepG2	Ung thư gan ở người
HMEC	Biểu mô tuyến vú của người
IC <sub>50</sub>	Nồng độ ức chế 50%
KB	Ung thư biểu mô ở người
KEGG	Cơ sở dữ liệu Kyoto Encyclopedia of Genes and Genomes
LLC	Ung thư biểu mô phổi Lewis
MCF7	Ung thư vú ở người
MIC	Nồng độ ức chế tối thiểu (Minimum Inhibitory Concentration)
NCBI	Trung tâm Thông tin Công nghệ Sinh học Quốc gia (National Center for Biotechnology Information)
NGS	Giải trình tự thế hệ mới (Next-generation sequencing)
NRPS	Peptide non-ribosomal
OEC-M1	Tế bào ung thư miệng ở người
PCR	Phản ứng chuỗi polyme (Polymerase chain reaction)

PGAP	Prokaryotic Genomes Annotation Pipeline
rRNA	Ribosomal ribonucleic acid
SK-LU-1	Ung thư biểu mô tuyến phổi của con người
smBGC	Cụm gen sinh tổng hợp các chất chuyển hóa thứ cấp (Secondary metabolite biosynthetic gene clusters)
T1PKS	Polyketide dạng 1 (type 1 polyketide synthase)
T2PKS	Polyketide dạng 2 (Type 2 polyketide synthase)
T3PKS	Polyketide dạng 3 (Type 3 polyketide synthase)
TIR	Trình tự lặp đảo ở đoạn kết thúc (Terminal inverted repeats sequence)
VSV	Vi sinh vật
WGS	Giải trình tự hệ gen (Whole genome sequencing)
XKNS	Xạ khuẩn nội sinh

**DANH MỤC BẢNG**

Bảng 3.1. Hoạt tính gây độc tế bào ung thư của 2 chủng xạ khuẩn nội sinh với 3 loại tế bào ung thư thử nghiệm. ....	27
Bảng 3.2. Dữ liệu thô từ mẫu giải trình tự <i>S. cacaoi</i> VCCM 22255. ....	32
Bảng 3.3. So sánh đặc điểm hệ gen của <i>S. cacaoi</i> VCCM 22255 với các hệ gen <i>Streptomyces</i> khác. ....	34
Bảng 3.4. Các cụm gen sinh tổng hợp được dự đoán trong hệ gen <i>S. cacaoi</i> VCCM 22255. ....	40

## DANH MỤC HÌNH

Hình 1.1. Giải trình tự hệ gen bằng công nghệ Illumina.....	10
Hình 1.2. Bản đồ hệ gen chủng <i>Streptomyces xiamenensis</i> 318.....	13
Hình 1.3. (a) Cụm gen sinh tổng hợp desferrioxamine ở <i>Streptomyces coelicolor</i> . (b) Con đường sinh tổng hợp tiền chất HAC và HSC. (c) Sinh tổng hợp desferrioxamine E và B. ....	15
Hình 1.4. Cây màng tang ( <i>Litsea cubeba</i> Lour. Pers). ....	16
Hình 3.1. Sàng lọc hoạt tính kháng khuẩn của các chủng xạ khuẩn nội sinh trên cây màng tang. ....	25
Hình 3.2. Điện di đồ sản phẩm PCR khuếch đại gen 16S rRNA của chủng VCCM 22255 trên gel agarose 1,0%.....	29
Hình 3.3. Cây phát sinh chủng loại dựa trên phân tích trình tự gen 16S rRNA của chủng VCCM 22255 và các chủng xạ khuẩn tham chiếu đại diện. ....	30
Hình 3.4. Điện di đồ DNA tổng số của chủng <i>S. cacaoi</i> VCCM 22255.....	31
Hình 3.5. Bản đồ hệ gen chủng <i>S. cacaoi</i> VCCM 22255.....	33
Hình 3.6. Dự đoán các nhóm chức năng gen trong hệ gen <i>S. cacaoi</i> VCCM 22255.....	35
Hình 3.7. Tỷ lệ các gen chức năng của chủng <i>S. cacaoi</i> VCCM 22255 dựa trên cơ sở dữ liệu EggNOG.....	36
Hình 3.8. Số lượng các nhóm gen CAZyme của chủng <i>S. cacaoi</i> VCCM 22255.....	38
Hình 3.9. Các cụm gen sinh tổng hợp các chất chống ung thư được xác định trong hệ gen của <i>S. cacaoi</i> VCCM 22255. ....	44
Hình 3.10. Các cụm gen sinh tổng hợp các chất kháng khuẩn được xác định trong hệ gen của <i>S. cacaoi</i> VCCM 22255. ....	46
Hình 3.11. Các cụm gen sinh tổng hợp các chất kháng nấm được xác định trong hệ gen của <i>S. cacaoi</i> VCCM 22255. ....	48



## MỤC LỤC

LỜI CAM ĐOAN .....	i
LỜI CẢM ƠN .....	ii
DANH MỤC CHỮ VIẾT TẮT .....	iii
DANH MỤC BẢNG .....	v
DANH MỤC HÌNH.....	vi
MỤC LỤC .....	vii
MỞ ĐẦU .....	1
Chương 1. TỔNG QUAN TÀI LIỆU .....	3
1.1. TỔNG QUAN VỀ XẠ KHUẨN NỘI SINH TRÊN CÂY DƯỢC LIỆU... 3	
1.1.1. Định nghĩa xạ khuẩn và xạ khuẩn nội sinh..... 3	
1.1.2. Phân loại xạ khuẩn nội sinh dựa trên phân tích trình tự gen 16S rRNA 4	
1.1.3. Xạ khuẩn <i>Streptomyces</i> nội sinh trên cây dược liệu..... 5	
1.2. KHẢ NĂNG SINH TỔNG HỢP CÁC CHẤT CÓ HOẠT TÍNH Y DƯỢC TỪ XẠ KHUẨN <i>STREPTOMYCES</i> NỘI SINH..... 7	
1.2.1. Các chất kháng vi sinh vật .....	7
1.2.2. Các chất gây độc tế bào ung thư..... 8	
1.3. GIẢI MÃ HỆ GEN XẠ KHUẨN BẰNG CÔNG NGHỆ GIẢI TRÌNH TỰ HỆ GEN THỂ HỆ MỚI .....	9
1.3.1. Các công nghệ giải trình tự hệ gen được áp dụng hiện nay .....	9
1.3.2. Đặc điểm hệ gen xạ khuẩn thuộc chi <i>Streptomyces</i> .....	11
1.3.3. Các cụm gen sinh tổng hợp sản phẩm thứ cấp từ xạ khuẩn <i>Streptomyces</i> .....	13
1.4. CÂY MÀNG TANG VÀ TIỀM NĂNG KHAI THÁC .....	16
1.5. TÌNH HÌNH NGHIÊN CỨU XẠ KHUẨN NỘI SINH TRÊN CÂY DƯỢC LIỆU TẠI VIỆT NAM..... 17	
Chương 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU..... 19	
2.1. ĐỐI TƯỢNG .....	19
2.1.1. Chủng giống..... 19	
2.1.2. Hóa chất và thiết bị .....	19
2.1.3. Môi trường .....	20
2.2. PHƯƠNG PHÁP NGHIÊN CỨU .....	20
2.2.1. Đánh giá hoạt tính kháng vi sinh vật kiểm định của xạ khuẩn..... 20	
2.2.2. Đánh giá khả năng gây độc tế bào ung thư..... 21	

2.2.3. Phân loại xạ khuẩn dựa trên phân tích trình tự gen 16S rRNA .....	22
2.2.4. Giải trình tự và phân tích hệ gen của chủng xạ khuẩn tuyển chọn.....	23
2.2.5. Xử lý thông kê số liệu.....	24
Chương 3. KẾT QUẢ VÀ THẢO LUẬN.....	25
3.1. KHẢ NĂNG KHÁNG KHUẨN CỦA XẠ KHUẨN NỘI SINH TRÊN CÂY MÀNG TANG.....	25
3.2. ĐÁNH GIÁ KHẢ NĂNG GÂY ĐỘC TẾ BÀO UNG THƯ CỦA CÁC CHỦNG XẠ KHUẨN TIỀM NĂNG .....	27
3.3. PHÂN LOẠI CHỦNG VCCM 22255 DỰA TRÊN PHÂN TÍCH TRÌNH TỰ GEN 16S rRNA.....	28
3.4. GIẢI TRÌNH TỰ HỆ GEN VÀ PHÂN TÍCH MỐI LIÊN QUAN GIỮA ĐẶC TÍNH DI TRUYỀN VÀ ĐẶC TÍNH KHÁNG SINH, GÂY ĐỘC TẾ BÀO CỦA CHỦNG VCCM 22255 .....	31
3.4.1. Giải trình tự hệ gen chủng <i>S. cacaoi</i> VCCM 22255.....	31
3.4.2. Lắp ghép <i>de novo</i> và chú giải hệ gen VCCM 22255.....	32
3.4.3. Phân tích đặc điểm di truyền hệ gen VCCM 22255.....	34
3.4.4. Các cụm gen sinh tổng hợp sản phẩm bậc hai của chủng <i>S. cacaoi</i> VCCM 22255.....	39
KẾT LUẬN VÀ KIẾN NGHỊ .....	49
TÀI LIỆU THAM KHẢO .....	50
PHỤ LỤC.....	62

## MỞ ĐẦU

Ngày nay, thực trạng kháng kháng sinh và sự gia tăng của bệnh ung thư đang trở thành vấn đề y tế toàn cầu, đặc biệt ở các nước đang phát triển. Tại Việt Nam, vi khuẩn đa kháng thuốc kháng sinh đã xuất hiện do lạm dụng kháng sinh, khiến vấn đề kháng thuốc kháng sinh đang ngày càng nghiêm trọng. Theo thống kê của Tổ chức ung thư toàn cầu GLOBOCAN trong năm 2020, ước tính có 182.563 ca mắc mới và 122.690 ca tử vong do ung thư ở Việt Nam. Vì vậy, việc tìm ra các loại thuốc điều trị bệnh truyền nhiễm và bệnh ung thư đang là cuộc chạy đua không ngừng giữa các nhà khoa học và các hãng dược phẩm trên thế giới.

Xạ khuẩn nội sinh là nhóm vi sinh vật cư trú bên trong các mô tế bào thực vật mà không cạnh tranh dinh dưỡng hay gây hại tới cây chủ. Các nghiên cứu gần đây đã chứng minh các chủng xạ khuẩn *Streptomyces* nội sinh trên cây dược liệu sinh ra các hợp chất không những có tính đa dạng về mặt số lượng mà còn hoạt tính sinh học như hoạt tính kháng vi khuẩn gây bệnh, gây độc tế bào ung thư, chống oxy hóa và kiểm soát sinh học. Đặc biệt, tỷ lệ phát hiện các hoạt chất mới từ xạ khuẩn nội sinh trên thực vật cao hơn nhiều lần so với xạ khuẩn biến và đất. Tuy nhiên, các nghiên cứu giải mã cơ chế sinh tổng hợp các chất có hoạt tính sinh học từ xạ khuẩn ở mức độ kiểu gen vẫn còn hạn chế.

Công nghệ giải trình tự gen thế hệ mới (Next Generation Sequencing, NGS) hiện đang là chủ đề nóng và thu hút sự quan tâm của các nhà vi sinh vật học trên thế giới. Giải trình tự hệ gen bằng NGS được gọi là giải trình tự thông lượng cao cho phép đồng thời giải mã hàng triệu trình tự DNA cùng lúc, nhờ vậy, nâng cao hiệu suất của quá trình giải mã bộ gen sinh vật. Xạ khuẩn *Streptomyces* chứa trung bình khoảng 8-83 cụm gen sinh tổng hợp các chất chuyển hóa thứ cấp nhưng chỉ có 2-3 chất được xác định trong môi trường lên men. Các cụm gen tiềm năng sinh tổng hợp chất mới thường ở dạng ẩn và chỉ tích lũy một lượng nhỏ trong tế bào. Công nghệ giải trình tự hệ gen (Whole genome sequencing, WGS) giúp sàng lọc, giải mã các cụm gen mới và xa hơn tạo tiền đề phát triển công nghệ chỉnh sửa gen nhằm khai thác các hoạt chất mới. Tại Việt Nam, xạ khuẩn sinh kháng sinh đã được khai thác từ lâu, tuy nhiên chưa có nhiều nghiên cứu về hệ gen xạ khuẩn nội sinh trên cây dược liệu được công bố. Với các triển vọng nghiên cứu trên, đề tài

**“Nghiên cứu hoạt tính kháng sinh, gây độc tế bào ung thư và đặc điểm di truyền hệ gen của xạ khuẩn *Streptomyces* sp. VCCM 22255”** được thực hiện nhằm cung cấp thông tin hệ gen và tìm ra mối liên hệ di truyền hệ gen liên quan đến sinh tổng hợp các chất kháng khuẩn và gây độc tế bào ung thư mới.

***Mục tiêu của đề tài:***

Tuyển chọn và đánh giá hoạt tính kháng khuẩn, gây độc tế bào ung thư và xác định được mối liên hệ di truyền hệ gen liên quan đến các hoạt tính y dược của chủng xạ khuẩn *Streptomyces* sp. VCCM 22255.

***Nội dung nghiên cứu:***

- (1) Tuyển chọn và đánh giá hoạt tính kháng khuẩn và gây độc tế bào ung thư từ chủng xạ khuẩn nội sinh.
- (2) Định danh chủng xạ khuẩn VCCM 22255 bằng giải trình tự và phân tích gen 16S rRNA.
- (3) Giải trình tự, chú giải và phân tích các gen/cụm gen tham gia sinh tổng hợp chất kháng sinh, gây độc tế bào ung thư.

## Chương 1. TỔNG QUAN TÀI LIỆU

### 1.1. TỔNG QUAN VỀ XẠ KHUẨN NỘI SINH TRÊN CÂY DƯỢC LIỆU

#### 1.1.1. Định nghĩa xạ khuẩn và xạ khuẩn nội sinh

Xạ khuẩn (Streptomycetaceae) là nhóm vi khuẩn có cấu tạo tế bào dạng sợi phân nhánh như nấm và hình thành bào tử. Thuật ngữ "Actinomycetes" có nguồn gốc từ tiếng Hy Lạp với *aktis* (tia) và *mykes* (nấm), bởi có các đặc điểm của cả vi khuẩn và nấm. Tế bào xạ khuẩn phát triển sâu vào môi trường dinh dưỡng nhằm hấp thụ chất dinh dưỡng được gọi là khuẩn ty cơ chất [1]. Khuẩn ty cơ chất này còn có thể sinh sắc tố làm thay màu môi trường xung quanh, tùy thuộc vào loài xạ khuẩn. Trái ngược, khuẩn ty khí sinh thiết lập hệ sợi cấu thành bề mặt của xạ khuẩn và sinh bào tử. Khác biệt với các nhóm vi khuẩn khác, đặc trưng của xạ khuẩn là có tỷ lệ GC cao, đạt trên 55%. Xạ khuẩn được tìm thấy ở nhiều loại môi trường khác nhau như đất, nước, thực vật, động vật và con người. Xạ khuẩn có nhiều đặc tính đa dạng và được nghiên cứu chuyên sâu vì khả năng sản xuất các enzym ngoại bào và đặc biệt là các chất chuyển hóa thứ cấp mang lại nhiều giá trị ứng dụng trong dược phẩm, y tế và các lĩnh vực khác [2,3]. Cụ thể, đến nay đã có hơn 10.000 chất kháng sinh, chiếm khoảng 73% tổng số các chất kháng sinh, được thu nhận từ xạ khuẩn.

Thuật ngữ “nội sinh” (endophyte) được đặt ra lần đầu tiên bởi De Bary (1866), liên quan đến sự tồn tại của các vi sinh vật bên trong các mô thực vật [4]. Tuy nhiên, các nghiên cứu về nhóm vi sinh vật này bị lãng quên trong một thời gian dài. Từ thập niên 90, nhiều báo cáo đã phát hiện rằng xạ khuẩn thuộc chi *Streptomyces* và *Microspora* có thể sống trong thực vật như tác nhân cộng sinh, kí sinh và hoại sinh [5–7]. Do vậy, rất nhiều định nghĩa về vi sinh vật nội sinh được đưa ra. Định nghĩa được sử dụng rộng rãi nhất cho đến nay được đưa ra bởi Stone và cộng sự [8]. Theo đó, vi sinh vật nội sinh là vi sinh vật sinh trưởng trong mô tế bào thực vật, không gây ra những tác động xấu tới cây chủ, mà còn tăng cường khả năng trao đổi chất, kích thích sinh trưởng, chống côn trùng, miễn dịch cho vật chủ bằng cách tổng hợp các sản phẩm trao đổi chất [3,9]. Bổ sung thêm định nghĩa này, xạ khuẩn nội sinh được phân loại thành các nhóm khác nhau, gồm bắt buộc và có liên quan [10]. Xạ khuẩn nội sinh bắt buộc chỉ phụ thuộc vào quá trình trao đổi chất của thực

vật để tồn tại và cơ chế thích nghi cũng như tiến hoá dựa vào quá trình chuyển gen ngang (Horizontal gene transfer) [3]. Xạ khuẩn nội sinh có khả năng sinh sản trải qua một số giai đoạn nhất định trong vòng đời của chúng, độc lập với cây chủ. Chúng được liên kết gián tiếp với thực vật thông qua môi trường đất và khí quyển lân cận. Các nghiên cứu mở rộng đã được thực hiện để tìm ra nguồn gốc của xạ khuẩn nội sinh trong các loài thực vật khác nhau. Các xạ khuẩn phát tán theo dạng bào tử trong không khí và môi trường đất vùng rễ được coi là nguồn gốc chính của xạ khuẩn nội sinh [11].

### **1.1.2. Phân loại xạ khuẩn nội sinh dựa trên phân tích trình tự gen 16S rRNA**

Trong hệ thống phân loại, họ Streptomycetaceae thuộc lớp Actinobacteria và chứa hơn 10 chi như *Actinomyces*, *Nocardia*, *Streptomyces*, *Thermoactinomyces*, *Waksmania*, *Thermopolyspora*, *Micromonospora*, *Thermomonospora*, *Actinoplanes* và *Streptosporangium* [12,13]. Trong đó, *Streptomyces* là chi xạ khuẩn phổ biến và lớn với sự có mặt của 708 loài. Từ thế kỷ trước, các nhà khoa học đã tìm ra công cụ để phân loại xạ khuẩn đó là phân loại bằng lai DNA, lai RNA và phân tích trình tự gen mã hóa 16S rRNA. Trong đó, giải trình tự gen 16S rRNA được sử dụng chủ yếu trong các nghiên cứu về phân loại xạ khuẩn. Nguyên nhân do gen mã hóa 16S rRNA có kích thước khoảng 1540 bp có độ dài vừa đủ để so sánh với các trình tự tham chiếu. Sự giống nhau về gen trong 16S rRNA đã được coi là một đặc điểm phân tử quan trọng trong phân loại sinh vật nhân sơ bởi độ phổ biến, độ ổn định và tính bảo tồn cao. Gen 5S rRNA có kích thước ngắn (120 bp) và không đặc hiệu, trong khi đó gen 23S rRNA với kích thước khoảng 2900 bp gây khó khăn cho quá trình tách dòng và phân tích trình tự gen. Bổ sung, trình tự gen 16S rRNA được giải bằng Sanger, là công nghệ được sử dụng phổ biến cho giải trình tự DNA từ những năm 1980 cho đến nay.

Phân tích trình tự gen 16S rRNA là sự so sánh nét đặc trưng về trình tự nucleotide giữa các chủng xạ khuẩn với nhau. Nếu độ tương đồng giữa hai trình tự 16S rRNA là 98,6% thì xác suất để mức độ giống nhau trong phép lai DNA thấp hơn 70% [14]. Vì vậy, ngưỡng được coi để phân biệt hai loài khác

nhau là giá trị tương đồng 98,6% của trình tự 16S rRNA. Tuy nhiên, lập luận trên đã thay đổi từ khi công nghệ giải trình tự toàn bộ hệ gen thế hệ mới xuất hiện. Phân tích trình tự hệ gen của chủng xạ khuẩn *Streptomyces* sp. For3<sup>T</sup> phân lập từ đất vùng Champenoux, Pháp có độ tương đồng tới 99% với *Streptomyces pratensis* ATCC 33331<sup>T</sup>, *Streptomyces anulatus* ATCC 27416<sup>T</sup>, *Streptomyces setonii* NRRL ISP-5322<sup>T</sup> và *Kitasatospora papulosa* NRRL B-16504<sup>T</sup> [15]. Tuy nhiên, phân tích cây phân loại bằng trình tự hệ gen cho thấy, chủng For3<sup>T</sup> có mối quan hệ gần gũi nhất với chủng *Streptomyces atroolivaceus* NRRL ISP-5137<sup>T</sup> nhưng tách biệt thành một nhánh riêng biệt. Trên cơ sở những phân tích bổ sung như lai DNA, chất béo thành tế bào, sinh hoá, chủng For3<sup>T</sup> được định danh là loài mới *Streptomyces silvae*. Tới nay, phân loại và hệ thống hoá xạ khuẩn vẫn đang trên tiến trình chuẩn hoá, do vậy, để phân loại chính xác xạ khuẩn cần những phương pháp cập nhật và nghiên cứu chuyên sâu.

### 1.1.3. Xạ khuẩn *Streptomyces* nội sinh trên cây dược liệu

Cho tới nay, xạ khuẩn đã được phân lập từ nhiều loài thực vật khác nhau như cây lương thực (lúa mì, lúa, khoai tây), cây ngập mặn (vẹt, đước, sù) đến cây dược liệu (long huyết, chiêu liên, quế). Sự đa dạng của xạ khuẩn nội sinh thể hiện rõ ràng khi chúng được phát hiện từ các loài thực vật ở các vùng khí hậu khác nhau như: vùng khô hạn, ngập mặn, hệ sinh thái dưới nước và các hốc sinh thái khác [16]. Không giống như các nghiên cứu về tính đa dạng nấm, cho đến nay, chưa có định nghĩa và thước đo cụ thể về xạ khuẩn nội sinh. Trong các loại thực vật, cây dược liệu là một nguồn khai thác phong phú các hợp chất có hoạt tính sinh học quý giá. Do mối quan hệ mang tính lâu dài với cây chủ, xạ khuẩn nội sinh cũng có thể tham gia vào quá trình trao đổi chất và tăng cường hoạt tính sinh học tự nhiên của chính nó hoặc có thể thu được một số thông tin di truyền để tạo ra hợp chất mới. Chính vì vậy, cây dược liệu nguồn khai thác xạ khuẩn mới và chất mới.

Trong các chi xạ khuẩn, *Streptomyces* là chi thường được phân lập từ cây dược liệu và có nhiều tiềm năng phát hiện các hoạt chất mới. Tổng số chất chuyển hóa có hoạt tính sinh học do VSV tạo ra là khoảng 23.000 trong đó 10.000 (45% tổng số chất chuyển hóa có hoạt tính sinh học) được tạo ra

bởi riêng xạ khuẩn và trong đó có 7600 (76 %) hợp chất được báo cáo từ chi *Streptomyces* [17]. Từ 7 cây dược liệu thu thập ở Chiang Mai, Thái Lan, 330 chủng xạ khuẩn thuộc 4 chi *Streptomyces*, *Microbispora*, *Nocardia* và *Micromonospora* đã được phân lập [18]. Trên cây neem Ấn Độ (*Azadirachta indica*) trưởng thành, chi *Streptomyces* chiếm tỷ lệ lớn nhất (49,1%), tiếp theo là *Streptosporangium* (14,5%), *Microbispora* (10,9%) *Streptoverticillium* (5,5%), *Saccharomonospora* (5,5%) và *Nocardia* (3,6%) [19]. Tương tự, chi *Streptomyces* chiếm 68% trên tổng số lượng xạ khuẩn phân lập từ 4 dược liệu ở miền Nam Ấn Độ [20]. Nghiên cứu đa dạng cộng đồng vi khuẩn không qua nuôi cấy trên cây dược liệu họ đậu *Senna italica* bằng công nghệ metagenome cũng chứng minh *Streptomyces* là chi phổ biến và chiếm đa số trong nhóm xạ khuẩn [21]. Sự phân bố và đa dạng của *Streptomyces* trên cây dược liệu có thể do đặc tính sinh trưởng nhanh và hệ gen chứa nhiều gen mã hoá enzyme có thể chuyển hoá, tương tác với cây chủ.

Nhiều nghiên cứu đã chứng minh *Streptomyces* nội sinh trên cây dược liệu có nguồn gốc từ đất. Dịch tiết từ rễ cây đậu gà được bổ sung chủng xạ khuẩn *Streptomyces* sp. CP200B và theo dõi sự di chuyển của chủng xạ khuẩn trên cây. Kết quả cho thấy *Streptomyces* sp. CP200 được tìm thấy trong hạt của cây đậu gà 16 tuần sau đó, chứng minh sự di chuyển từ rễ lên ngọn [22]. Do vậy, dịch tiết rễ đóng vai trò then chốt trong sự thu hút xạ khuẩn *Streptomyces* có ích trên thực vật. Năm 2009, Shimizu và cộng sự đã xác định rằng *Streptomyces* sp. MBCu-56 xâm nhập vào lớp biểu bì của lá dưa chuột bằng thể sợi dày đặc bao trùm các điểm giao của tế bào biểu bì trên bề mặt lá và thâm nhập vào lớp biểu bì [23]. Ngoài ra, các enzym như cellulase, xylanases và pectinases sản xuất bởi xạ khuẩn cũng có thể tạo điều kiện xâm nhập vào cây chủ bằng cách phân hủy thành tế bào thực vật [24]. Đặc biệt là khả năng sinh ra expansin, một protein có nguồn gốc thực vật có vai trò "nới lỏng" các vi sợi cellulose ở thành tế bào, hỗ trợ việc xâm nhập của xạ khuẩn. Có thể thấy, cơ chế nội sinh và xâm nhiễm của xạ khuẩn *Streptomyces* khá đa dạng và cần nhiều nghiên cứu chứng minh các cơ chế này.



## 1.2. KHẢ NĂNG SINH TỔNG HỢP CÁC CHẤT CÓ HOẠT TÍNH Y DƯỢC TỪ XẠ KHUẨN *STREPTOMYCES* NỘI SINH

### 1.2.1. Các chất kháng vi sinh vật

Đến nay, đã có rất nhiều công bố về kháng sinh mới được tách chiết từ xạ khuẩn trên cây dược liệu. Chi *Streptomyces* là nhà sản xuất kháng sinh tiềm năng. Những loại kháng sinh này hiệu quả trong điều trị bệnh truyền nhiễm gây bởi vi sinh vật do các đặc tính dược động học mong muốn cho việc sử dụng lâm sàng [25]. Năm 2005, Guan và cộng sự đã công bố 3 acid mới có hoạt tính kháng sinh là p-aminoacetophenic của 7- (4-aminophenyl)-2,4-dimethyl-7-oxo-hept-5-enoic acid (9-(4-aminophenyl)-7-hydroxy-2,4,6-trimethyl-9-oxo-non-2-enoic acid và 12-(4-aminophenyl)-10-hydroxy-6-(1-hydroxyethyl)-7,9-dimethyl-12-oxo-dodeca-2,4 dienic acid [26]. Đặc biệt, nhóm nghiên cứu của Li và cộng sự tại Đại học Vân Nam (Trung Quốc) đã phát hiện được nhiều kháng sinh mới như: 9-hydroxybafilomycin D, 29-hydroxybafilomycin D, bafilomycin D, bafilomycin E, bafilomycin A1, bafilomycin B1, bafilomycin B2, bafilomycin C1, bafilomycin C2, bafilomycin C1 amide và bafilomycin C2 amide; caryolane-1,7 $\alpha$ -diol, 1,6,11-eudesmanetriol, 11-eudesmene-1,6-diol, 7,4 dihydroxy-8-(hydroxymethyl)-1methoxy-isoflavones và tripstretin từ *Streptomyces* [27,28]. Kakadumycin A kháng lại vi khuẩn bệnh than *Bacillus anthracis* được sản xuất bởi *Streptomyces* sp. 30566 [29]. Thử nghiệm mở rộng trên các chủng vi khuẩn khác (*Staphylococcus simulans*, *Enterococcus faecalis*, *Listeria monocytogenes*, *Shigella dysenteriae*, *Staphylococcus epidermidis*, *Staphylococcus aureus*, *Streptococcus pneumoniae*) cho thấy kakadumycin A có phổ kháng khuẩn mạnh.

Bên cạnh đó, nhóm xạ khuẩn nội sinh cũng có khả năng sinh tổng hợp các chất kháng nấm như chất 6–prenylindole tách chiết từ xạ khuẩn *Streptomyces* sp. TP-A0595 phân lập từ cây rêu tản (*Hepaticae* sp.) [30]. Coronamycin được sản xuất bởi chủng *Streptomyces* sp. ở nồng độ 2  $\mu$ g/ml (MIC) có tác dụng chống lại nấm gây bệnh trên người *Cryptococcus*

*neoformans* (MIC 4  $\mu\text{g/ml}$ ) [31]. Hai hợp chất mới là cedarmycins A và B được tách chiết từ dịch lên men chủng *Streptomyces* sp. TP- A0456 phân lập từ dịch chiết cây tuyết tùng trong đó chất cedarmycins A có hoạt tính kháng mạnh đối với nấm *Candida glabrata* gây bệnh trên người với giá trị nồng độ ức chế tối thiểu (MIC) đạt 0,4  $\mu\text{g/ml}$  [28]. Gần đây, chất saadamycin chống nấm gây bệnh hoắc lào và một số bệnh do nấm ký sinh trên da người được phát hiện từ dịch lên men chủng *Streptomyces* sp. Hedaya48 [32].

### 1.2.2. Các chất gây độc tế bào ung thư

Bên cạnh kháng sinh ức chế vi sinh vật gây bệnh, các chất có hoạt tính gây độc tế bào ung thư cũng được công bố từ xạ khuẩn nội sinh. Các cơ chế gần đây trong quá trình sinh sản của các tế bào khối u và sự chuyên đổi của chúng thành di căn đã dẫn đến việc phân loại và ứng dụng điều trị các hợp chất chống ung thư. Năm 2000, Caruso và cộng sự đã phân lập được các chủng *Streptomyces* sản xuất taxane, tiền chất của chất điều trị ung thư taxol, từ mô vỏ của *Taxus baccata* và *T. brevifolia* [33]. Các hợp chất chống ung thư đang được phân lập và thử nghiệm về độ hiệu quả với các dòng tế bào khác nhau. Kháng sinh peptide coronamycin từ *Streptomyces* sp. MSU-2110 cho thấy khả năng gây độc tế bào ngang với taxol bằng cách ức chế các dòng tế bào HMEC và BT20 với giá trị  $\text{IC}_{50}$  dao động từ 5-10  $\mu\text{g/ml}$ ) [31]. Chất 4-arylcoumarin từ *S. aureofaciens* CMUAc130 cho thấy hoạt tính chống ung thư nhờ tác dụng ức chế của chúng đối với LLC được cấy ghép ở chuột. Các cơ chế sinh hóa để ức chế sự tăng sinh khối u đã được thử nghiệm với dòng tế bào ung thư biểu mô tuyến A427 đối với sự biểu hiện quá mức của oncoprotein, điều này cho thấy rằng 5,7-dimethoxy-4-phenylcoumarin đã làm chậm tín hiệu nội bào dẫn đến sự tăng sinh khối u [34]. Lupinacidin A-B là chất chống ung thư thuộc nhóm anthraquinones được phân lập từ chủng *M. lupine* nội sinh trong rễ cây *Lupinus angustifolius* từ vùng trung-tây của Tây Ban Nha [35]. Lupinacidin A và B cho thấy khả năng ức chế mạnh trên tế bào ung thư biểu mô ruột kết 26-L5 mà không ảnh hưởng đến sự sinh trưởng của tế bào với giá trị  $\text{IC}_{50}$  lần lượt của lupinacidin A và B là 0,07 và 0,3  $\mu\text{g/ml}$ .

### 1.3. GIẢI MÃ HỆ GEN XẠ KHUẨN BẰNG CÔNG NGHỆ GIẢI TRÌNH TỰ HỆ GEN THẾ HỆ MỚI

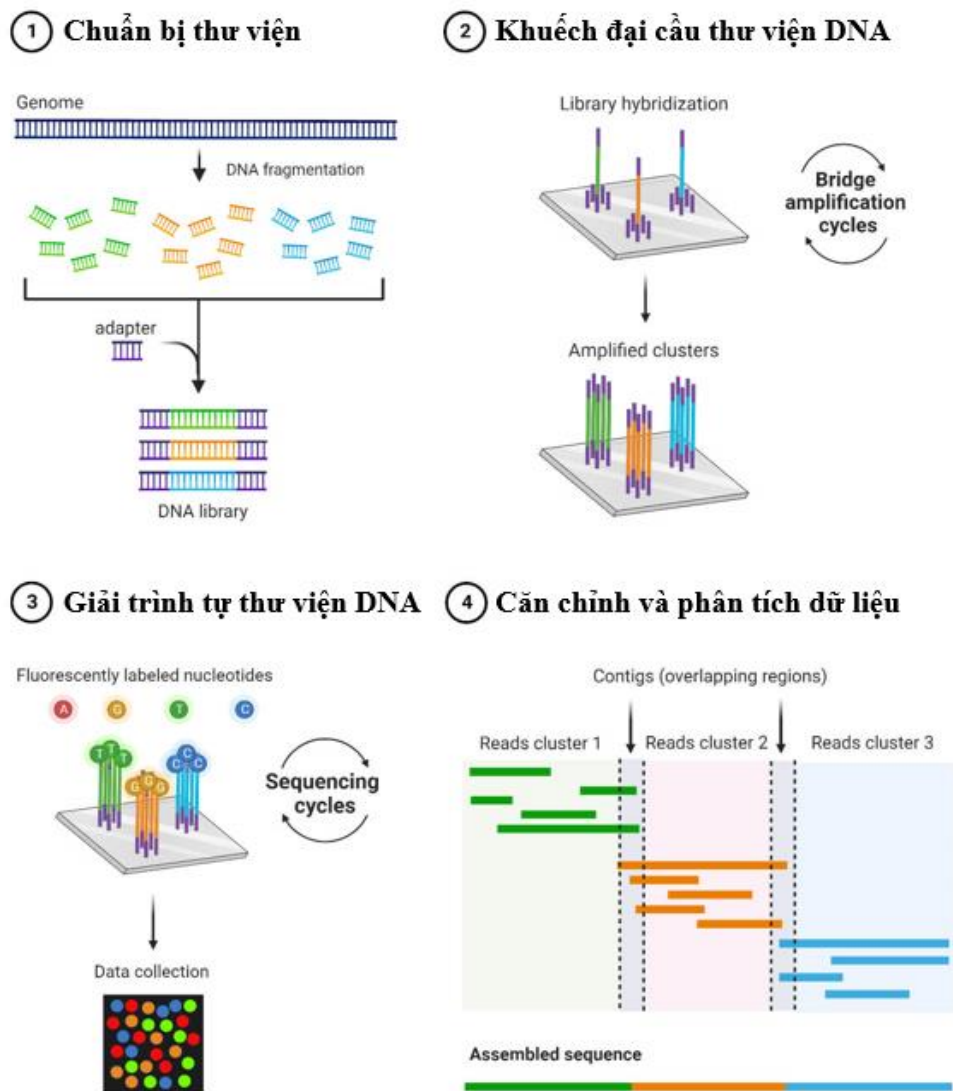
#### 1.3.1. Các công nghệ giải trình tự hệ gen được áp dụng hiện nay

Sự xuất hiện của các công nghệ giải trình tự đã đóng một vai trò quan trọng trong việc giải mã trình tự bộ gen của các sinh vật. Công nghệ giải trình tự đầu tiên được phát triển vào năm 1977 bởi Sanger và cộng sự [36]. Giải trình tự Sanger đã trở thành kỹ thuật giải trình tự được áp dụng nhiều vì hiệu quả cao và độ phóng xạ thấp và đã được thương mại hóa và tự động hóa với tên gọi "Công nghệ giải trình tự Sanger". Tuy nhiên, công nghệ này chỉ phù hợp với nghiên cứu đoạn gen có kích thước nhỏ hơn là với hệ gen. Cụ thể, đến năm 2004, chỉ có 192 hệ gen vi khuẩn có kích thước nhỏ được công bố và mỗi hệ gen tiêu tốn khoảng trên 50.000 đô la Mỹ [37].

Tuy nhiên, mọi thứ đã thay đổi do trình tự bộ gen người đầu tiên đã được giải mã bằng phương pháp Sanger vào năm 2003 với sự hợp tác của 15 quốc gia do Mỹ đứng đầu, đã tiêu tốn 3 tỷ đô la và kéo dài tới 13 năm [38]. Câu hỏi được đặt ra là, làm thế nào để có thể rút ngắn thời gian và giảm chi phí giải trình tự toàn bộ hệ gen. Với lý do này, Viện Nghiên cứu hệ gen người quốc gia (National Human Genome Research Institute, NHGRI – Hoa Kỳ) đã khởi động chương trình đầu tư với mục tiêu làm giảm chi phí giải mã hệ gen người xuống 1000 USD trong 10 năm. Đây là động lực thúc đẩy sự phát triển và thương mại hóa các công nghệ giải trình tự hệ gen thế hệ mới (Next Generation Sequencing, NGS). Giải trình tự thế hệ mới cho phép đồng thời giải mã hàng triệu trình tự DNA cùng lúc, nhờ vậy, nâng cao hiệu suất của quá trình giải mã bộ gen sinh vật nói chung và bộ gen xạ khuẩn nói riêng. Cụ thể, năm 2014, đã có 1000 hệ gen *Escherichia coli* được công bố, nhưng chỉ sau 3 năm con số đã lên tới 100.000 hệ gen [39]. Do đó, NGS đã tạo ra một cuộc cách mạng trong nghiên cứu hệ gen và sinh học phân tử.

Hiện tại có ba công nghệ được sử dụng rộng rãi để giải trình tự hệ gen xạ khuẩn là Oxford Nanopore, Illumina MiSeq và PacBio. Vào năm 2012, Illumina đã đạt giải trình tự tốt nhất với độ dài đọc chính xác cao, chất lượng

cao nhất và chi phí trên mỗi base thấp nhất. Công nghệ này có thể phân tích được các trình tự của một đoạn nucleotide lặp lại liên tục, đồng thời giúp cho thư viện DNA được mã hóa và tách riêng trong toàn bộ quá trình phân tích kết quả. Công nghệ của Illumina có thể phù hợp với hầu hết các nghiên cứu do độ chính xác cao, lượng DNA đầu vào thấp và thời gian giải trình tự ngắn. Giải trình tự Illumina tạo ra số lượng lớn các đoạn đọc (read) ngắn (< 300 bp) có chất lượng cao, rất tốt trong việc phát hiện các biến thể. Công nghệ Illumina được áp dụng rộng rãi cho các nghiên cứu hệ gen của virus, vi khuẩn, động vật và thực vật, với nhiều kiểu nguyên liệu đầu vào như DNA và RNA. Quá trình giải trình tự bằng công nghệ Illumina được thể hiện trong Hình 1.1.



**Hình 1.1. Giải trình tự hệ gen bằng công nghệ Illumina [40].**

Đầu tiên, các mẫu DNA được phân đoạn ngẫu nhiên thành các trình tự và các adapter được nối vào cả hai đầu của mỗi trình tự. Ngoài ra, các adapter này được cố định vào các adapter bổ sung tương ứng và được nối vào một thanh trượt với nhiều biến thể của adapter (bổ sung) được đặt trên một giá cố định. Sau đó, mỗi trình tự gắn được khuếch đại bằng khuếch đại cầu nối PCR tạo ra một số bản sao giống hệt nhau của mỗi trình tự; một tập hợp các trình tự được tạo ra từ cùng một trình tự ban đầu được gọi là một cụm. Mỗi cụm chứa khoảng một triệu bản sao của cùng một chuỗi ban đầu. Để xác định từng nucleotide trong trình tự, Illumina sử dụng phương pháp giải trình tự bằng phương pháp tổng hợp sử dụng các đầu cuối có thể đảo ngược trong đó bốn nucleotide đã sửa đổi, đoạn mỗi trình tự và DNA polymerase được thêm vào dưới dạng hỗn hợp và đoạn mỗi được lai với trình tự. Hiện nay, polymerase được sử dụng để mở rộng đoạn mỗi bằng cách sử dụng các nucleotide đã biến đổi (Hình 1.1). Mỗi loại nucleotide được gắn huỳnh quang cụ thể để mỗi loại là duy nhất. Các nucleotide có nhóm 3'-hydroxyl không hoạt động, đảm bảo rằng chỉ một nucleotide được kết hợp. Các cụm được kích thích bằng tia laser để phát ra tín hiệu ánh sáng dành riêng cho từng nucleotide, tín hiệu này sẽ được phát hiện bởi máy ảnh của thiết bị tích điện kép và các chương trình máy tính sẽ dịch các tín hiệu này thành chuỗi nucleotide. Quá trình này tiếp tục với việc loại bỏ chất kết thúc có nhãn huỳnh quang và bắt đầu một chu kỳ mới với sự kết hợp mới.

### **1.3.2. Đặc điểm hệ gen xạ khuẩn thuộc chi *Streptomyces***

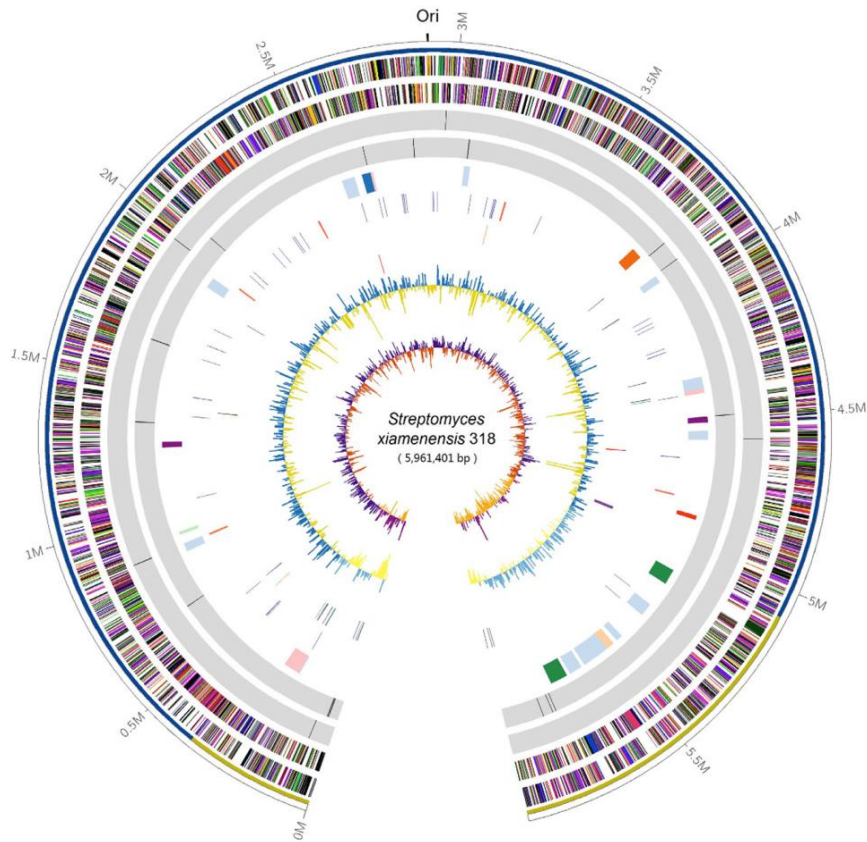
Hệ gen hai chủng xạ khuẩn *S. coelicolor* A3 (2) và *S. avermitilis* được công bố đầu tiên bằng công nghệ Sanger [41,42]. Công nghệ NGS đã cách mạng hóa lĩnh vực nghiên cứu xạ khuẩn và gia tăng đáng kể số lượng hệ gen *Streptomyces* kể từ năm 2013. Theo cơ sở dữ liệu RefSeq, tổng cộng 1749 hệ gen *Streptomyces* đã được đăng ký tính đến ngày 6 tháng 2 năm 2020 và hơn 73% bộ gen được giải trình tự bằng công nghệ NGS. Trong đó, 1549 hệ gen sơ bộ (draft genome) được giải trình tự bằng Illumina, 454, MinION và 200 hệ gen hoàn thiện (complete genome) được giải bằng PacBio. Số lượng hệ

gen xạ khuẩn *Streptomyces* ngày càng tăng theo các năm, hứa hẹn tạo ra cơ sở dữ liệu lớn để nghiên cứu xạ khuẩn.

Xạ khuẩn có đặc điểm hệ gen khác biệt với vi khuẩn. Hệ gen xạ khuẩn dạng thẳng và có tỷ lệ GC cao. Vùng sao chép *oriC* nằm ở trung tâm nhiễm sắc thể dạng thẳng và các trình tự lặp đảo ở đoạn kết thúc (terminal inverted repeats sequence, TIR) [43]. Đáng chú ý, các trình tự TIR có sự khác nhau giữa các loài và không tuyến tính với kích thước hệ gen. Đặc điểm khác biệt nhất của bộ gen *Streptomyces* là nhiễm sắc thể có độ ổn định thấp, dẫn đến hiện tượng xóa và sắp xếp lại tự phát thường xuyên, đặc biệt là ở phần cuối của nhiễm sắc thể. Ví dụ, khoảng 0,5% bào tử đang nảy mầm của *Streptomyces lividans* trải qua quá trình loại bỏ tới 25% bộ gen (~2 Mbp) trong điều kiện nuôi cấy trong phòng thí nghiệm. Chính vì vậy, các gen quan trọng liên quan đến duy trì tế bào, bao gồm phiên mã, dịch mã và sao chép DNA, nằm trong vùng “lõi” của nhiễm sắc thể [44]. Ngược lại, các gen tham gia thích nghi và chuyển hóa thứ cấp thường nằm trong vùng “cánh tay” của nhiễm sắc thể. Chính vì vậy, cụm gen mã hoá sản phẩm trao đổi thứ cấp thường bị biến đổi và không bảo tồn trong cùng một loài.

Xạ khuẩn *Streptomyces* là chi có kích thước hệ gen dao động lớn. Hệ gen có kích thước lớn nhất được ghi nhận ở *Streptomyces hygrosopicus* XM201 (12,0 Mb) và thấp nhất được quan sát ở *Streptomyces xiamenensis* 318 (5,96 Mb) [45,46]. Tương ứng, các gen mã hoá protein dao động từ 5.100 – 9.385. Tỷ lệ GC đạt khoảng 71%. Hệ gen *S. xiamenensis* 318 mã hoá cho 5491 protein chia thành 22 nhóm chức năng [47]. Số lượng gen mã hoá protein chưa rõ chức năng chiếm 52,7%. Kích thước lõi gen so với các chủng *Streptomyces* khác là 4,8 Mb và kích thước 2 vùng thay đổi là 0,36 Mb và 0,8 Mb. Đáng chú ý, *S. hygrosopicus* KCTC 1717 có 2 nhiễm sắc thể. Phân tích 121 hệ gen *Streptomyces* đã tìm ra 633 gen “lõi”, 1.080 gen “lõi mềm”, 6.040 gen “vỏ” và 137.709 gen “đám mây” [48]. Đặc biệt là sự phát hiện của 81.568 cụm gen duy nhất chỉ được phát hiện ở một trong số các hệ gen phân tích. Kết

quả trên chứng minh độ khác biệt về hệ gen *Streptomyces* so với các vi khuẩn khác.



**Hình 1.2. Bản đồ hệ gen chủng *Streptomyces xiamenensis* 318 [47].**

### **1.3.3. Các cụm gen sinh tổng hợp sản phẩm thứ cấp từ xạ khuẩn *Streptomyces***

Số lượng cụm gen sinh tổng hợp các chất chuyển hóa thứ cấp (secondary metabolite biosynthetic gene clusters, smBGC) trong xạ khuẩn dao động từ 8 – 83. Một số chủng xạ khuẩn có số lượng smBGC cao như *Streptomyces rhizosphaericus* NRRL B-24304 (83), *Streptomyces* sp. NRRL B-1347 (82), *Streptomyces* sp. PRh5 (82), *Streptomyces milbemycinicus* NRRL 5739 (81) và *Streptomyces* sp. NBS 14/10 (79) [49]. Trái ngược, số lượng smBGC được xác định ở *Streptomyces gilvigriseus* MUSC 26 (8), *Streptomyces thermoautotrophicus* H1 (9) and *Streptomyces thermoautotrophicus*UBT1 (11). Mặc dù số lượng thay đổi nhưng smBGC luôn chiếm khoảng 9 - 25% kích thước hệ gen [50]. Kết quả phân tích 121 hệ gen *Streptomyces* cho thấy 5289 smBGC khác nhau được tìm thấy, trong đó các smBGC có khung là peptide tổng hợp không thông qua ribosome (non-ribosomal peptide

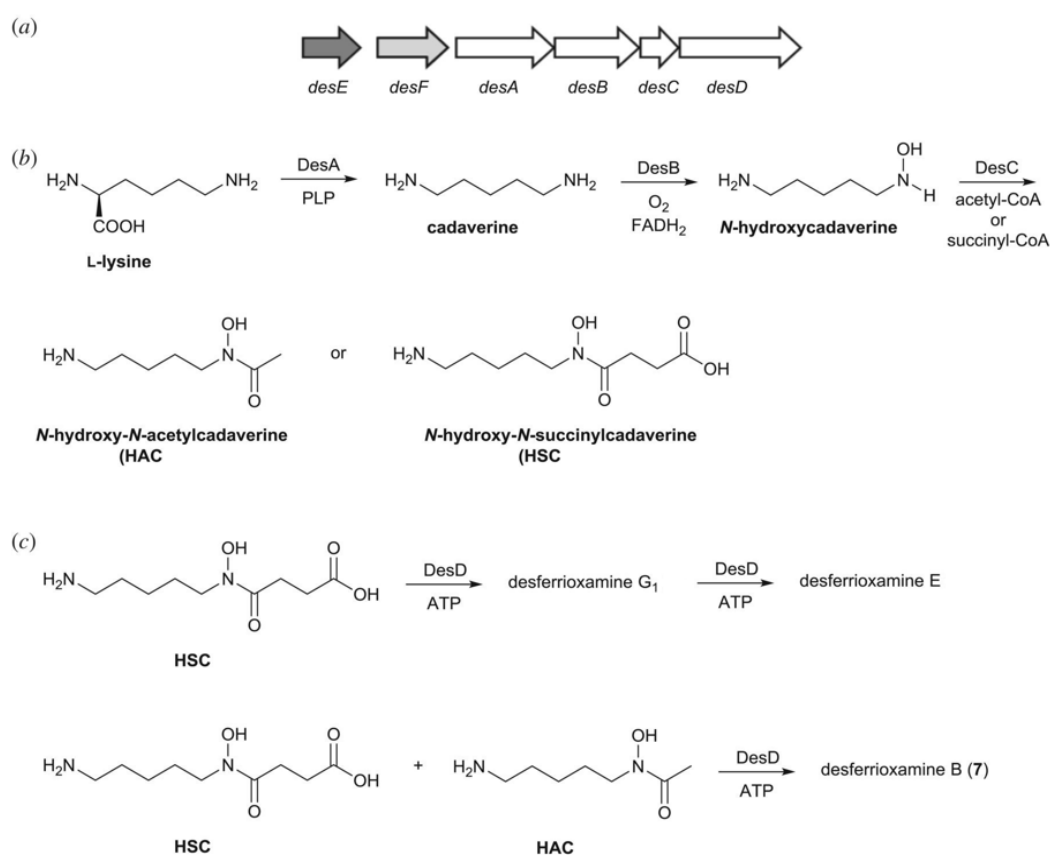
synthetase, NRPS), terpene, polyketide dạng 1 (type 1 polyketide synthase, T1PKS) và siderophore chiếm tới 50% tổng số smBGC được dự đoán. Ở mức độ thấp hơn là butyrolactone, polyketide dạng 2 (type 2 polyketide synthase, T2PKS) bacteriocin và polyketide dạng 3 (type 3 polyketide synthase, T3PKS). Ngoài ra, một số smBGC có kích thước lớn được tổng hợp từ nhiều loại enzym được gọi là smBGC lai. Mặc dù nguồn gốc và vai trò của các cụm gen này chưa rõ, các cụm gen này có khả năng sinh tổng hợp nhiều chất có hoạt tính kháng sinh. Trong 1110 hệ gen được phân tích, 1088 hệ gen mang ít nhất một smBGC lai [49]. Nguyên nhân do loài *Streptomyces* có cấu trúc hệ gen mở và các smBGC dễ bị tác động từ điều kiện ngoại cảnh, điều này tạo nên sự khác biệt của chi *Streptomyces* và các sản phẩm thứ cấp từ chúng so với các chi xạ khuẩn khác.

Trong vài thập kỷ qua, một số cơ sở dữ liệu đã được phát triển để khai thác smBGC. Các phần mềm tin sinh như BAGEL, ClustScan, NP.searcher, SMURF, antiSMASH, ClusterFinder, PRISM, EvoMining, RODEO và ARTS được xây dựng để khai thác smBGC từ hệ gen xạ khuẩn. Các công cụ này thực hiện các thuật toán để xác định ranh giới BGC và phát hiện các BGC tiềm năng dựa trên nhiều chỉ số như miền protein đặc trưng, mô tả của các enzym chuyển hóa chính và các dấu hiệu tiến hóa [51]. Các BGC tiềm năng sau đó được tách dòng và biểu hiện bởi các vật chủ. Các chất sau đó được tách chiết và xác định cấu trúc. Ngoài ra, công nghệ chỉnh sửa gen (CRISPR-Cas9) được sử dụng để cải biến di truyền trực tiếp trên tế bào xạ khuẩn.

smBGC tiêu biểu là desferrioxamine E, một chất có vai trò quan trọng trong sự hấp thu sắt, có khả năng gây độc tế bào ung thư. Desferrioxamine E được tổng hợp từ ba phân tử N-hydroxy-N-succinylcadaverine (HSC) trong *S. coelicolor* bởi enzyme DesD (Hình 1.3a) [52–54]. Quá trình này bao gồm sự adenyl hóa nhóm cacboxyl trong một đơn vị HSC, sau đó ngưng tụ với nhóm amin trong một đơn vị khác. Chất trung gian homodime thu được được adenylat hóa và ngưng tụ với phân tử HSC thứ ba để tạo ra desferrioxamine G1. Desferrioxamine E được hình thành từ homotrimer này thông qua một



vòng adenylation tiếp theo sau quá trình macrolactam hóa. DesD tổng hợp desferrioxamine B (7) thông qua ngưng tụ homodimer HSC adenylation hóa với N-hydroxy-N-acetylcadaverine (HAC). Mặc dù N-hydroxycadaverine, phát sinh từ quá trình khử carboxyl của l-lysine được xúc tác bởi DesA, sau đó là quá trình N-hydroxyl hóa được xúc tác bởi DesB của cadaverine, được biết đến là chất trung gian trong quá trình sinh tổng hợp desferrioxamine (Hình 1.3b), còn cơ chế các nhóm hydroxamate thải sắt chủ chốt được tích hợp vào desferrioxamine vẫn còn phải được xác định. Gen *desC* trong cụm gen sinh tổng hợp desferrioxamine mã hóa một acyl transferase phụ thuộc acyl-CoA giả định đã được đề xuất để xúc tác quá trình acyl hóa N-hydroxycadaverine với succinyl- và acetyl- CoA để tạo thành HSC và HAC, tương ứng (Hình 1.3c) [52]. Tuy nhiên, giả thuyết này vẫn chưa được chứng minh.



**Hình 1.3. (a) Cụm gen sinh tổng hợp desferrioxamine ở *Streptomyces coelicolor*. (b) Con đường sinh tổng hợp tiền chất HAC và HSC. (c) Sinh tổng hợp desferrioxamine E và B.**

#### 1.4. CÂY MÀNG TANG VÀ TIỀM NĂNG KHAI THÁC

Cây màng tang (*Litsea cubeba* (Lour.) Pers là cây dược liệu thường mọc hoang rải rác ở các vùng núi Hà Giang, Yên Bái, Điện Biên, Gia Lai, Kon Tum và Đà Lạt. Theo Đông y cổ truyền, cây màng tang có tác dụng ôn trung hạ khí, trừ thấp, giảm đau, tán phong hàn. Trong Y học hiện đại, màng tang còn có một số khả năng như: an thần, chống loạn nhịp tim, thiếu máu cơ tim,... Thành phần hóa học chính trong tinh dầu màng tang là monoterpenes (citral, limonene, citronellal,...) được chứng minh có nhiều hoạt tính sinh học [55].



**Hình 1.4. Cây màng tang (*Litsea cubeba* Lour. Pers).**

Nguồn: [www.ydhvn.com](http://www.ydhvn.com)

Tinh dầu từ quả màng tang đã được phân tích về hoạt tính kháng khuẩn kháng lại các vi khuẩn Gram dương và Gram âm. Nồng độ diệt khuẩn tối thiểu (MBC) của màng tang với *E. coli* là 0,38 µg/mL, với *B. subtilis* là 390,6 µg/mL và các giá trị MBC của màng tang đối với VSV kiểm định khác nhau là khác nhau. Màng tang còn thể hiện khả năng gây chết ấu trùng với giá trị LC<sub>50</sub> 31,62 µg/mL và hoạt tính DPPH với giá trị IC<sub>50</sub> 1628,85 µg/mL. Kết quả này cho thấy màng tang có hoạt tính kháng khuẩn và chống ung thư mạnh cũng như đặc tính chống oxy hóa nhẹ [56].

Tinh dầu lá và quả màng tang đã được đánh giá về hoạt tính gây độc tế bào ung thư ở người. Tinh dầu quả thể hiện hoạt tính chống ung thư đối với các tế bào ung thư miệng (OEC-M1), gan (J5) và phổi (A549) ở người với giá trị IC<sub>50</sub> đạt lần lượt 50, 50 và 100 ppm, còn tinh dầu lá không thể hiện hoạt tính chống tăng sinh đối với các tế bào ung thư được thử nghiệm [57]. Năm

2012, từ thân cây màng tang thu thập tại Quý Châu (Trung Quốc), Zhang và cộng sự đã tách chiết được 5 hợp chất mới thuộc nhóm isoquinoline alkaloid có khả năng kháng khuẩn với độ tinh sạch >90%. Hai chất (+)-N-(methoxycarbonyl)-N-norbulbodione và (+)-N-(methoxycarbonyl)-N-norisocorydione có hoạt tính ức chế 6 dòng tế bào ung thư khác nhau với giá trị IC<sub>50</sub> dao động 9,54-12,22  $\mu$ M và 9,83-11,96  $\mu$ M [58].

### 1.5. TÌNH HÌNH NGHIÊN CỨU XẠ KHUẨN NỘI SINH TRÊN CÂY DƯỢC LIỆU TẠI VIỆT NAM

Tại Việt Nam đã có nhiều nghiên cứu về xạ khuẩn từ môi trường đất, suối nước nóng, biển nhưng các công trình chỉ dừng lại ở phân lập và sàng lọc hoạt tính kháng sinh hay gây độc tế bào ung thư. Năm 2013, từ các mẫu bùn đất (Cần Giờ), Hoàng Thị Hồng và Nguyễn Ngọc Phương đã phân lập được 55 chủng xạ khuẩn và sàng lọc được 11/55 chủng xạ khuẩn có khả năng đối kháng nấm *Fusarium* sp [59]. Năm 2014, Dương Minh Lam và cộng sự đã phân lập được 52 chủng XKNS trên ba loại cây ngập mặn tại Nam Định, trong đó cao chiết xuất thô của 2 chủng 2E20 và 2E29 có hoạt tính kháng 4 dòng tế bào ung thư KB, SK-LU-1, HepG2 và MCF7 ở nồng độ 128  $\mu$ g/ml [60].

Ở Việt Nam, những nghiên cứu về xạ khuẩn nội sinh bắt đầu được thực hiện. Từ thân cây long huyết (*Dracaena cochinchinensis*) thu thập tại rừng Cúc Phương, chủng xạ khuẩn *Streptomyces* sinh kháng sinh actinomycin D ức chế vi khuẩn gây bệnh (*Staphylococcus epidermis* ATCC35984 kháng methicillin, *Staphylococcus aureus* ATCC25923 kháng methicillin, *Escherichia coli* ATCC25922 và *Klebsiella pneumoniae* ATCC13883) và tế bào ung thư (gan Hep G2 và phổi MCF7) được phân lập [61]. Năm 2019, Vũ Thị Hạnh Nguyên và cộng sự đã phân lập được 111 chủng xạ khuẩn nội sinh trên cây quế thu thập tại các tỉnh Hoà Bình. Tỷ lệ xạ khuẩn phân bố chủ yếu trên thân cây và các chủng được phân loại thuộc về các chi *Streptomyces*, *Microbacterium* và *Nocardia*. Trong nghiên cứu này, có 38/111 chủng xạ khuẩn có phổ kháng vi sinh vật kiểm định mạnh và rộng [62,63].

Công nghệ NGS hiện đang bước đầu được áp dụng trong nghiên cứu và khai thác hệ gen vi sinh vật tại Việt Nam trong vài năm gần đây. Tương tự, Trần Xuân Thạch và cộng sự (2019) đã sử dụng WGS để phác thảo cơ chế lây nhiễm của chủng vi khuẩn *Neisseria meningitidis* gây bệnh viêm màng não mô cầu [64]. Trong nghiên cứu về vi khuẩn biển, Đoàn Văn Thước và cộng sự (2020) đã giải mã cơ chế sinh tổng hợp nhựa sinh học polyhydroxyalkanoate và ectoine từ chủng vi khuẩn ưa mặn *Salinivibrio proteolyticus* M318 phân lập từ mắm tôm [65]. Xạ khuẩn biển *Streptomyces* sp. PDH23 phân lập từ bọt biển *Rhabdastrella globostellata* được giải trình tự bằng công nghệ Illumina [66]. Ở khía cạnh giải mã cơ chế thích nghi của vi sinh vật, Quách Ngọc Tùng và cộng sự đã chứng minh rằng chủng xạ khuẩn *Streptomyces parvulus* VCCM 22513 đã sử dụng 26 gen sinh mycothiol, ergothioneine; 43 gen tham gia cân bằng oxy hoá, 87 gen chống lại sự thâm thấu và 55 gen phân huỷ các chất thơm (aromatic compounds) để thích nghi với môi trường nội sinh trên cây sù. So với các chủng xạ khuẩn phân lập từ biển hay trên cạn, con đường phân huỷ protocatechuate, gentisate, 4-hydroxyphenylpyruvate, cinnamate, 3-phenylpropionate, styrene lần đầu tiên được phát hiện ở *S. parvulus* VCCM 22513 [67].

Cây màng tang (*Litsea cubeba* Lour. Pers) là cây dược liệu được sử dụng để tách chiết tinh dầu và khai thác các hoạt chất chống tế bào ung thư. Quách Ngọc Tùng và cộng sự (2021) đã nghiên cứu xạ khuẩn nội sinh trên cây màng tang [68]. Tuy nhiên, các chủng xạ khuẩn *Streptomyces* khác sinh chất kháng khuẩn và gây độc tế bào ung thư chưa được khai thác. Ngoài ra, hệ gen các chủng *Streptomyces* cũng chưa được nghiên cứu. Với những tiềm năng nghiên cứu trên, đề tài luận văn thạc sĩ: “**Nghiên cứu hoạt tính kháng sinh, gây độc tế bào ung thư và đặc điểm di truyền hệ gen của xạ khuẩn *Streptomyces* sp. VCCM 22255**” được thực hiện nhằm giải mã mối liên hệ di truyền hệ gen liên quan đến các hoạt tính kháng sinh và gây độc tế bào ung thư của chủng xạ khuẩn nội sinh trên cây màng tang. Ngoài ra, nghiên cứu

cũng tạo cơ sở để phát triển công nghệ cải biến di truyền hệ gen xạ khuẩn trong tương lai.

## **Chương 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU**

### **2.1. ĐỐI TƯỢNG**

#### **2.1.1. Chủng giống**

Bộ chủng xạ khuẩn phân lập từ cây màng tang (*Litsea cubeba* (Lour.) Pers) thu thập tại Phú Thọ gồm 15 chủng thuộc Bộ sưu tập giống của Trung tâm Giống và Bảo tồn nguồn gen Vi sinh vật, Viện Công nghệ sinh học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, trong đó chủng *Streptomyces* sp. LCP02 được ký gửi dưới mã số VCCM 22255. Các chủng xạ khuẩn *Streptomyces* sp. VCCM 22255 được lưu trữ trong 30% (v/v) glycerol ở -80°C.

Các chủng vi khuẩn kiểm định nhận từ Bộ sưu tập giống của Trung tâm Giống và Bảo tồn nguồn gen Vi sinh vật, Viện Công nghệ sinh học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, bao gồm: *Salmonella enterica* subsp. *enterica* serovar Typhimurium ATCC 14028 (*S. Typhimurium* ATCC 14028), *Escherichia coli* ATCC 11105, *Bacillus cereus* ATCC 11778, *Pseudomonas auroginosa* ATCC 9027, *Enterobacter aerogenes* ATCC 13048, *Staphylococcus aureus* kháng methicillin ATCC 25923 (MRSA), *Staphylococcus epidermidis* kháng methicilin ATCC 35984 (MRSE), *Proteus vulgaris* ATCC 49132.

Các dòng tế bào ung thư được thử nghiệm gồm: A549 (ung thư phổi người), MCF7 (ung thư vú người) và Hep3B (ung thư gan)

#### **2.1.2. Hóa chất và thiết bị**

Một số hóa chất chính được sử dụng trong đề tài: cao nấm men (Himedia, Ấn Độ), tryptone (Himedia, Ấn Độ), cao malt (Himedia, Ấn Độ), NaCl (Trung Quốc), glucose (Trung Quốc), ethyl acetate (Merck, Đức), dimethyl sulfoxide (DMSO) (Merck, Đức) và một số hóa chất thí nghiệm khác.

Các bộ kit được sử dụng trong nghiên cứu: bộ kit tinh sạch PureLink™ – DNA Purification (Invitrogen, Mỹ), bộ kit tách hệ gen của xạ khuẩn PureLink® Genomic DNA Kit (Thermo Fisher Scientific, Mỹ).

Các thiết bị, dụng cụ được sử dụng trong nghiên cứu này: Đĩa petri, micropipettes, đầu tip (Isolab, Đức); máy PCR (Applied Biosystem 9700, Mỹ); máy ly tâm lạnh (Biofuge fresco, Kendro, Đức); máy ổn nhiệt (Labnet, Mỹ); bộ điện di agarose, máy soi gel (Bio-Rad, Mỹ); máy vortex (vision scientific, Hàn Quốc); máy khuấy từ (vision scientific, Hàn Quốc); tủ cấy an toàn sinh học Class II (Esco, Anh); tủ nuôi (Sanyo, Nhật Bản); cân kỹ thuật (Precisa XT2200A, Thụy Điển); pH kế Melter Toledo (Đức); tủ sấy Cornthem (New Zealand); máy giải trình tự ABI PRISM®3100-Avant Genetic Analyzer (Applied Biosystems, Mỹ); máy quang phổ NanoDrop Spectrophotometer (Thermo Fisher Scientific Inc., Mỹ); máy đọc trình tự Ion Torrent PGM (Life Technologies, Mỹ); các phần mềm tin sinh học và một số thiết bị, dụng cụ nghiên cứu khác.

### 2.1.3. Môi trường

Môi trường LB (Luria-Bertani) (g/l): Cao nấm men 5,0; tryptone 10,0; NaCl 10,0; thạch 15,0; nước cất 1000 ml; pH 6,5-7,0.

Môi trường YIM38 (g/l): Cao malt 4,0; cao nấm men 4,0; glucose 4,0; thạch 20,0; nước cất 1000 ml; pH 7,2.

## 2.2. PHƯƠNG PHÁP NGHIÊN CỨU

### 2.2.1. Đánh giá hoạt tính kháng vi sinh vật kiểm định của xạ khuẩn

Hoạt tính kháng vi khuẩn kiểm định của chủng xạ khuẩn được xác định bằng phương pháp đục lỗ thạch có hiệu chỉnh theo phương pháp của Saadoun và Muhana [69]. Chủng xạ khuẩn được nuôi cấy trên môi trường YIM38 trên máy lắc tốc độ 200 vòng/phút trong 5 ngày ở nhiệt độ 30°C. Sau 5 ngày, thu hồi và ly tâm dịch nuôi cấy các chủng xạ khuẩn với tốc độ 6.000-8.000 vòng/phút ở 4°C. Trên môi trường LB thạch, cấy trang vi khuẩn kiểm định có mật độ tế bào đạt  $5.10^8$  CFU/ml, sau đó đục lỗ thạch bằng dụng cụ đục lỗ thạch. Nhỏ 100 µl dịch nuôi cấy vào mỗi lỗ và để đĩa ở 4-5°C trong 4-6 giờ, sau đó đem nuôi ở tủ ấm 36°C trong 18-24 giờ. Hoạt tính kháng VSV kiểm định được xác định theo kích thước vòng kháng khuẩn (VKK, mm) xuất hiện xung quanh lỗ thạch đã nhỏ dịch, vòng càng lớn thì hoạt tính càng mạnh theo công thức:

$$VKK = D - d \text{ (mm)}$$

Trong đó: D: Đường kính vòng vô khuẩn (mm);

d: Đường kính lỗ thạch (mm).

Biểu đồ heatmap về trung bình đường kính vòng kháng khuẩn của các chủng được xây dựng trên Heatmapper theo phương pháp liên kết trung bình Average linkage (<http://www.heatmapper.ca/expression/>).

### 2.2.2. Đánh giá khả năng gây độc tế bào ung thư

Phương pháp tách chiết các hoạt chất sinh học từ xạ khuẩn được tiến hành theo mô tả của Saadoun và Muhana [69]. Chủng xạ khuẩn được nuôi lắc ở 30°C trong 5 ngày trong 300 ml môi trường YIM38. Sau khi lên men, dịch nuôi được lọc qua bốn lớp vải thưa để loại bỏ tế bào. Phần dịch lọc được chiết với ethyl acetate theo tỉ lệ 1:1, lặp lại hai lần và thu phần dung môi phía trên, sau đó cô quay ở 60°C, thu được cao chiết thô. Cao chiết thô được bảo quản ở 4°C cho thí nghiệm tiếp theo.

Phương pháp thử độ độc tế bào *in vitro* được Viện Ung thư Quốc gia Hoa Kỳ (National Cancer Institute – NCI), là phép thử độ độc tế bào chuẩn nhằm sàng lọc và phát hiện các chất có khả năng kìm hãm sự phát triển hoặc diệt tế bào ung thư (TBUT) ở điều kiện *in vitro*, được thực hiện theo phương pháp của Skehan và cộng sự (1990) [70]. Phép thử tiến hành xác định hàm lượng protein tế bào tổng số dựa vào mật độ quang học (OD – Optical Density) đo được khi thành phần protein của tế bào được nhuộm bằng Sulforhodamine B (SRB). Giá trị OD tỉ 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:

- 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.
- Cao chiết ở nồng độ 100 µg/mL được đưa vào các giếng của đĩa 96 giếng. Giếng không có chất thử nhưng có TBUT (190 µL) được sử dụng làm đối chứng ngày 0. Sau 1 giờ, giếng đối chứng ngày 0 tế bào được cố định bằng Trichloroacetic acid – TCA 20%.
- 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, rửa 3 lần bằng acetic acid rồi để khô ở nhiệt độ phòng.

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

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

- Phép thử được lặp lại 3 lần để đảm bảo tính chính xác. Camptothecin ở nồng độ 10  $\mu\text{M}$  được sử dụng làm chất đối chứng.
- DMSO 1% luôn được sử dụng như đối chứng âm (nồng độ cuối cùng trong giếng thử là 0.05%). Giá trị  $\text{IC}_{50}$  (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), cao chiết được coi có hoạt tính tốt với  $\text{IC}_{50} \leq 20 \mu\text{g/ml}$ , chất tinh khiết được coi có hoạt tính tốt khi  $\text{IC}_{50} \leq 5 \mu\text{M}$  [71].

### 2.2.3. Phân loại xạ khuẩn dựa trên phân tích trình tự gen 16S rRNA

DNA tổng số của chủng xạ khuẩn được tách chiết theo hướng dẫn của bộ kit tách DNA PureLink® Genomic DNA Kit. Trình tự gen 16S rRNA của xạ khuẩn được khuếch đại bằng phản ứng PCR bằng cặp mồi 27F 5'-TAACACATGCAAGTCGAACG-3' và 1429R 5'-GGTGTGACGGGCGGTGTGTA-3' với chu trình nhiệt theo Salam và cộng sự [72]. Sản phẩm của phản ứng PCR được kiểm tra bằng điện di trên gel agarose 1,0%. Kích thước của các đoạn DNA thu được sau phản ứng PCR được so sánh với thang DNA chuẩn (Thermo scientific, Mỹ). Sản phẩm PCR được tinh sạch bằng bộ kit PureLink™ – DNA Purification (Invitrogen, Mỹ) và gửi đến First BASE Laboratories Sdn. Bhd. (Malaysia) để giải trình tự.

Dùng phần mềm BioEdit 7.2 để kiểm tra trình tự gen hai chiều, kết quả căn chỉnh được so sánh với các gen tham chiếu bằng công cụ BLAST trên ngân hàng cơ sở dữ liệu GenBank NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). Cây phân loại được xây dựng thông qua phần mềm MEGA 11.0 dựa trên cơ sở khoảng cách di truyền theo Tamura-Nei, bằng việc sử dụng phương pháp Neighbor-joining với giá trị bootstrap 1000 lần thử (Chu Văn Mẫn 2011).



## 2.2.4. Giải trình tự và phân tích hệ gen của chủng xạ khuẩn tuyển chọn

### 2.2.4.1. Giải trình tự hệ gen chủng *Streptomyces* sp. VCCM 22255

Chủng xạ khuẩn VCCM 22255 được nuôi trong môi trường YIM38 trong 48 giờ, ly tâm thu tế bào ở 10.000 vòng/phút trong 5 phút. Hệ gen của chủng xạ khuẩn được tách chiết theo sách hướng dẫn của kit PureLink® Genomic DNA (Invitrogen, CA, USA). Nồng độ và chất lượng của DNA được đánh giá bằng điện di DNA trên 0,9% agarose, máy quang phổ Nanodrop UV-Vis (Thermo Fisher Scientific, Hoa Kỳ) và Qubit 2.0 fluorometer (Life Technologies, Mỹ).

Mẫu DNA đạt tiêu chuẩn được giải trình tự toàn bộ hệ gen bằng hệ thống máy Illumina MiSeq với kích thước trình tự đọc là 2x150bp và độ bao phủ 81X. Thư viện DNA được tạo nhờ phân cắt DNA tổng số thành các đoạn ngắn theo hướng dẫn của hãng Illumina. Sau khi điện di để chọn các mảnh DNA có kích thước khoảng 200-300 bp để chuẩn bị giải trình tự, DNA được xử lý để mặc định gắn thêm A vào đầu 3' và 5' của mỗi mạch đơn để gắn với adaptor và index có trình tự đặc thù của Illumina p5 (CACTGC) và p7 (GCGCTA) (IDT Gen Custom Blocking Oligos). Làm giàu các đoạn DNA mục tiêu bằng phản ứng PCR sử dụng cặp mồi đặc hiệu của Illumina. Để kiểm tra chất lượng của quá trình làm giàu mục tiêu, phản ứng qPCR được thực hiện với cặp mồi đặc hiệu cho mỗi vùng mục tiêu. Cuối cùng tất cả các mẫu DNA mục tiêu được gộp chung lại với nhau để giải trình tự. Sau khi giải trình tự, các đoạn trình tự thô (reads) được phân tích bằng phần mềm FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Các read có chất lượng thấp ( $Q < 15$ ) và adapter bị loại bỏ bằng phần mềm Fastp (<https://github.com/OpenGene/fastp>).

### 2.2.4.2. Lắp ghép *de novo* và chú giải hệ gen xạ khuẩn

Quá trình lắp ráp hệ gen dựa trên các Illumina reads được thực hiện phần mềm Spades (<https://github.com/ablab/spades>) và MegaHit (<https://github.com/voutcn/megahit>). Sử dụng BUSCO v.3 để đánh giá mức độ hoàn thiện của trình tự lắp ghép *de novo*. Hệ gen chủng xạ khuẩn VCCM 22255 được chú giải bằng RASTtk trên nền cơ sở dữ liệu SEED [73]. Các gen mã hoá protein được phân theo chức năng được phân tích tích dựa trên

EggNOG-mapper v1.0.3 [74]. Dùng phần mềm PATRIC (Bacterial and Viral Bioinformatics Resource Center (BV-BRC)) để tạo bản đồ hệ gen [75].

Đặc điểm hệ gen chủng xạ khuẩn *Streptomyces* sp. VCCM 22255 được so sánh với hệ gen tham chiếu từ các chủng *Streptomyces cacaoi* DSM 40057 (JABELW000000000.1), *Streptomyces avertimilis* ATCC 31267 (BJHY000000000.1), *Streptomyces griseus* IFO 13350 (AP009493.1), *Streptomyces coelicolor* A3(2) (AL645882.2), *Streptomyces fulvissimus* DSM40593 (AP017424.1), *Streptomyces albus* DSM 40763 (RCIY000000000.1).

### 2.2.4.3. Phân tích gen và cụm gen sinh tổng hợp sản phẩm trao đổi chất bậc hai và mã hoá enzym chức năng

Các gen mã hoá enzym chức năng (Carbohydrate-active enzyme, CAZymes) được phân nhóm và phân tích dựa trên cơ sở dữ liệu dbCAN [76]. Các smBGC được dự đoán bằng cách sử dụng cơ sở dữ liệu AntiSMASH 7 (<https://antismash.secondarymetabolites.org/#!/start>). Dùng trình tự protein tham chiếu để phân tích lần lượt các khung đọc mở dựa theo công cụ BLASTp (giá trị e-off-off =  $1e^{-5}$ ). Các trình tự axit amin được giống với thông số mặc định. Các gen đã dự đoán được chú thích với kết quả mô tả chất lượng tốt nhất (với tỷ lệ % chiều dài giống cao nhất và độ tương đồng cao nhất). Tất cả các gen đã chú giải sau đó được phân loại dựa trên các lớp COG của chúng.

### 2.2.5. Xử lý thống kê số liệu

Kết quả nghiên cứu được xử lý theo phương pháp thống kê sinh học trên phần mềm Microsoft Excel 2016. Phần mềm sử dụng các hàm tính toán dựa trên những công thức cơ bản của thống kê, trong đó có công thức:

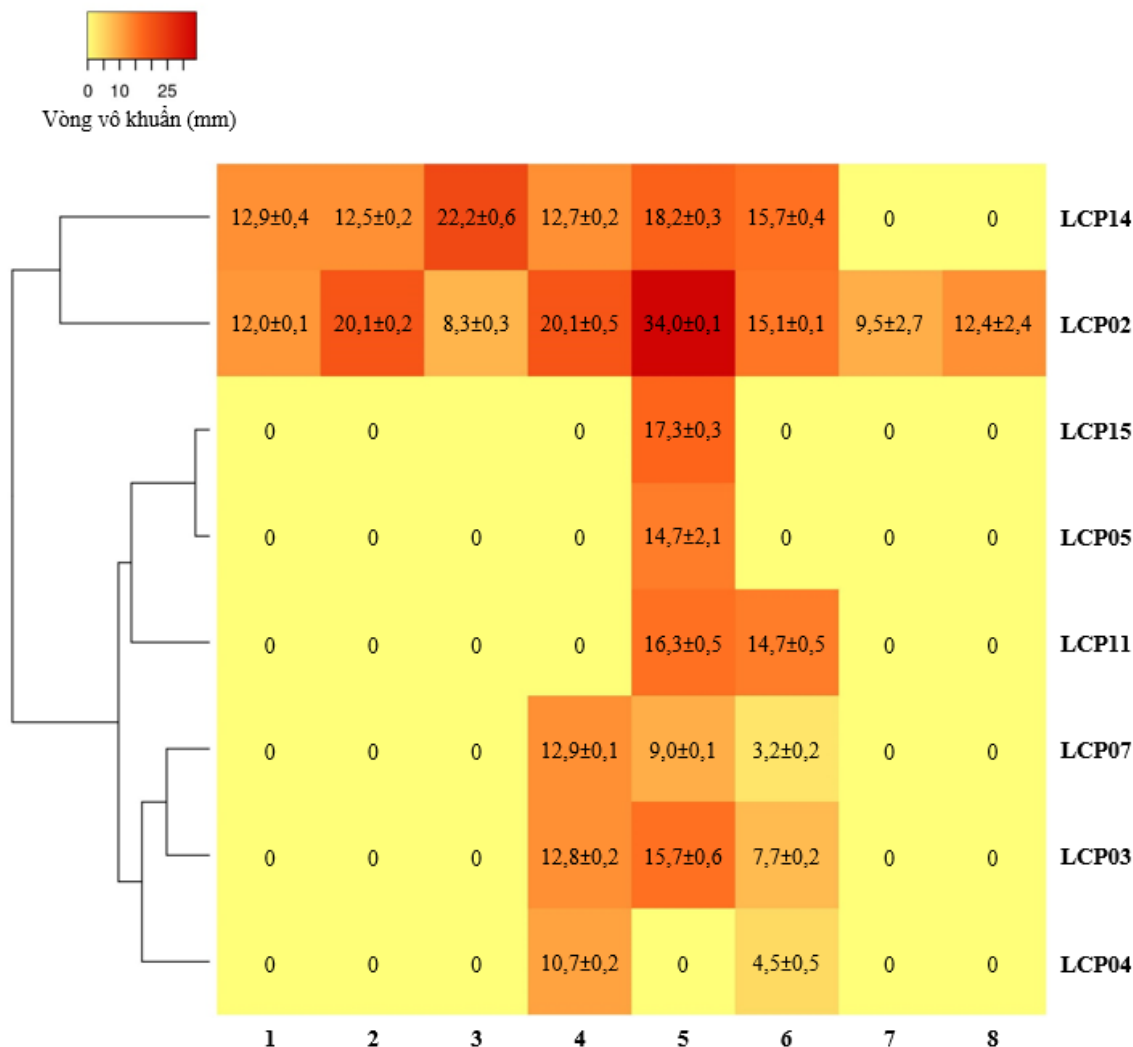
$$\text{Giá trị trung bình (X): } X = \frac{1}{n} \sum_{i=1}^n X_i;$$

$$\text{Độ lệch chuẩn (}\delta\text{): } \delta = \sqrt{\frac{\sum_{i=1}^n (X_i - X)^2}{n-1}}; \text{ Sai số (m): } m = \pm \frac{\delta}{m}$$

### Chương 3. KẾT QUẢ VÀ THẢO LUẬN

#### 3.1. KHẢ NĂNG KHÁNG KHUẨN CỦA XẠ KHUẨN NỘI SINH TRÊN CÂY MÀNG TANG

Từ 15 chủng xạ khuẩn nội sinh trên cây màng tang (*Litsea cubeba* (Lour.) Pers), nghiên cứu sàng lọc khả năng kháng vi khuẩn kiểm định của các chủng xạ khuẩn được tiến hành và thể hiện ở Hình 3.1.



**Hình 3.1. Sàng lọc hoạt tính kháng khuẩn của các chủng xạ khuẩn nội sinh trên cây màng tang.**

Ghi chú: (1) *S. Typhimurium* ATCC 14028; (2) *E. coli* ATCC 11105; (3) *MRSA* ATCC 25923; (4) *MRSE* ATCC 35984; (5) *P. aeruginosa* ATCC 9027; (6) *B. cereus* ATCC 11778; (7) *P. vulgaris* ATCC 49132; (8) *E. aerogenes* ATCC 13048.

Kết quả cho thấy có 8/15 chủng xạ khuẩn thể hiện hoạt tính kháng khuẩn đối với ít nhất một chủng vi khuẩn gây bệnh, chiếm khoảng 53,33% tổng số chủng xạ khuẩn thử nghiệm (Hình 3.1). Trong số các chủng thể hiện

hoạt tính, 4/8 chủng thể hiện phổ kháng khuẩn với ít nhất 3 chủng vi khuẩn kiểm định. Quan sát phổ kháng khuẩn với cả 8 vi khuẩn kiểm định thử nghiệm cho thấy *P. aeruginosa* ATCC 9027 và *B. cereus* ATCC 11778 nhạy cảm nhất với các dịch nuôi cấy xạ khuẩn.

Đáng chú ý, hai chủng xạ khuẩn LCP02 và LCP14 thể hiện phổ kháng khuẩn rộng và mạnh với ít nhất 6 vi khuẩn kiểm định. Cụ thể, chủng LCP02 ức chế mạnh nhất đối với *P. auroginosa* ATCC 9027 với đường kính vòng kháng khuẩn đạt  $34,0 \pm 0,1$  mm, tiếp đến là *E. coli* ATCC 11105 và MRSE ATCC 35984 với đường kính khoảng 20,1 mm. Tuy nhiên, hoạt tính kháng khuẩn với *S. Typhimurium* ATCC 14028, MRSA ATCC 25923, *B. cereus* ATCC 11778, *P. vulgaris* ATCC 49132 và *E. aerogenes* ATCC 13048 chỉ nằm trong khoảng từ 8,3-15,1 mm (Hình 3.1). Ở mức độ thấp hơn, chủng LCP14 ức chế *S. Typhimurium* ATCC 14028 ( $12,9 \pm 0,4$  mm), *E. coli* ATCC 11105 ( $12,5 \pm 0,2$  mm), MRSA ATCC 25923 ( $22,2 \pm 0,6$  mm), MRSE ATCC 35984 ( $12,7 \pm 0,2$  mm), *P. auroginosa* ATCC 9027 ( $18,2 \pm 0,3$  mm), *B. cereus* ATCC 11778 ( $15,7 \pm 0,4$  mm).

Xạ khuẩn nội sinh trên cây dược liệu, đặc biệt là tại các vùng nhiệt đới có thể là nguồn nguyên liệu tiềm năng cung cấp các chất kháng khuẩn đặc hiệu. Năm 2015, Peng và cộng sự đã phân lập được 119 chủng xạ khuẩn từ 13 cây dược liệu tại tỉnh Tứ Xuyên (Trung Quốc), trong đó có 12/80 chủng (chiếm 15%) thể hiện hoạt tính kháng ít nhất một VSV gây bệnh; 41,67% chủng có tính kháng *P. aeruginosa* và 33,33% kháng *S. epidermidis* [77]. Cũng năm 2015, Gohain và cộng sự đã phân lập 76 chủng XKNS từ 6 cây dược liệu tại Assam (Ấn Độ). Kết quả thử hoạt tính kháng VSV kiểm định cho thấy có 21/76 chủng (27,63%) có tính kháng với ít nhất một chủng VSV kiểm định. Đặc biệt, trong đó có 12 chủng (chiếm 15,79%) thể hiện hoạt tính kháng *S. aureus* và không có chủng nào thể hiện hoạt tính kháng *E. coli* và *P. aeruginosa* [78]. Năm 2020, trong số 111 chủng XKNS phân lập từ các mẫu cây quế *Cinnamomum cassia* Presl (gồm rễ, thân, lá), có 38 chủng (chiếm 34,2%) có hoạt tính kháng khuẩn đối với ít nhất 1 trong 9 VSV kiểm định với vòng vô khuẩn rộng từ 10-20 mm, đặc biệt có ghi nhận những chủng tạo vòng >20 mm [63]. Như vậy, có thể thấy rằng tỷ lệ xạ khuẩn nội sinh trên cây màng tang có hoạt tính kháng khuẩn là cao.

### 3.2. ĐÁNH GIÁ KHẢ NĂNG GÂY ĐỘC TẾ BÀO UNG THƯ CỦA CÁC CHỦNG XẠ KHUẨN TIỀM NĂNG

Ngoài khả năng sinh kháng sinh ức chế vi khuẩn gây bệnh, xạ khuẩn nội sinh trên cây ngập mặn còn là nguồn cung cấp các chất gây độc tế bào ung thư ứng dụng trong điều trị lâm sàng và dược phẩm. Dịch lên men hai chủng LCP02 và LCP14 được tách chiết bằng dung môi ethyl acetat và cô bằng hệ thống cô quay chân không. Mẫu cao chiết đã được đánh giá hoạt tính gây độc tế bào ung thư đối với 3 dòng tế bào, gồm tế bào ung thư phổi A549, ung thư vú MCF7 và ung thư gan Hep3B. Kết quả thể hiện ở Bảng 3.1.

**Bảng 3.1. Hoạt tính gây độc tế bào ung thư của 2 chủng xạ khuẩn nội sinh với 3 loại tế bào ung thư thử nghiệm.**

Tên chủng	Nồng độ cao chiết thô	Tỷ lệ tế bào sống sót (%)		
		A549	MCF7	Hep3B
LCP02	100 µg/ml	28,3±1,9	20,2±1,1	17,9±1,5
LCP14		15,6±1,1	24,9±4,7	23,9±4,2
LCP02	30 µg/ml	33,4±0,4	31,4±2,2	42,4±1,8
LCP14		39,7±2,1	41,9±3,3	58,8±0,9
Camptothecin (10 µM)		41,7±1,3	41,9±2,9	28,3±2,6

*Ghi chú: Đối chứng dương là camptothecin.*

Cao chiết chủng LCP02 có tác dụng ức chế đáng kể ở cả hai nồng độ sử dụng với khả năng sống sót của 3 dòng tế bào dao động từ 17,9 - 42,4% (Bảng 3.1). Cụ thể, 30 µg/ml cao chiết thô thể hiện hoạt tính gây độc tế bào mạnh đối với MCF7 và A549 với khả năng sống sót của tế bào được ghi nhận lần lượt là 31,4±2,2% và 33,4±0,4%. Tăng nồng độ dịch chiết lên 100 µg/ml làm tăng đáng kể hoạt tính gây độc tế bào đối với cả 3 dòng tế bào thử nghiệm. Tương tự, cao chiết thô LCP14 cũng thể hiện hoạt tính ức chế đối với 3 dòng tế bào ung thư. Tại nồng độ 30 µg/ml, khả năng ức chế trên dòng ung thư gan Hep3B và ung thư vú MCF7 đều không cao. Khi tăng nồng độ lên 100 µg/mL, khả năng sống sót của tế bào chỉ còn lần lượt 15,6±1,1%, 23,9±4,2% và 24,9±4,7% với tế bào A549, Hep3B và MCF7, tương ứng (Bảng 3.1).

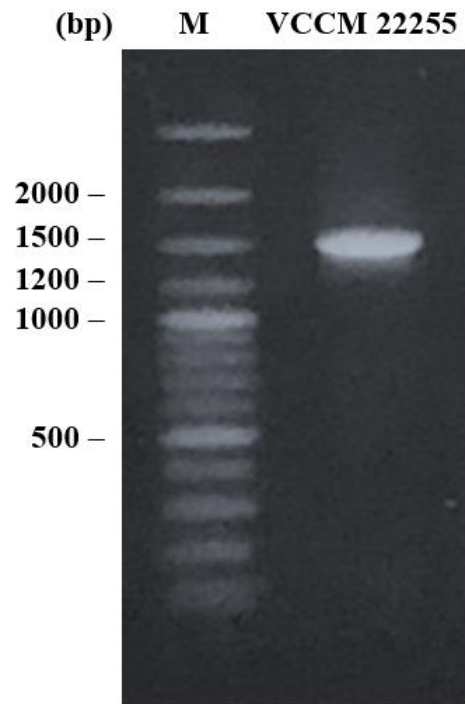
Nhiều loài xạ khuẩn nội sinh được biết đến là nguồn sản xuất các hợp chất có hoạt tính y dược, đặc biệt là khả năng ức chế tế bào ung thư. Năm 2019, nghiên cứu của Janaki đã sử dụng 200 µg/mL dịch chiết thô của chủng

*S. cacaoi* M20 phân lập từ đất rừng ngập mặn thể hiện khả năng ức chế tăng trưởng 49% tế bào ung thư MCF7 [79]. Trong tổng số 119 chủng xạ khuẩn phân lập từ các cây dược liệu bản địa Trung Quốc, có 80 chủng có hoạt tính gây độc tế bào đối với ung thư HepG2, trong đó 26 chủng (32,5%) thể hiện hoạt động chống ung thư mạnh, phần lớn thuộc chi *Streptomyces* [77]. Một alkaloid là flavensomycinoic acid từ chủng *Streptomyces* sp. YIM66017 nội sinh trên cây *Alpinia oxyphylla*, đã được phát hiện có hoạt tính gây độc tế bào mạnh đối với dòng tế bào ung thư vú MCF7 với giá trị  $IC_{50}$  là 17,0 mM [80]. Năm 2015, Savi và cộng sự đã công bố chất 1-vinyl-b-carboline-3-carboxylic acid có hoạt tính gây độc tế bào ung thư tuyến tiền liệt PC3 và dòng tế bào ung thư biểu mô phổi A549 với các giá trị  $IC_{50}$  lần lượt là 9,45 và 24,67 mM [81]. Chủng XKNS *Streptomyces* sp. neau-D50 được chứng minh sinh dẫn xuất indole prenyl hóa mới là 3-acetylidene-7-prenylindolin-2-one có tác dụng gây độc tế bào A549 với giá trị  $IC_{50}$  là 3,3 mg/ml [82]. Hiện nay, rất nhiều chất chống u, tế bào ung thư được phát hiện từ xạ khuẩn nội sinh trên cây dược liệu nhưng số lượng được điều chế thành thuốc đặc trị vẫn còn hạn chế.

Như vậy, với khả năng ức chế rõ rệt sự phát triển của các dòng tế bào ung thư và vi khuẩn, chủng xạ khuẩn LCP02 có tiềm năng khai thác hoạt chất y dược và được lựa chọn cho những nghiên cứu tiếp theo. Quan sát hình thái khuẩn lạc cho thấy chủng LCP02 thuộc chi *Streptomyces* (Phụ lục 2). Đồng thời, chủng xạ khuẩn LCP02 được ký gửi ở Bộ sưu tập giống của Trung tâm Giống và Bảo tồn nguồn gen Vi sinh vật, Viện Công nghệ sinh học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam dưới mã số VCCM 22255. Do vậy, tên *Streptomyces* sp. VCCM 22255 được sử dụng cho các nghiên cứu tiếp theo.

### **3.3. PHÂN LOẠI CHỦNG VCCM 22255 DỰA TRÊN PHÂN TÍCH TRÌNH TỰ GEN 16S rRNA**

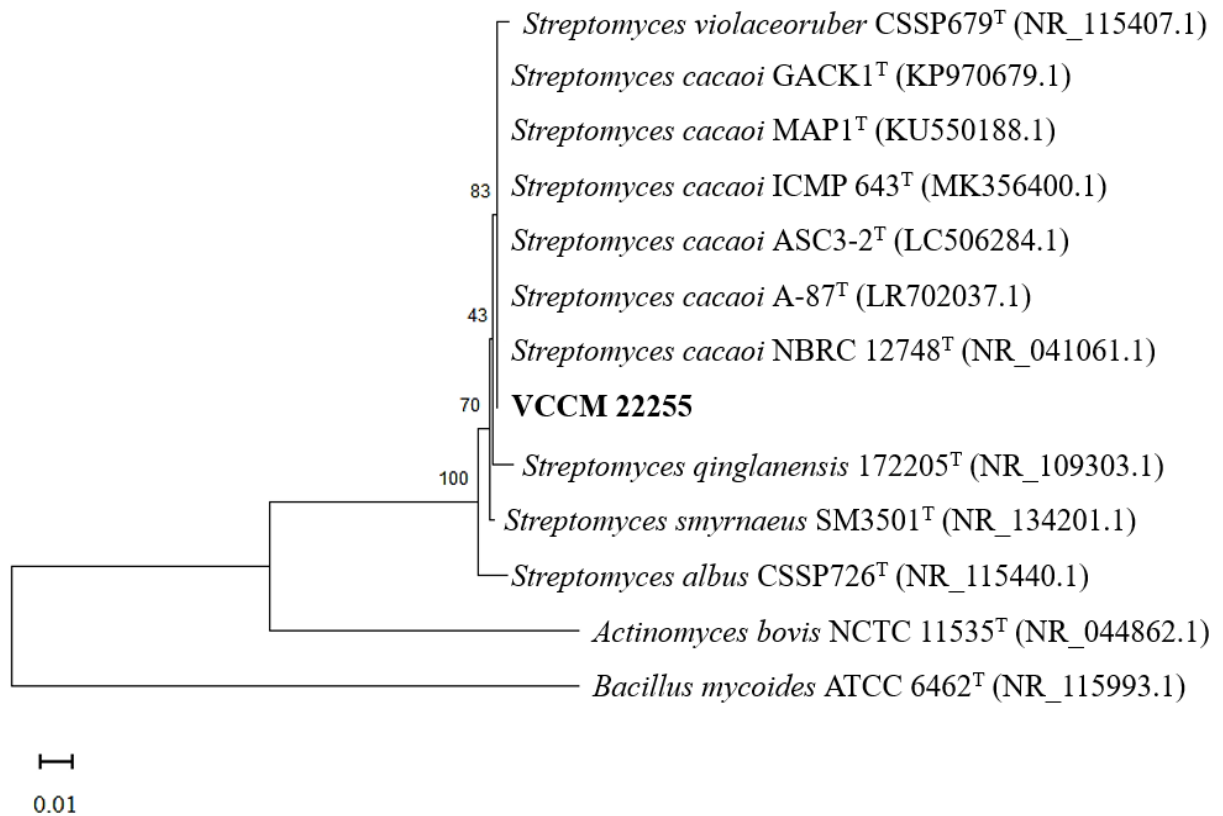
Để phân loại tới loài, chủng xạ khuẩn VCCM 22255 được tách chiết DNA tổng số và khuếch đại gen 16S rRNA bằng phản ứng PCR (Hình 3.2). Kết quả điện di trên gel agarose 1% thu được một băng DNA sáng, rõ nét và không tạo vệt, với kích thước khoảng 1500 bp. Sản phẩm PCR được tinh sạch và gửi giải trình tự tại Công ty First BASE Laboratories Sdn. Bhd. (Malaysia).



**Hình 3.2. Điện di đồ sản phẩm PCR khuếch đại gen 16S rRNA của chủng VCCM 22255 trên trên gel agarose 1,0%.**

*Ghi chú: M: Thang DNA chuẩn (DNA marker).*

Trình tự gen 16S rRNA của chủng xạ khuẩn VCCM 22255 được trình bày trong Phụ lục 1. Dựa trên công cụ BLAST so sánh với các trình tự gen công bố trên cơ sở dữ liệu GenBank cho thấy trình tự gen 16S rRNA của chủng VCCM có độ tương đồng cao với trình tự 16S rRNA của các chủng *Streptomyces cacaoi* NBRC 12748<sup>T</sup> (100%), *S. cacaoi* A-87<sup>T</sup> (100%), *S. cacaoi* ASC3-2<sup>T</sup> (100%), *S. cacaoi* ICMP 643<sup>T</sup> (100%), *S. cacaoi* MAP1<sup>T</sup> (100%) và *S. cacaoi* GACK1<sup>T</sup> (100%) (Phụ lục 3). Phân tích cây phát sinh chủng loại cũng cho thấy chủng xạ khuẩn VCCM 22255 tạo thành một nhánh riêng biệt với các chủng tham chiếu *S. cacaoi* với giá trị bootstrap là 83% (Hình 3.3). Như vậy, chủng VCCM 22255 được định danh là *S. cacaoi* VCCM 22255.



**Hình 3.3. Cây phát sinh chủng loại dựa trên phân tích trình tự gen 16S rRNA của chủng VCCM 22255 và các chủng xạ khuẩn tham chiếu đại diện.**

*S. cacaoi* lần đầu được phát hiện và phân lập từ hạt cacao, do Waksman mô tả ban đầu vào năm 1932, và chỉnh sửa lần cuối bởi Lanoot và cộng sự năm 2002 [83]. Xạ khuẩn nội sinh trên cây cacao *S. cacaoi* subsp. *cacaoi* NBRC 12748<sup>T</sup> được chứng minh sinh chất kháng khuẩn pentaminomycin C [84]. Năm 1965, Isono và cộng sự đã tách chiết được polyoxins B và polyoxins D có hoạt tính diệt nấm từ cao chiết thô của chủng *Streptomyces cacaoi* var. *asoensis*. Cơ chế kháng nấm của polyoxins B và D nhờ hoạt tính ức chế sinh tổng hợp chitin, một chất cần thiết trong thành tế bào nấm [85]. Bên cạnh đó, các chất polyether K41A và inophores dạng polyether với khả năng gây độc tế bào ung thư A549 đã được tách chiết từ một chủng *S. cacaoi* từ biển [86].

Gần đây, chủng *S. cacaoi* AM1 phân lập từ cây mắm (*Avicennia marina*) được chứng minh có tác dụng ức chế mạnh đối với 6 loại vi sinh vật kiểm định với vòng vô khuẩn từ 7,5 - 22,3 mm. Đồng thời cao chiết etyl axetat của *S. cacaoi* AM1 ở nồng độ 100 µg/ml cho thấy khả năng gây độc tế bào mạnh đối với ung thư vú MCF7 và ung thư phổi A549 với khả năng sống

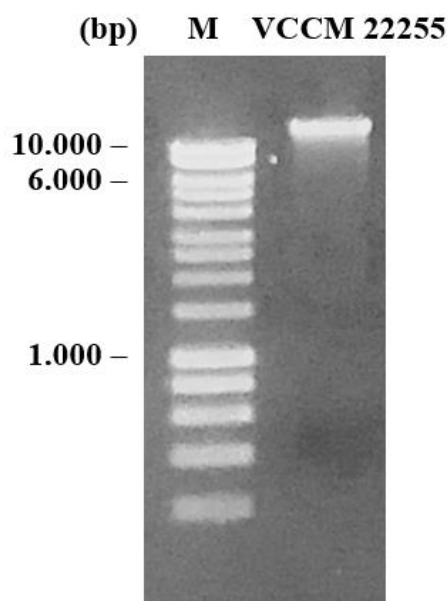


của tế bào lần lượt là  $16,5 \pm 1,28\%$  và  $17,69 \pm 2,3\%$  [87]. Một nghiên cứu trước đây nhấn mạnh khả năng gây độc tế bào ung thư của chủng *S. cacaoi* M20 có nguồn gốc từ rừng ngập mặn với tỷ lệ ức chế dòng tế bào MCF7 là 49% với nồng độ cao chiết là 200  $\mu\text{g/ml}$  [79]. So với chủng *S. cacaoi* M20 và *S. cacaoi* AM1 khả năng gây độc tế bào của chủng VCCM 22255 cao hơn đáng kể. Mặc dù *S. cacaoi* đã được phân lập từ nhiều nguồn khác nhau, tiềm năng nghiên cứu các chất kháng khuẩn và gây độc tế bào ung thư vẫn còn rất lớn.

### 3.4. GIẢI TRÌNH TỰ HỆ GEN VÀ PHÂN TÍCH MỐI LIÊN QUAN GIỮA ĐẶC TÍNH DI TRUYỀN VÀ ĐẶC TÍNH KHÁNG SINH, GÂY ĐỘC TẾ BÀO CỦA CHỦNG VCCM 22255

#### 3.4.1. Giải trình tự hệ gen chủng *S. cacaoi* VCCM 22255

Chất lượng DNA tổng số đóng vai trò quan trọng trong giải trình tự hệ gen. Các mẫu DNA không đạt tiêu chuẩn được loại bỏ sau quá trình kiểm tra chất lượng. Do vậy, chủng *S. cacaoi* VCCM 22255 được tách chiết DNA tổng số bằng bộ kit tách hệ gen của xạ khuẩn PureLink® Genomic DNA Kit. Đầu tiên, DNA tổng số trên gel agarose 0,8% cho một băng đậm, sáng rõ nét, ít đứt gãy và không chứa RNA. Sau đó, nồng độ DNA được kiểm tra trên máy NanoDrop. Mẫu DNA có nồng độ đạt 43 ng/ $\mu\text{l}$  với độ tinh sạch  $\text{OD}_{260/280}$  dao động từ 1,9 - 2,0. Kết quả nồng độ và chất lượng DNA tổng số cho thấy mẫu đủ điều kiện để giải trình tự trên hệ thống giải trình tự Illumina.



**Hình 3.4. Điện di đồ DNA tổng số của chủng *S. cacaoi* VCCM 22255.**

*Ghi chú: M: Thang DNA chuẩn (DNA marker).*

Công nghệ Illumina là công nghệ giải trình tự đoạn ngắn, cho phép giải trình tự toàn bộ hệ gen của xạ khuẩn. Tới nay, hàng trăm hệ gen tham chiếu của xạ khuẩn đã được công bố, cung cấp thông tin cơ bản về trình tự hệ gen của các chủng trong cùng 1 chi hoặc loài, cho thấy sự biến đổi của hệ gen theo tiến hoá và tiềm năng sinh sản phẩm thứ cấp. Từ DNA tổng số, thư viện được chuẩn bị bằng kit theo hướng dẫn của Illumina. Mẫu VCCM 22255 được giải trình tự và dữ liệu thô được thu nhận. Chủng 22255 có 6.207.812 trình tự (read) và tổng số base là 626.443.311 (Bảng 3.2). Tỷ lệ Q30 và GC đạt lần lượt là 99,3% và 73,3%. Kết quả cho thấy, số liệu thô của mẫu khá tốt và đủ để tiến hành bước lắp ghép *de novo*.

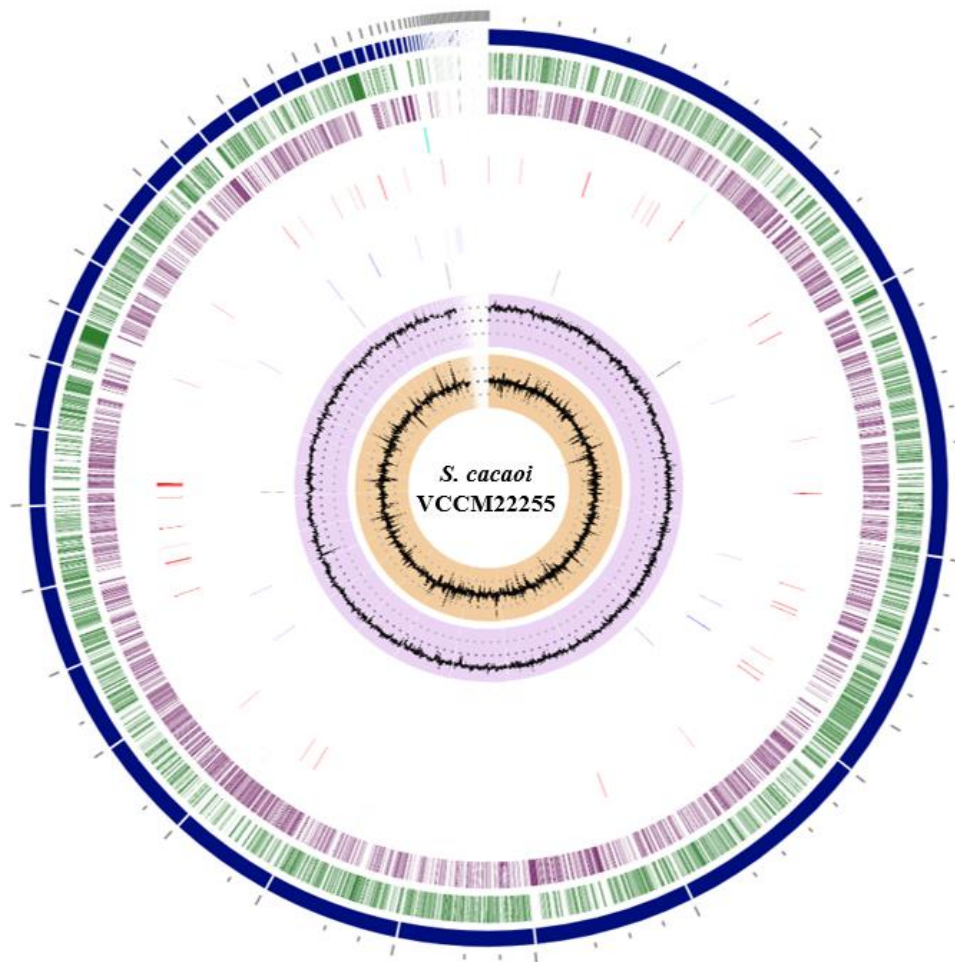
**Bảng 3.2. Dữ liệu thô từ mẫu giải trình tự *S. cacaoi* VCCM 22255.**

Chỉ tiêu phân tích	<i>S. cacaoi</i> VCCM 22255
Tổng số base	626.443.311
Tổng số read	6.207.812
GC (%)	73,3
Q20 (%)	99,4
Q30 (%)	99,3

*Ghi chú: Tỷ lệ phần trăm GC được xác định theo công thức  $(G+C)/(A+T+G+C)*100%$ ; Q20 (%): Tỷ lệ các base có chất lượng phred trên 20; Q30 (%): Tỷ lệ các base có chất lượng phred trên 30.*

### 3.4.2. Lắp ghép *de novo* và chú giải hệ gen VCCM 22255

Kết quả lắp ghép *de novo* và chú giải hệ gen cho thấy hệ gen của chủng *S. cacaoi* VCCM 22255 dạng thẳng, có kích thước là 8,44 Mb và hàm lượng GC là 73,3% (Hình 3.5). Tổng số có 63 RNA trong đó có 3 rRNA và 60 tRNA. Trong đó, tổng số trình tự gen mã hóa protein là 7359. Kết quả đánh giá độ hoàn thiện của hệ gen bằng BUSCO cho thấy hệ gen có độ hoàn thiện đạt 98,6%.



**Hình 3.5. Bản đồ hệ gen chủng *S. cacaoi* VCCM 22255.**

Từ ngoài vào: Thứ tự các contig (thể hiện bằng màu xanh nước biển); các trình tự mã hóa theo chuỗi xuôi và ngược (trương ứng là màu xanh lá cây và màu tím); các yếu tố không mã hóa dọc theo nhiễm sắc thể (màu xanh lam); các gen liên quan đến kháng kháng sinh (màu đỏ); các gen độc lực khác (màu cam); các gen mã hóa protein xuyên màng (màu xanh đậm); các gen mã hóa các mục tiêu thuốc (màu đen); hàm lượng GC dọc theo các chuỗi xuôi và ngược (hai vòng tròn ở tâm).

Để có cái nhìn rõ hơn về đặc điểm hệ gen của chủng xạ khuẩn nghiên cứu, hệ gen *S. cacaoi* VCCM 22255 được so sánh với hệ gen các chủng *S. cacaoi* SCSIO 68063, *S. avertimilis* ATCC 31267, *S. griseus* IFO 13350, *S. coelicolor* A3(2), *S. fulvissimus* DSM40593, *S. albus* J1074. Hiện nay chỉ có duy nhất trình tự hệ gen chủng *S. cacaoi* SCSIO 68063 được công bố trên GenBank (NCBI). Hệ gen chủng VCCM 22255 (8,44 Mb) có kích thước nhỏ hơn *S. cacaoi* SCSIO 68063 (8,50). Hàm lượng GC khá tương đồng, đạt 73%. Tuy nhiên, số lượng protein được dự đoán ở *S. cacaoi* SCSIO 68063 thấp hơn *S. cacaoi* VCCM 22255 là 535. Số lượng rRNA giữa các chủng so sánh có sự

khác biệt, một đặc điểm quen thuộc của xạ khuẩn Nhiều copy của rRNA có thể có ý nghĩa trong quá trình thích nghi của xạ khuẩn trước điều kiện thay đổi về chất dinh dưỡng hoặc nhiệt độ tuy nhiên việc này không phải thiết yếu để duy trì sự sống của vi sinh vật [88].

Hệ gen xạ khuẩn khá lớn, dao động từ 5,9 – 12,0 Mb. Chủng xạ khuẩn có kích thước lớn nhất được giải trình tự là *S. hygroscopicus* XM201 [45,46]. Nhìn chung, kích thước hệ gen *S. cacaoi* VCCM 22255 đạt mức trung bình với 8,44 Mb (Bảng 3.3). Một số hệ gen có kích thước thấp hơn như *S. fulvissimus* DSM40593 và *S. albus* J1074. Đáng chú ý, số lượng gen mã hóa protein không hoàn toàn tuyến tính với kích thước hệ gen. Nguyên nhân do nhiều gen mới chưa được dự đoán chức năng.

**Bảng 3.3. So sánh đặc điểm hệ gen của *S. cacaoi* VCCM 22255 với các hệ gen *Streptomyces* khác.**

STT	Tên chủng	Kích thước (Mb)	GC (%)	Gen mã hóa protein	tRNA	rRNA
1	<i>S. cacaoi</i> VCCM 22255	8,44	73,3	7359	60	3
2	<i>S. cacaoi</i> DSM 40057	8,39	73,3	6316	61	12
3	<i>S. avertimilis</i> ATCC 31267	9,03	70,7	7583	68	6
4	<i>S. griseus</i> IFO 13350	8,55	72,2	7138	66	18
5	<i>S. coelicolor</i> A3(2)	8,67	72,2	7825	63	6
6	<i>S. fulvissimus</i> DSM40593	7,91	71,5	6729	73	18
7	<i>S. albus</i> DSM 40763	8,10	72,5	6562	61	19

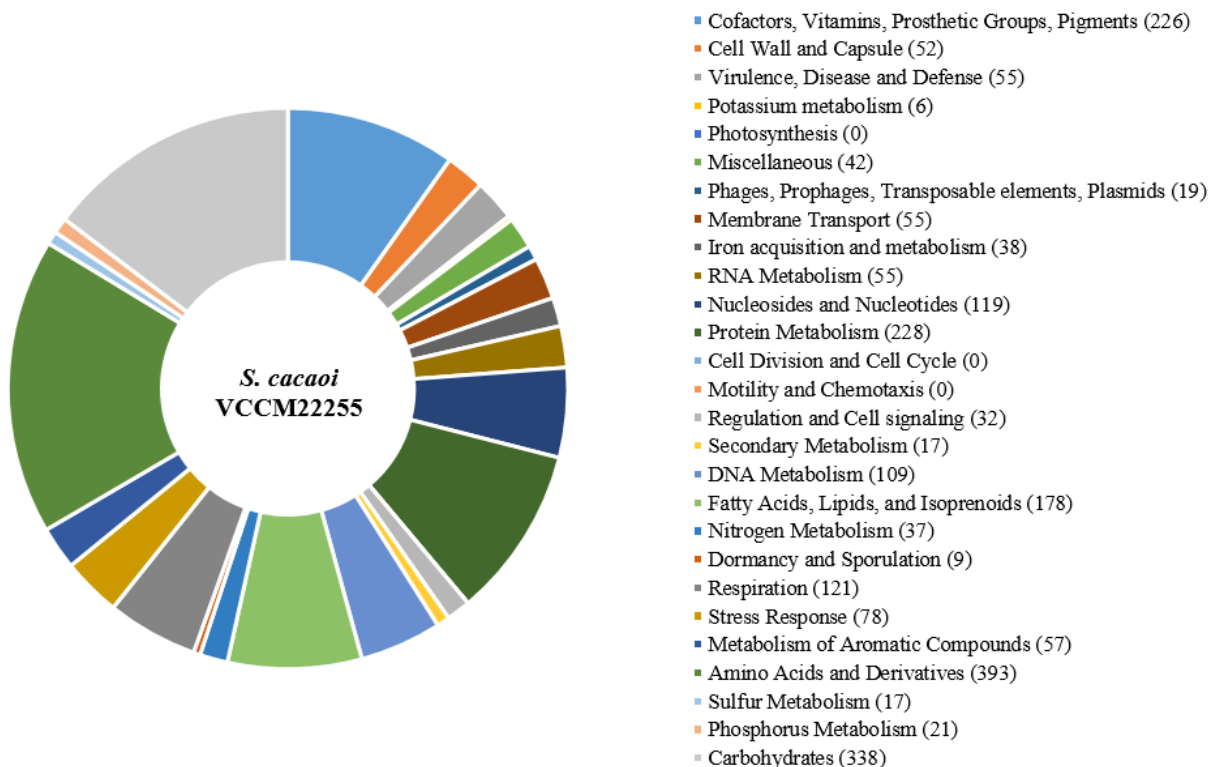
### 3.4.3. Phân tích đặc điểm di truyền hệ gen VCCM 22255

#### 3.4.3.1. Phân nhóm các gen chức năng

Các gen mã hoá protein chức năng là chỉ tiêu đầu tiên để đánh giá đặc điểm di truyền hệ gen xạ khuẩn. Cơ sở dữ liệu RAST đã phân loại protein trong hệ gen VCCM 22255 thành 27 nhóm chức năng chứa 324 nhóm nhỏ (subsystem). Trong đó, nhóm “Amino acids và dẫn xuất” và “Carbohydrates” chiếm số lượng gen cao nhất với 393 và 338 gen, tương ứng. Điều này cũng gợi ý rằng các chủng xạ khuẩn nội sinh từ cây màng tang có khả năng sử dụng nhiều nguồn dinh dưỡng từ carbohydrates, amino acids và protein có sẵn trong môi trường sống. Bốn nhóm có số lượng thấp hơn là “Protein

metabolism” (228 gen); “Cofactors, vitamins, prosthetic groups, pigments” (226 gen); Fatty acids, lipids, isoprenoids (178 gen); “Nucleosides and Nucleotides” (119 gen) (Hình 3.5). Nhóm có số lượng gen dự đoán thấp nhất là “Potassium metabolism” với 6 gen, phù hợp với đặc điểm của *Streptomyces* vì khả năng chuyển hoá kali trong tế bào thường thấp.

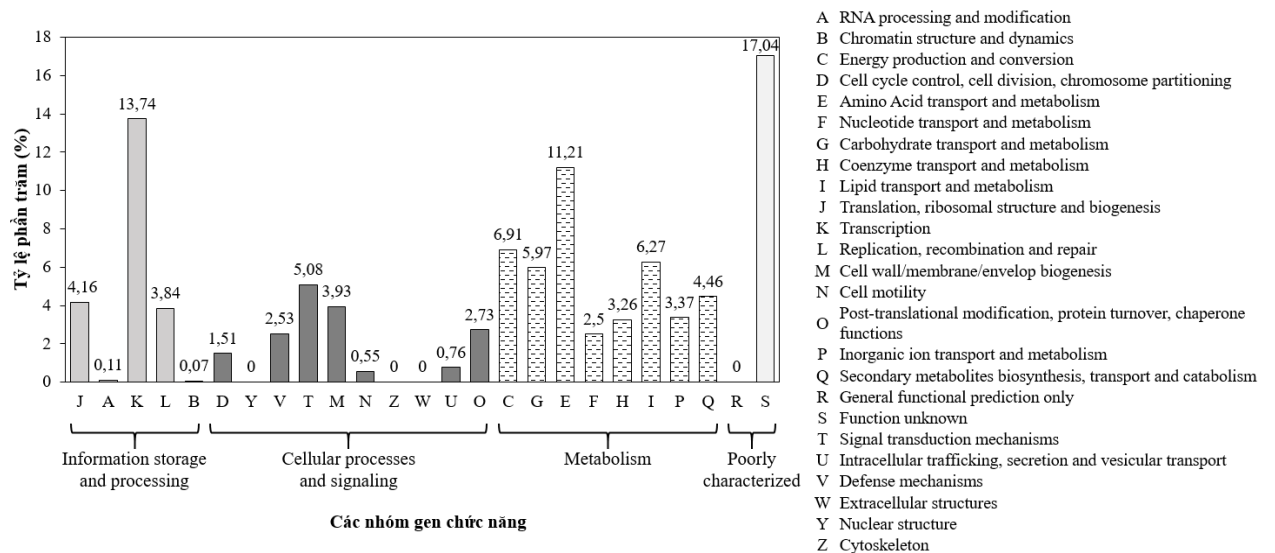
Đặc điểm đáng chú ý là nhóm “Stress response” (78 gen) đóng vai trò quan trọng với sự sinh trưởng và phát triển của xạ khuẩn trong các môi trường khác nhau. Xạ khuẩn nội sinh thường chịu các loại stress khác nhau từ cây chủ như pH và nhiệt độ thay đổi, ánh sáng, hợp chất vòng thơm, hoạt chất trong cây, đặc biệt là các cây dược liệu. Các stress này là nguyên nhân dẫn đến sự hình thành các gốc oxi hoá tự do (Reactive oxygen species). Việc sản sinh quá mức ROS dẫn đến nguy cơ phá huỷ phân tử sinh học từ DNA, RNA và lipid, dẫn đến phá huỷ tế bào. Các chủng xạ khuẩn phân lập từ đất thường có số lượng gen trong nhóm “Stress response” thấp như *S. coelicolor* (16 gen), *S. avermitilis* (17 gen), *S. griseus* (15 gen) [89].



**Hình 3.6. Dự đoán các nhóm chức năng gen trong hệ gen *S. cacaoi* VCCM 22255.**

Bên cạnh cơ sở dữ liệu RASK, hệ gen chủng *S. cacaoi* VCCM được phân tích bổ sung bằng cơ sở dữ liệu EggNOG. Hệ gen chủng *S. cacaoi* VCCM 22255 được dự đoán có 5646 gen chia thành 4 nhóm chức năng chính,

gồm: (1) information storage and processing (xử lý và lưu trữ thông tin) (1238 gen, tương ứng chiếm 21,93%), (2) cellular processes and signaling (tín hiệu và quá trình tế bào) (965 gen, tương ứng 17,09%), (3) metabolism (trao đổi chất) (2481 gen, tương ứng 43,94%), và (4) poorly characterized (các chức năng khác) (962 gen, tương ứng 17,04%) (Hình 3.7). Trong đó, tỷ lệ gen không được xác định chức năng chiếm tỷ lệ cao nhất với 17,04%. Khác biệt, xạ khuẩn *S. spectabilis* KCTC9218<sup>T</sup> and *Streptomyces* sp. AN091965 phân lập từ đất có tỷ lệ gen không xác định chức năng đạt lần lượt 37,92% và 38,13 [90]. Tiếp theo, nhóm “Transcription” (13,74%) và “Amino acid transport and metabolism” (12,21%) chiếm tỷ lệ thấp hơn. Tỷ lệ gen trong nhóm “Amino acid transport and metabolism” từ *S. spectabilis* KCTC9218<sup>T</sup> and *Streptomyces* sp. AN091965 thấp hơn một nửa so với *S. cacaoi* VCCM 22255. Sự phân bố của các nhóm gen trong 2 chủng trên khá tương đồng với các chủng *Streptomyces* khác, tuy nhiên chủng *S. cacaoi* VCCM 22255 tương đối khác biệt. Điều này khẳng định có nhiều sự thay đổi trong hệ gen chủng *S. cacaoi* VCCM 22255 cần được nghiên cứu.



**Hình 3.7. Tỷ lệ các gen chức năng của chủng *S. cacaoi* VCCM 22255 dựa trên cơ sở dữ liệu EggNOG.**

### 3.4.3.2. Phân nhóm enzym trong hệ gen

Enzym đóng vai trò quan trọng trong hoạt động trao đổi chất, sinh trưởng của xạ khuẩn. Chính vì vậy, phân tích hệ enzyme hoạt động với carbohydrate (Carbohydrate-Active enZymes, CAZyme) của chủng *S. cacaoi* VCCM 22255 được tiến hành. Kết quả cho thấy, số lượng CAZyme được dự đoán công cụ h-mmer và diamond là 128 enzym (Phụ lục 4). Trong đó, nhóm có số lượng enzym dự đoán nhiều nhất thuộc về nhóm GH (Glycoside Hydrolase) (56 enzym) và GT (Glycosyl Transferase) (55 enzym) (Hình 3.8). Các nhóm còn lại chiếm số lượng thấp hơn, gồm AA (Auxiliary Activitie) (3 enzym), PL (Polysaccharide Lyase) (3 enzym), CE (Carbohydrate Esterase) (7 enzym) và CBM (Carbohydrate-Binding Module) (3 enzym). Các CAZymes chịu trách nhiệm tổng hợp và phân hủy glycoconjugates, oligosacarit và polysacarit, tương ứng với 1-5% gen của một sinh vật sống [76]. Glycoconjugates, oligosacarit và polysacarit không chỉ đóng vai trò là dự trữ năng lượng và cấu trúc mà còn liên quan tới nhận dạng trong và ngoài tế bào.

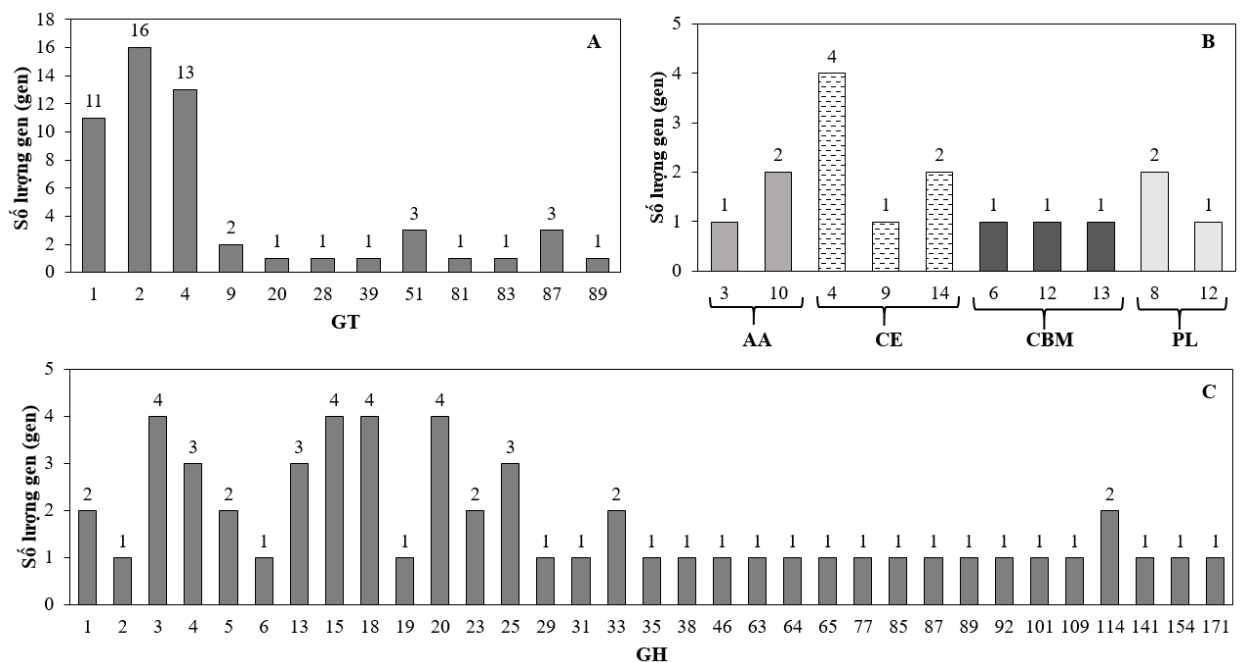
Quá trình sinh tổng hợp disacarit, oligosacarit và polysacarit liên quan đến hoạt động của hàng trăm GT khác nhau (EC 2.4.x.y), các enzyme xúc tác quá trình chuyển các gốc đường từ các phân tử chất cho, được kích hoạt sang các phân tử chất nhận cụ thể, hình thành các liên kết glycosid. Trong số 116 họ GT, đã tìm được 54 loại enzyme GT thuộc 12 họ GT khác nhau có trong chủng *S. cacaoi* VCCM 22255. Chiếm đa số là GT1 (11 enzym, tương ứng chiếm 20%), GT2 (16 enzym, chiếm 29,09%) và GT4 (13 enzym, chiếm 23,64%). Các họ còn lại gồm GT9, 20, 28, 39, 51, 81, 83, 87 và 89 chỉ có từ 1 – 3 enzym (Hình 3.8A). Bên cạnh đó, nhóm hoạt động phụ trợ (AA) bao gồm các enzyme oxi hóa khử hoạt động cùng với các CAZyme, trong số 17 họ AA, từ hệ gen chủng *S. cacaoi* VCCM 22255 tìm thấy 2 họ là AA3 (1 enzym) và AA10 (2 enzym) (Hình 3.8B).

GH là một nhóm phổ biến các enzyme thủy phân liên kết glycosid giữa hai hoặc nhiều carbohydrate hoặc giữa một carbohydrate và một nửa không

phải carbohydrate. Trong số 180 họ GH, đã tìm được 56 loại enzyme GH thuộc 33 họ GH khác nhau có trong chủng *S. cacaoi* VCCM 22255. Chiếm đa số là 4 họ GH3, GH15, GH18, GH20 (4 enzyme, tương ứng chiếm 7,41%), tiếp đó là 3 họ GH4, GH 13, GH25 (3 enzyme, chiếm 5,56%), các họ còn lại chỉ có từ 1 – 2 enzyme (Hình 3.8C).

Nhóm CE xúc tác quá trình khử-O hoặc khử-N-acyl hóa các sacarit thay thế. Hệ gen chủng *S. cacaoi* VCCM 22255 chứa CE4 (4 enzym), CE9 (1 enzym) và CE14 (2 enzym) (Hình 3.8B). CBM là trình tự axit amin liền kề trong enzyme hoạt động với carbohydrate với nếp gấp kín có hoạt tính liên kết với carbohydrate. Trong cơ sở dữ liệu CAZyme, 97 CBM khác nhau được ghi nhận. Hệ gen VCCM 22255 chỉ có CBM6, CBM12 và CBM13 (Hình 3.8B).

PL là một nhóm enzyme tách chuỗi polysacarit chứa axit uronic thông qua cơ chế loại bỏ  $\beta$  để tạo ra gốc axit hexenuronic không bão hòa và một đầu khử mới. PL thường chỉ chiếm số lượng nhỏ trong hệ gen vi khuẩn. PL8 (2 enzym) và PL2 (1 enzym) được tìm thấy trong VCCM 22255. (Hình 3.8B).



**Hình 3.8. Số lượng các nhóm gen CAZyme của chủng *S. cacaoi* VCCM 22255.**



### 3.4.4. Các cụm gen sinh tổng hợp sản phẩm bậc hai của chủng *S. cacaoi* VCCM 22255

Qua phân tích các cụm gen sinh tổng hợp smBGC bằng công cụ AntiSMASH, có tổng cộng 31 cụm gen mã hóa đã được dự đoán, trong đó 17 cụm gen có độ tương đồng trên 20% với cụm gen tham chiếu, bao gồm các cụm gen trong nhóm terpene (4 cụm gen); non-ribosomal peptide synthetases (NRPS) (3 cụm gen); lantipeptides (1 cụm gen); type II polyketide synthase (T1PKS) (3 cụm gen); ectoine (1 cụm gen); butyrolactone (1 cụm gen); butyrolactone (1 cụm gen); prodigiosin (1 cụm gen); lassopeptide (1 cụm gen) và cụm gen khác (1 cụm gen) (Bảng 3.4). Kết quả cho thấy, các smBGC của VCCM 22255 có độ tương đồng 100% với các cụm gen sinh tổng hợp ectoine, geosmin, desferrioxamine E và A-factor. Các cụm gen có độ tương đồng từ 61 – 96% bao gồm hopene (61%), ibomycin (63%), aborycin (64%), dudomycin A (65%), pseudouridimycin (68%), sapB (75%), xantholipin (81%), tambjamine BE-18591 (96%). Với các cụm gen có độ tương đồng > 68% bao gồm ectoine, geosmin, desferrioxamine E, sapB, xantholipin chắc chắn được sinh ra trong môi trường nuôi cấy. Điều này phù hợp với các nghiên cứu trước đây chứng minh cụm gen ectoine, geosmin và sapB có độ bảo thủ cao trong chi *Streptomyces* [91].

Tổng quan nghiên cứu 17 smBGC mã hoá các chất có hoạt tính sinh học cho thấy 10 cụm gen có một hoặc một vài hoạt tính quan trọng như gây độc tế bào ung thư, kháng khuẩn, kháng nấm (Bảng 3.4). Các cụm gen tiềm năng trên được phân nhóm theo hoạt tính sinh học và được thể hiện ở Hình 3.9 – 3.11.

**Bảng 3.4. Các cụm gen sinh tổng hợp được dự đoán trong hệ gen *S. cacaoi* VCCM 22255.**

Cụm gen	Kích thước (bp)	Loại cụm gen	Cụm gen tham chiếu	Hoạt tính/Chức năng	Độ tương đồng (%)	Chủng tham chiếu
1.1	10.405	Ectoine	Ectoine	Chất thẩm thấu	100	<i>Streptomyces</i> sp.
1.2	117.340	Terpene	Xantholipin	Chống ung thư, kháng khuẩn	81	<i>S. flavogriseus</i>
1.3	22.232	Terpene	Geosmin	Liên quan đến tạo bào tử	100	<i>S. coelicolor</i> A3(2)
2.1	37.951	T1PKS	Selvamicin	Kháng nấm	20	<i>Pseudonocardia</i> sp. HH130630-07
2.2	26.705	Terpene	Hopene	Protein màng	61	<i>S. coelicolor</i> A3(2)
4.4	24.462	Terpene	Isorenieratene	Sắc tố	45	<i>S. collinus</i> Tu 365
5.1	48.587	NRPS	Griseobactin	Chống ung thư	53	<i>Streptomyces</i> sp. ATCC 700974
5.2	85.601	T1PKS	Ibomycin	Kháng nấm	63	<i>Streptomyces</i> sp. WAC2288
8.1	35.008	Prodigiosin	Tambjamine BE-18591	Chống ung thư, kháng khuẩn, kháng nấm	96	<i>S. albus</i> NRRL B-2362
9.2	8.811	NRPS	Desferrioxamine E	Vận chuyển sắt, kháng ung thư	100	<i>Streptomyces</i> sp. ID38640

Cụm gen	Kích thước (bp)	Loại cụm gen	Cụm gen tham chiếu	Hoạt tính/Chức năng	Độ tương đồng (%)	Chủng tham chiếu
10.1	22.715	Lanthipeptide	SapB	Peptide hình thái	75	<i>S. coelicolor</i> A3(2)
10.2	41.239	Nucleoside	Pseudouridimycin	Kháng khuẩn	68	<i>Streptomyces</i> sp. ID38640
14.3	69.248	NRPS	Thiazostatin	Kháng khuẩn	46	<i>S. venezuelae</i> ATCC 10712
17.1	19.265	Lasso peptide	Aborycin	Kháng khuẩn	64	<i>Streptomyces</i> sp. ZS0098
22.1	44.074	NRPS	Dudomycin A	Không xác định	65	<i>S. albus</i> subsp. chlorinus
24.1	10.345	Butyrolactone	A-factor	Protein điều hòa	100	<i>S. griseus</i> subsp. griseus NBRC 13350
33.1	12.549	T1PKS	Selvamicin	Kháng nấm	20	<i>Pseudonocardia</i> sp. HH130630-07

#### 3.4.4.1. Cụm gen sinh tổng hợp chất gây độc tế bào ung thư

Các smBGC sinh tổng hợp chất gây độc tế bào ung thư là cụm gen số 1.2 (xantholipin), 5.1 (griseobactin), 8.1 (tambjamine BE-18591) và 9.2 (desferrioxamine E).

Desferrioxamine E không chỉ có chức năng vận chuyển sắt vào tế bào, mà còn có hoạt tính chống ung thư. Khi đối chiếu với chủng tham chiếu *Streptomyces* sp. ID38640, cụm gen 9.2 từ *S. cacaoi* VCCM 22255 có độ tương đồng lên tới 100% với smBGC sinh desferrioxamine E. Cụm gen này có kích thước nhỏ (8.811 bp) gồm các gen đặc trưng *desABCD*, trong đó siderophore synthetase *desD* chịu trách nhiệm chính cho sinh tổng hợp desferrioxamine E (Hình 3.9). Tế bào ung thư được biết đến với nhu cầu sử dụng sắt cao, trong khi đó desferrioxamine cạnh tranh sắt với tế bào ung thư từ đó hạn chế sự phát triển của các tế bào ung thư như ung thư máu, ung thư vú, ung thư gan và ung thư buồng trứng [92]. Ngoài vai trò là iron chelator, desferrioxamine còn được cho là có ảnh hưởng đến sự biểu hiện của gen N-

myc downstream-regulated gene 1 (*ndrg1*) mã hóa cho protein ức chế quá trình di căn của tế bào ung thư [93]. Sự có mặt của cluster desferrioxamine E trong hệ gen của VCCM 22255 gợi ý khả năng sinh tổng hợp chất này và phù hợp với hoạt tính kháng ung thư đã quan sát được. Desferrioxamin E gần đây cũng được xác định trong dịch nuôi cấy chủng *S. parvulus* 03 phân lập từ cây ngập mặn *Kandelia candel* [94].

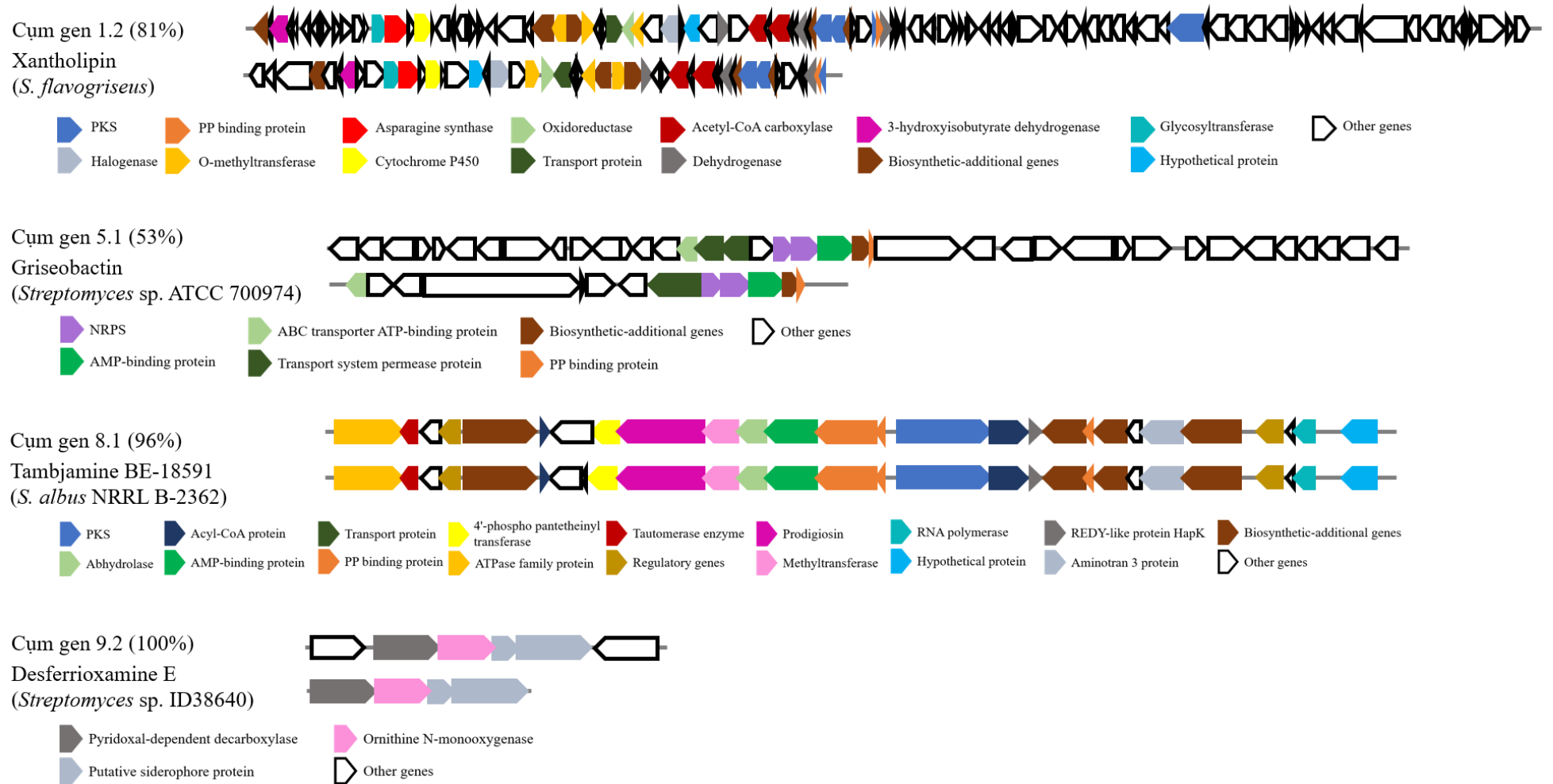
Tương tự, một cụm gen khác cũng mã hóa sinh tổng hợp siderophore có hoạt tính chống ung thư là cụm gen 5.1. Cụm gen này có kích thước 48,6 kb, gần gấp đôi so với cluster tham chiếu griseobactin (26 kb). So sánh cho thấy, cụm gen 5.1 của *S. cacaoi* VCCM 22255 có độ tương đồng 53% với cụm gen tham chiếu griseobactin ở *Streptomyces* sp. ATCC 700974. Cụm gen tham chiếu griseobactin bao gồm các gen *dhbABCEG*, có liên quan đến quá trình sinh tổng hợp 2,3-dihydroxybenzoate (DHBA) và kết hợp DHBA vào siderophores thông qua một NRPS. Bên cạnh các gen sinh tổng hợp là các gen mã hóa protein vận chuyển MFS griseobactin exporter *grisC*, hấp thu siderophore uptake membrane transporter *grisA*, griseobactin-binding lipoprotein *grisG*, ATPase of griseobactin uptake ABC transporter *grisH* và phân hủy griseobactin gồm griseobactin esterase *grisB* và amidohydrolase *grisF* [95]. Trong số các gen, chỉ có các gen *dhbABCEG*, *grisH* được xác định với tương đồng từ 51 - 74%. Mặc dù nhiều gen vẫn chưa rõ chức năng, chủng *S. cacaoi* VCCM 22255 vẫn có thể sinh ra griseobactin do chất này cũng cần thiết cho sinh trưởng của chủng (Hình 3.9).

Cụm gen số 8.1 có độ tương đồng lên tới 96% với cụm gen sinh tổng hợp tambjamine BE-18591, là một chất có phổ hoạt tính rộng với các tế bào ung thư [96]. Tambjamines và dẫn xuất thuộc lớp alkaloids gây độc tế bào ung thư nhờ hoạt tính khả năng cắt đoạn xoắn kép DNA khi có mặt phân tử oxy và ion đồng [97]. Cụm gen sinh tambjamine BE-18591 ở chủng *S. albus* NRRL B-2362 chứa 18 gen gồm *tabD*, *tabE*, *tabA*, *tabC*, *tabP*, *tabQ*, *tabR*, *tabY*, *tabW*, *tabX*, *tabO*, *tabN*, *tabM*, *tabJ*, *tabI*, *tabH*, *tabU* và *tabV* [98]. Các gen này đều có mặt ở hệ gen chủng VCCM 22255 với độ tương đồng dao động từ 74-100% (Hình 3.9). Cụm gen liên quan đến prodigiosin ở *S. coelicolor* A3(2) có phần lớn các gen tương đồng (15/18) với cụm gen tambjamine BE-18591 ở *S. albus* NRRL B-2362, tuy nhiên, hai cụm gen này

mã hóa sinh tổng hợp hai chất khác nhau [98]. Do độ tương đồng rất cao, VCCM 22255 chỉ có thể sản xuất tambjamine thay vì prodigiosin khác.

Cụm gen số 1.2 có độ tương đồng 81% với smBGC sinh xantholipin, một hợp chất vòng thơm. Mặc dù hoạt tính kháng khuẩn được kiểm chứng, hoạt tính gây độc tế bào ung thư là ưu điểm nổi trội nhất của chất này. Xantholipin được tổng hợp từ cụm gen có kích thước 52 kb chứa khung PKSII, gen điều hòa (regulator) và gen kéo dài chuỗi (tailoring polyketide) [99]. So sánh về kích thước, cụm gen 1.2 của chủng *S. cacaoi* VCCM 22255 có chiều dài gấp 2,3 lần so với xantholipin của chủng tham chiếu *S. flavogriseus* (Hình 3.9). Nhiều gen chưa được dự đoán chức năng. Có hai giả thuyết đưa ra rằng: (i) cụm gen 1.2 sinh xantholipin, (ii) xantholipin đóng vai trò là tiền chất và cấu trúc chất bị xúc tác phản ứng tạo thành chất mới.

Cho tới nay, chưa có nhiều nghiên cứu về các cụm gen sinh tổng hợp những chất có hoạt tính kháng ung thư ở loài *S. cacaoi*. Chất ung thư ansamycins thuộc lớp polyketide được sinh từ *S. cacaoi* subsp. *asoensis* H2S5, [100,101] tuy nhiên, cụm gen sinh ansamycin chưa được nghiên cứu. Do griseobactin, tambjamine BE-18591 và desferrioxamine E có hoạt tính kháng tế bào ung thư không cao, nhiều khả năng xantholipin và có thể dẫn xuất là nguyên nhân chính gây ức chế tế bào ung thư ở *S. cacaoi* VCCM 22255.



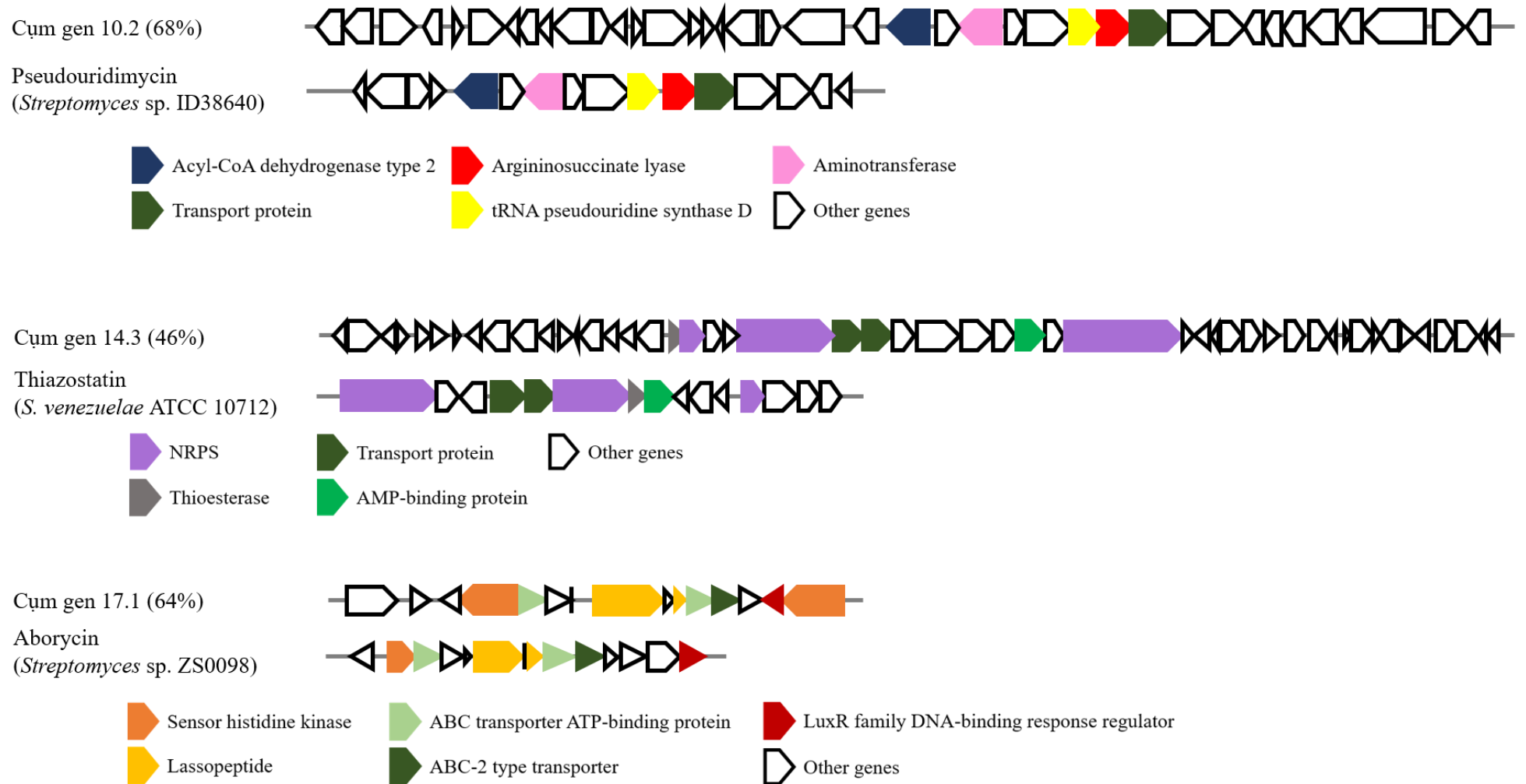
**Hình 3.9. Các cụm gen sinh tổng hợp các chất chống ung thư được xác định trong hệ gen của *S. cacaoi* VCCM 22255.**

### 3.4.4.2. Cụm gen sinh tổng hợp chất kháng khuẩn

Hình 3.10 biểu diễn độ tương đồng giữa các smBGC từ *S. cacaoi* VCCM 22255 và chủng tham chiếu. Các cụm gen liên quan đến sinh các chất kháng khuẩn gồm cụm gen 10.2 (pseudouridimycin), cụm gen 14.3 (thiazostatin) và cụm gen 17.1 (aborycin).

Cụm gen 10.2 có kích thước lớn (41kb) thể hiện độ tương đồng 68% với cụm gen tạo pseudouridimycin của chủng tham chiếu *Streptomyces* sp. ID38640 (Hình 3.10). Trong số 16 gen thuộc cụm gen tham chiếu, cluster 12 chứa 11 gen tương đồng gồm *pumEFGHIJKLMNO* với hai gen cốt lõi tRNAse pseudouridine synthase D *pumJ* và flavin-dependent oxidoreductase *pumE* có độ tương đồng lần lượt là 75% và 81%. Đặc biệt, gen *pumJ* là đóng vai trò quyết định trong con đường sinh tổng hợp pseudouridimycin [102]. Pseudouridimycin là một chất ức chế RNA polymerase của vi khuẩn Gram (+) và Gram (-). Pseudouridimycin và cụm gen sinh tổng hợp chất này được báo cáo lần đầu tiên từ hai chủng *Streptomyces* sp. ID38640 và ID38673 có nguồn gốc từ đất, và sau đó là chủng *Streptomyces albus* DSM 40763 [103]. Như vậy, *S. cacaoi* VCCM 22255 là chủng nội sinh thuộc loài *S. cacaoi* đầu tiên được báo cáo với cụm gen pseudouridimycin. Thêm vào đó, cụm gen 10.2 có hai gen mã hóa cytochrome P450 không thuộc cụm gen pseudouridimycin, gợi ý khả năng sinh tổng hợp các hợp chất khác ngoài pseudouridimycin.

Cụm gen 17.1 có độ tương đồng 64% với cluster sinh tổng hợp aborycin của *Streptomyces* sp. ZS0098, là một chất tổng hợp ribosome của các sản phẩm tự nhiên lassopeptide loại I (Hình 3.10). Cụm gen 17.1 chứa hai gen chính thuộc lassopeptide, tương tự với tham chiếu. Bên cạnh đó, theo kết quả AntiSMASH, cluster 14 giống 57% với cluster sinh tổng hợp siamycin I, một chất lassopeptide khác, do đó cụm gen 17.1 có khả năng sinh tổng hợp hợp chất thứ cấp mới thuộc lassopeptide.



**Hình 3.10.** Các cụm gen sinh tổng hợp các chất kháng khuẩn được xác định trong hệ gen của *S. cacaoi* VCCM 22255.



Tương tự, cụm gen 14.3 chỉ có độ tương đồng 46% với cluster sinh tổng hợp thiazostatin, là một chất kháng khuẩn (Hình 3.11). Cụm gen 14.3 không có khả năng sinh chất watasemycin, cũng là một chất kháng khuẩn mạnh, do không chứa gen *svn0515* để kích hoạt quá trình methyl hóa thiazoline C bằng chất tương đồng methylase gốc-SAM loại B tạo watasemycin [104]. Do chỉ tương đồng 3/4 gen NRPS so với cluster của *S. venezuelae* ATCC 10712, nên khả năng cao chủng *S. cacaoi* VCCM 22255 tạo ra các chất sinh học mới thay vì thiazostatin.

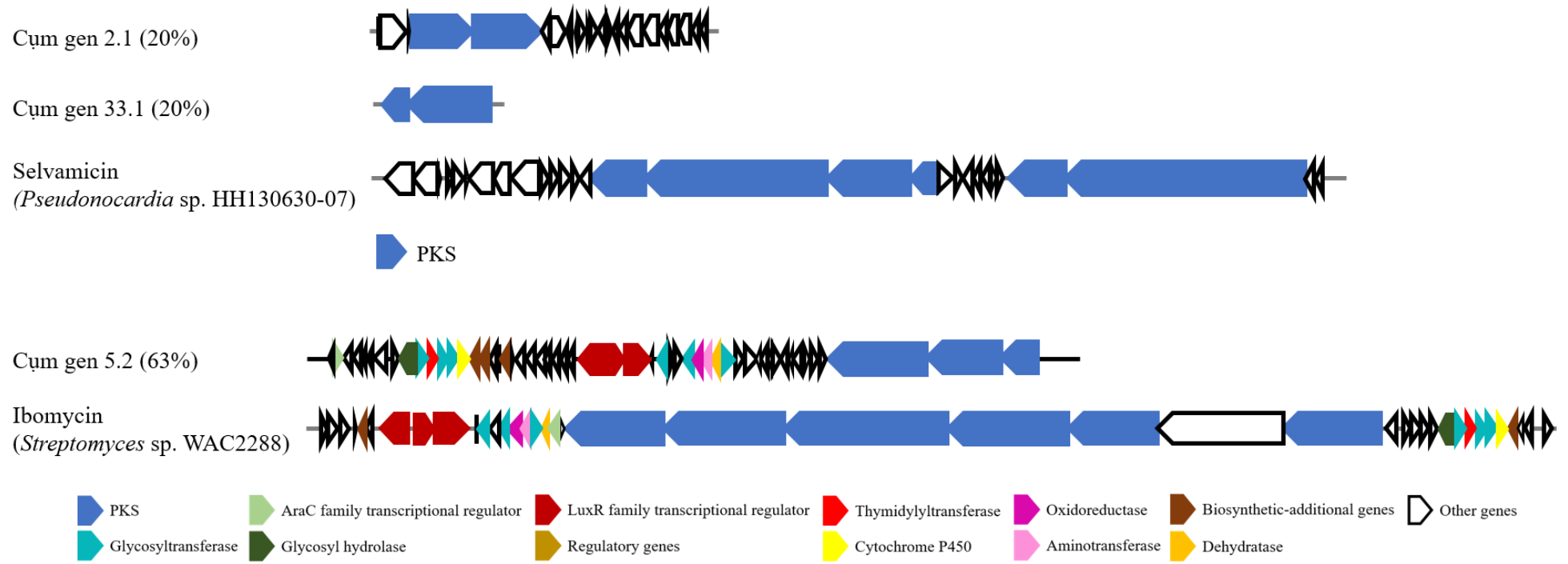
#### 3.4.4.3. Cụm gen sinh tổng hợp chất kháng nấm

Trong hệ gen *S. cacaoi* VCCM 22255, 3 cụm gen liên quan đến sinh tổng hợp chất kháng nấm được dự đoán, bao gồm cụm gen 2.1 và 33.1 (selvamicin), cụm gen 5.2 (ibomycin).

Cụm gen 5.2 được dự đoán là một phức hợp gồm các PKS và các gen khác và có độ tương đồng 63% với cụm gen tham chiếu sinh ibomycin từ *Streptomyces* sp. WAC2288 (Hình 3.11). Khung polyketide của ibomycin được tổng hợp bởi 86 miền xúc tác trên 20 mô-đun PKS trải rộng trên bảy gen PKS đa miền và tất cả các enzyme trong cụm ibomycin phải hoạt động với thứ tự của các mô-đun phù hợp để sản xuất một polyketide một cách tuyến tính [105]. Bên cạnh đó, cụm gen 5.2 cũng thể hiện độ tương đồng tới 55% với cụm gen sinh chất kháng nấm aculeximycin của *Kutzneria albida* DSM 43870.

Từ hệ gen *S. cacaoi* VCCM 22255 có hai cụm gen 2.1 và 33.1 đều thể hiện độ tương đồng 20% với cụm gen tạo selvamicin của chủng *Pseudonocardia* sp. HH130630-07. Cả hai cluster chỉ chứa hai gen PKS, trong khi đó cần có một loạt cụm gen PKS I chịu trách nhiệm cho quá trình sinh tổng hợp của selvamicin bao gồm một phân cụm gen phù hợp với đường 4-O-methyldigitoxose của selvamicin [106].

Như vậy, nhờ công cụ antiSMASH, đã dự đoán được 17 cluster, trong đó có 10 cluster có tiềm năng sinh tổng hợp hợp chất thứ cấp có hoạt tính chống ung thư, kháng khuẩn và kháng nấm, phù hợp với kết quả khảo sát hoạt tính sơ bộ ban đầu của chủng *S. cacaoi* VCCM 22255 với khả năng kháng các chủng VSV kiểm định và ức chế một số dòng tế bào ung thư. Các kết quả trên chỉ thể hiện tiềm năng sinh các chất, do vậy cần những nghiên cứu chuyên sâu để xác định liệu có phải các cụm gen này sinh tổng hợp các hoạt chất mới.



**Hình 3.11. Các cụm gen sinh tổng hợp các chất kháng nấm được xác định trong hệ gen của *S. cacaoi* VCCM 22255.**

## KẾT LUẬN VÀ KIẾN NGHỊ

### KẾT LUẬN

- Đã xác định được chủng xạ khuẩn nội sinh trên cây màng tang LCP02 (VCCM 22255) thể hiện hoạt tính kháng khuẩn và gây độc tế bào ung thư tiềm năng. Chủng VCCM 22255 kháng 8 loại vi khuẩn kiểm định với vòng kháng khuẩn dao động từ 8,3 - 34,0 mm. Cao chiết thô cũng ức chế mạnh 3 dòng tế bào ung thư A549 và MCF7 ở nồng độ 30 µg/ml.
- Chủng nghiên cứu được định danh là *S. cacaoi* VCCM 22255 bằng phân tích trình tự gen 16S rRNA.
- Đã giải trình tự hệ gen chủng xạ khuẩn VCCM 22255 bằng công nghệ Illumina nhằm nghiên cứu đặc điểm di truyền liên quan đến hoạt tính kháng sinh và gây độc tế bào ung thư. Kích thước hệ gen đạt 8,44 Mb với hàm lượng GC là 73,3%, mã hoá cho 7359 protein. Trong đó, 10 cụm gen tiềm năng sinh chất chống ung thư và kháng vi sinh vật được dự đoán và phân tích.

### KIẾN NGHỊ

- Áp dụng công nghệ chỉnh sửa hệ gen CRISPR-CAS9 nhằm khai thác các cụm gen tiềm năng.

## TÀI LIỆU THAM KHẢO

1. Singh, R., and Dubey, A. K., 2015, Endophytic actinomycetes as emerging source for therapeutic compounds, *Indo Glob. J Pharm Sci*, **5**(2), pp. 106–116.
2. Ryan, R. P., Germaine, K., Franks, A., Ryan, D. J., and Dowling, D. N., 2008, Bacterial endophytes: recent developments and applications, *FEMS Microbiol. Lett.*, **278**(1), pp. 1–9.
3. Staniek, A., Woerdenbag, H. J., and Kayser, O., 2008, Endophytes: exploiting biodiversity for the improvement of natural product-based drug discovery, *J. Plant Interact.*, **3**(2), pp. 75–93.
4. Schulz, B., and Boyle, C., 2006, *What Are Endophytes?*, Springer.
5. Matsukuma, S., Okuda, T., and Watanabe, J., 1994, Isolation of actinomycetes from pine litter layers, *Actinomycetologica*, **8**(2), pp. 57–65.
6. Matsumoto, A., Takahashi, Y., Mochizuki, M., Seino, A., Iwai, Y., and Omura, S., 1998, Characterization of actinomycetes isolated from fallen leaves, *Actinomycetologica*, **12**(1), pp. 46–48.
7. Okazaki, T., 2003, *Studies on actinomycetes isolated from plant leaves*, Sel. Isol. Rare Actinomycetes.
8. Stone, J. K., Bacon, C. W., and White Jr, J. F., 2000, An overview of endophytic microbes: endophytism defined, *Microb. Endophytes*, pp. 17–44.
9. Nalini, M. S., and Prakash, H. S., 2017, Diversity and bioprospecting of actinomycete endophytes from the medicinal plants, *Lett. Appl. Microbiol.*, **64**(4), pp. 261–270.
10. Rosenblueth, M., and Martínez-Romero, E., 2006, Bacterial endophytes and their interactions with hosts, *Mol. Plant. Microbe Interact.*, **19**(8), pp. 827–837.
11. Alzahrani, K., Jastaniah, S., Amasha, R., and Aly, M. M., 2022, Diversity and colonization of endophytic actinomycetes in some medicinal plants: review, *J. Contemp. Med. Sci.*, **8**(3).
12. Zhi, X.-Y., Li, W.-J., and Stackebrandt, E., 2009, An update of the structure and 16S rRNA gene sequence-based definition of higher ranks of the class Actinobacteria, with the proposal of two new suborders and four new families and emended descriptions of the existing higher taxa, *Int. J. Syst. Evol. Microbiol.*, **59**(3), pp. 589–608.

13. Tang, S.-K., Zhi, X.-Y., Wang, Y., Shi, R., Lou, K., Xu, L.-H., and Li, W.-J., 2011, *Haloactinopolyspora alba* gen. nov., sp. nov., a halophilic filamentous actinomycete isolated from a salt lake, with proposal of *Jiangellaceae* fam. nov. and *Jiangellineae* subord. nov., *Int. J. Syst. Evol.*, **61**(1), pp. 194–200.
14. Keswani, J., and Whitman, W. B., 2001, Relationship of 16S rRNA sequence similarity to DNA hybridization in prokaryotes., *Int. J. Syst. Evol. Microbiol.*, **51**(2), pp. 667–678.
15. Besaury, L., Martinet, L., Mühle, E., Clermont, D., and Rémond, C., 2021, *Streptomyces silvae* sp. nov., isolated from forest soil, *Int. J. Syst. Evol. Microbiol.*, **71**(12).
16. El-Shatoury, S. A., El-Kraly, O. A., Trujillo, M. E., El-Kazzaz, W. M., El-Din, E.-S. G., and Dewedar, A., 2013, Generic and functional diversity in endophytic actinomycetes from wild *Compositae* plant species at South Sinai–Egypt, *Res. Microbiol.*, **164**(7), pp. 761–769.
17. Bérdy, J., 2012, Thoughts and facts about antibiotics: where we are now and where we are heading, *J. Antibiot. (Tokyo)*, **65**(8), pp. 385–395.
18. Taechowisan, T., Peberdy, J. F., and Lumyong, S., 2003, Isolation of endophytic actinomycetes from selected plants and their antifungal activity, *World J. Microbiol. Biotechnol.*, **19**, pp. 381–385.
19. Verma, V. C., Gond, S. K., Kumar, A., Mishra, A., Kharwar, R. N., and Gange, A. C., 2009, Endophytic actinomycetes from *Azadirachta indica* A. Juss.: isolation, diversity, and anti-microbial activity, *Microb. Ecol.*, **57**, pp. 749–756.
20. Akshatha, J. V., Prakash, H. S., and Nalini, M. S., 2016, Actinomycete endophytes from the ethno medicinal plants of Southern India: Antioxidant activity and characterization studies, *J. Biol. Act. Prod. Nat.*, **6**(2), pp. 166–172.
21. Alsaedi, Z. S., Ashy, R. A., Shami, A. Y., Majeed, M. A., Alswat, A. M., Baz, L., Baeshen, M. N., and Jalal, R. S., 2022, Metagenomic study of the communities of bacterial endophytes in the desert plant *Senna Italica* and their role in abiotic stress resistance in the plant, *Braz. J. Biol.*, **82**.
22. Xu, T., Vo, Q. T., Barnett, S. J., Ballard, R. A., Zhu, Y., and Franco, C. M., 2022, Revealing the underlying mechanisms mediated by endophytic

actinobacteria to enhance the rhizobia-chickpea (*Cicer arietinum* L.) symbiosis, *Plant Soil*, **474**(1–2), pp. 299–318.

23. Shimizu, M., Yazawa, S., and Ushijima, Y., 2009, A promising strain of endophytic *Streptomyces* sp. for biological control of cucumber anthracnose, *J. Gen. Plant Pathol.*, **75**, pp. 27–36.

24. Jog, R., Nareshkumar, G., and Rajkumar, S., 2016, Enhancing soil health and plant growth promotion by actinomycetes, *Plant Growth Promot. Actinobacteria New Ave. Enhancing Product. Soil Fertil. Grain Legum.*, pp. 33–45.

25. Farnet, C. M., and Zazopoulos, E., 2005, Improving drug discovery from microorganisms, *Nat. Prod. Drug Discov. Ther. Med.*, pp. 95–106.

26. Guan, S., Sattler, I., Lin, W., Guo, D., and Grabley, S., 2005, P-Aminoacetophenonic acids produced by a mangrove endophyte: *Streptomyces griseus* subsp., *J. Nat. Prod.*, **68**(8), pp. 1198–1200.

27. Strobel, G. A., 2003, Endophytes as sources of bioactive products, *Microbes Infect.*, **5**(6), pp. 535–544.

28. Qin, S., Xing, K., Jiang, J.-H., Xu, L.-H., and Li, W.-J., 2011, Biodiversity, bioactive natural products and biotechnological potential of plant-associated endophytic actinobacteria, *Appl. Microbiol. Biotechnol.*, **89**, pp. 457–473.

29. Castillo, U., Harper, J. K., Strobel, G. A., Sears, J., Alesi, K., Ford, E., Lin, J., Hunter, M., Maranta, M., and Ge, H., 2003, Kakadumycins, novel antibiotics from *Streptomyces* sp. NRRL 30566, an endophyte of *Grevillea pteridifolia*, *FEMS Microbiol. Lett.*, **224**(2), pp. 183–190.

30. Igarashi, Y., 2004, Screening of novel bioactive compounds from plant-associated actinomycetes, *Actinomycetologica*, **18**(2), pp. 63–66.

31. Ezra, D., Castillo, U. F., Strobel, G. A., Hess, W. M., Porter, H., Jensen, J. B., Condrón, M. A., Teplow, D. B., Sears, J., and Maranta, M., 2004, Coronamycins, peptide antibiotics produced by a verticillate *Streptomyces* sp.(MSU-2110) endophytic on *Monstera* sp., *Microbiology*, **150**(4), pp. 785–793.

32. El-Gendy, M. M., and El-Bondkly, A. M., 2010, Production and genetic improvement of a novel antimycotic agent, saadamycin, against dermatophytes and other clinical fungi from endophytic *Streptomyces* sp. Hedaya48, *J. Ind. Microbiol. Biotechnol.*, **37**(8), pp. 831–841.

33. Caruso, M., Colombo, A. L., Fedeli, L., Pavesi, A., Quaroni, S., Saracchi, M., and Ventrella, G., 2000, Isolation of endophytic fungi and actinomycetes taxane producers, *Ann. Microbiol.*, **50**(1), pp. 3–14.
34. Taechowisan, T., Chaisaeng, S., and Phutdhawong, W. S., 2017, Antibacterial, antioxidant and anticancer activities of biphenyls from *Streptomyces* sp. BO-07: an endophyte in *Boesenbergia rotunda* (L.) Mansf A., *Food Agric. Immunol.*, **28**(6), pp. 1330–1346.
35. Igarashi, Y., Yanase, S., Sugimoto, K., Enomoto, M., Miyanaga, S., Trujillo, M. E., Saiki, I., and Kuwahara, S., 2011, Lupinacidin C, an inhibitor of tumor cell invasion from *Micromonospora lupini*, *J. Nat. Prod.*, **74**(4), pp. 862–865.
36. Sanger, F., Nicklen, S., and Coulson, A. R., 1977, DNA sequencing with chain-terminating inhibitors, *Proc. Natl. Acad. Sci.*, **74**(12), pp. 5463–5467.
37. Land, M., Hauser, L., Jun, S.-R., Nookaew, I., Leuze, M. R., Ahn, T.-H., Karpinets, T., Lund, O., Kora, G., and Wassenaar, T., 2015, Insights from 20 years of bacterial genome sequencing, *Funct. Integr. Genomics*, **15**, pp. 141–161.
38. Consortium, I. H. G. S., 2004, Finishing the euchromatic sequence of the human genome, *Nature*, **431**(7011), pp. 931–945.
39. Ben Khedher, M., Ghedira, K., Rolain, J.-M., Ruimy, R., and Croce, O., 2022, Application and challenge of 3rd generation sequencing for clinical bacterial studies, *Int. J. Mol. Sci.*, **23**(3), p. 1395.
40. Selvakumar, S. C., Preethi, K. A., Ross, K., Tusubira, D., Khan, M. W. A., Mani, P., Rao, T. N., and Sekar, D., 2022, CRISPR/Cas9 and next generation sequencing in the personalized treatment of Cancer, *Mol. Cancer*, **21**(1), p. 83.
41. Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A.-M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., and Harper, D., 2002, Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2), *nature*, **417**(6885), pp. 141–147.
42. Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., and Ōmura, S., 2003, Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*, *Nat. Biotechnol.*, **21**(5), pp. 526–531.

43. Choulet, F., Gallois, A., Aigle, B., Mangenot, S., Gerbaud, C., Truong, C., Francou, F.-X., Borges, F., Fourier, C., and Guérineau, M., 2006, Intraspecific variability of the terminal inverted repeats of the linear chromosome of *Streptomyces ambofaciens*, *J. Bacteriol.*, **188**(18), pp. 6599–6610.
44. Dyson, P., and Schrempf, H., 1987, Genetic instability and DNA amplification in *Streptomyces lividans* 66, *J. Bacteriol.*, **169**(10), pp. 4796–4803.
45. Hopwood, D. A., 2006, Soil to genomics: the *Streptomyces* chromosome, *Annu Rev Genet*, **40**, pp. 1–23.
46. Kämpfer, P., 2006, The family *Streptomycetaceae*, part I: taxonomy, *The prokaryotes*, **3**, pp. 538–604.
47. Xu, M.-J., Wang, J.-H., Bu, X.-L., Yu, H.-L., Li, P., Ou, H.-Y., He, Y., Xu, F.-D., Hu, X.-Y., and Zhu, X.-M., 2016, Deciphering the streamlined genome of *Streptomyces xiamenensis* 318 as the producer of the anti-fibrotic drug candidate xiamenmycin, *Sci. Rep.*, **6**(1), p. 18977.
48. Caicedo-Montoya, C., Manzo-Ruiz, M., and Ríos-Estapa, R., 2021, Pan-genome of the genus *Streptomyces* and prioritization of biosynthetic gene clusters with potential to produce antibiotic compounds, *Front. Microbiol.*, **12**, p. 677558.
49. Belknap, K. C., Park, C. J., Barth, B. M., and Andam, C. P., 2020, Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria, *Sci. Rep.*, **10**(1), pp. 1–9.
50. Ziemert, N., Alanjary, M., and Weber, T., 2016, The evolution of genome mining in microbes—a review, *Nat. Prod. Rep.*, **33**(8), pp. 988–1005.
51. Lautru, S., Deeth, R. J., Bailey, L. M., and Challis, G. L., 2005, Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining, *Nat. Chem. Biol.*, **1**(5), pp. 265–269.
52. Barona-Gómez, F., Wong, U., Giannakopoulos, A. E., Derrick, P. J., and Challis, G. L., 2004, Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145, *J. Am. Chem. Soc.*, **126**(50), pp. 16282–16283.
53. Oves-Costales, D., Kadi, N., and Challis, G. L., 2009, The long-overlooked enzymology of a nonribosomal peptide synthetase-independent



pathway for virulence-conferring siderophore biosynthesis, *Chem. Commun.*, (43), pp. 6530–6541.

54. Kadi, N., Oves-Costales, D., Barona-Gomez, F., and Challis, G. L., 2007, A new family of ATP-dependent oligomerization-macrocyclization biocatalysts, *Nat. Chem. Biol.*, **3**(10), pp. 652–656.

55. Chen, Y., Wang, Y., Han, X., Si, L., Wu, Q., and Lin, L., 2013, Biology and chemistry of *Litsea cubeba*, a promising industrial tree in China, *J. Essent. Oil Res.*, **25**(2), pp. 103–111.

56. Bajracharya, G. B., and Pratigya, K. C., 2019, A high antibacterial efficacy of fruits of *Litsea cubeba* (Lour.) Pers from Nepal. GC-MS and antioxidative capacity analyses, *Pharmacogn. J.*, **11**(5).

57. Ho, C.-L., Jie-Ping, O., Liu, Y.-C., Hung, C.-P., Tsai, M.-C., Liao, P.-C., Wang, E. I.-C., Chen, Y.-L., and Su, Y.-C., 2010, Compositions and *in vitro* anticancer activities of the leaf and fruit oils of *Litsea cubeba* from Taiwan, *Nat. Prod. Commun.*, **5**(4), p. 1934578X1000500425.

58. Zhang, W., Hu, J.-F., Lv, W.-W., Zhao, Q.-C., and Shi, G.-B., 2012, Antibacterial, antifungal and cytotoxic isoquinoline alkaloids from *Litsea cubeba*, *Molecules*, **17**(11), pp. 12950–12960.

59. Hồng, H. T., and Phương, N. N., 2013, Phân lập và tuyển chọn chủng xạ khuẩn từ rừng ngập mặn Cần Giờ kháng nấm *Fusarium* sp., *Tạp Chí Khoa Học*, (51), p. 59.

60. Duong, M. L., Nguyen, D. V., and Tong, T. M., 2014, Screening for anticancer producing endophytic actinomycetes in three mangrove plant species in nam dinh province, *HNUE J. Sci. Biol. Sci.*, **59**(9), pp. 3–12.

61. Khieu, T.-N., Liu, M.-J., Nimaichand, S., Quach, N.-T., Chu-Ky, S., Phi, Q.-T., Vu, T.-T., Nguyen, T.-D., Xiong, Z., and Prabhu, D. M., 2015, Characterization and evaluation of antimicrobial and cytotoxic effects of *Streptomyces* sp. HUST012 isolated from medicinal plant *Dracaena cochinchinensis* Lour., *Front. Microbiol.*, **6**, p. 574.

62. Vu, H.-N. T., Nguyen, D. T., Nguyen, H. Q., Chu, H. H., Chu, S. K., Chau, M. V., and Phi, Q.-T., 2018, Antimicrobial and cytotoxic properties of bioactive metabolites produced by *Streptomyces cavourensis* YBQ59 isolated from *Cinnamomum cassia* Pries in Yen Bai Province of Vietnam, *Curr. Microbiol.*, **75**, pp. 1247–1255.

63. Vu, T. H. N., Nguyen, Q. H., Dinh, T. M. L., Quach, N. T., Khieu, T. N., Hoang, H., Chu-Ky, S., Vu, T. T., Chu, H. H., and Lee, J., 2020, Endophytic actinomycetes associated with *Cinnamomum cassia* Presl in Hoa Binh province, Vietnam: Distribution, antimicrobial activity and, genetic features, *J. Gen. Appl. Microbiol.*, **66**(1), pp. 24–31.
64. Tran, T. X., Le, T. T., Trieu, L. P., Austin, C. M., Van Quyen, D., and Nguyen, H. M., 2019, Whole-genome sequencing and characterization of an antibiotic resistant *Neisseria meningitidis* B isolate from a military unit in Vietnam, *Ann. Clin. Microbiol. Antimicrob.*, **18**, pp. 1–9.
65. Van Thuoc, D., Loan, T. T., Trung, T. A., Van Quyen, N., Tung, Q. N., Tien, P. Q., and Sudesh, K., 2020, Genome mining reveals the biosynthetic pathways of polyhydroxyalkanoate and ectoines of the halophilic strain *Salinivibrio proteolyticus* M318 isolated from fermented shrimp paste, *Mar. Biotechnol.*, **22**, pp. 651–660.
66. Le Ngoc Giang, L. T. H., Minh, V. T. Q., Nguyen Mai Anh, N. T. K., and Cuc, V. T. T. H., 2020, Genome mining of a marine-derived *Streptomyces* sp. PDH23 isolated from sponge in Da Nang sea for secondary metabolite gene clusters, *J. Biotechnol.*, **18**(4), pp. 709–721.
67. Quach, N. T., Vu, T. H. N., Bui, T. L., Le, T. T. X., Nguyen, T. T. A., Ngo, C. C., and Phi, Q.-T., 2022, Genomic and physiological traits provide insights into ecological niche adaptations of mangrove endophytic *Streptomyces parvulus* VCCM 22513, *Ann. Microbiol.*, **72**(1), p. 27.
68. Quach, N. T., Nguyen, Q. H., Vu, T. H. N., Le, T. T. H., Ta, T. T. T., Nguyen, T. D., Van Doan, T., Van Nguyen, T., Dang, T. T., and Nguyen, X. C., 2021, Plant-derived bioactive compounds produced by *Streptomyces variabilis* LCP18 associated with *Litsea cubeba* (Lour.) Pers as potential target to combat human pathogenic bacteria and human cancer cell lines, *Braz. J. Microbiol.*, **52**(3), pp. 1215–1224.
69. Saadoun, I., and Muhana, A., 2008, Optimal production conditions, extraction, partial purification and characterization of inhibitory compound (s) produced by *Streptomyces* Ds-104 isolate against multi-drug resistant *Candida albicans*, *Curr. Trends Biotechnol. Pharm.*, **2**(2), pp. 402–432.
70. Skehan, P., Storeng, R., Scudiero, D., Monks, A., McMahon, J., Vistica, D., Warren, J. T., Bokesch, H., Kenney, S., and Boyd, M. R., 1990,

New colorimetric cytotoxicity assay for anticancer-drug screening, *JNCI J. Natl. Cancer Inst.*, **82**(13), pp. 1107–1112.

71. Hughes, J. P., Rees, S., Kalindjian, S. B., and Philpott, K. L., 2011, Principles of early drug discovery, *Br. J. Pharmacol.*, **162**(6), pp. 1239–1249.

72. Salam, N., Khieu, T.-N., Liu, M.-J., Vu, T.-T., Chu-Ky, S., Quach, N.-T., Phi, Q.-T., Narsing Rao, M. P., Fontana, A., and Sarter, S., 2017, Endophytic actinobacteria associated with *Dracaena cochinchinensis* Lour.: isolation, diversity, and their cytotoxic activities, *BioMed Res. Int.*, **2017**.

73. Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., and Pusch, G. D., 2015, RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes, *Sci. Rep.*, **5**(1), p. 8365.

74. Davín, A. A., Tricou, T., Tannier, E., de Vienne, D. M., and Szöllősi, G. J., 2020, Zombi: a phylogenetic simulator of trees, genomes and sequences that accounts for dead lineages, *Bioinformatics*, **36**(4), pp. 1286–1288.

75. Davis, J. J., Wattam, A. R., Aziz, R. K., Brettin, T., Butler, R., Butler, R. M., Chlenski, P., Conrad, N., Dickerman, A., and Dietrich, E. M., 2020, The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities, *Nucleic Acids Res.*, **48**(D1), pp. D606–D612.

76. Huang, L., Zhang, H., Wu, P., Entwistle, S., Li, X., Yohe, T., Yi, H., Yang, Z., and Yin, Y., 2018, DbCAN-seq: a database of carbohydrate-active enzyme (CAZyme) sequence and annotation, *Nucleic Acids Res.*, **46**(D1), pp. D516–D521.

77. Peng, Q. I. U., Zhi-Xiang, F., Jie-Wei, T., Zu-Chao, L. E. I., Lei, W., Zhi-Gang, Z., Yi-Wen, C. H. U., and Yong-Qiang, T., 2015, Diversity, bioactivities, and metabolic potentials of endophytic actinomycetes isolated from traditional medicinal plants in Sichuan, China, *Chin. J. Nat. Med.*, **13**(12), pp. 942–953.

78. Gohain, A., Gogoi, A., Debnath, R., Yadav, A., Singh, B. P., Gupta, V. K., Sharma, R., and Saikia, R., 2015, Antimicrobial biosynthetic potential and genetic diversity of endophytic actinomycetes associated with medicinal plants, *FEMS Microbiol. Lett.*, **362**(19), p. fnv158.

79. Janaki, T., 2019, Anticancer activity of *Streptomyces cacaoi* subsp *cacaoi* M20 against breast cancer (MCF-7) cell lines, Int J Chemtech Res Httpdx Doi Org1020902IJCTR.
80. Zhou, H., Yang, Y., Zhang, J., Peng, T., Zhao, L., Xu, L., and Ding, Z., 2013, Alkaloids from an endophytic *Streptomyces* sp. YIM66017, Nat. Prod. Commun., **8**(10), p. 1934578X1300801012.
81. Savi, D. C., Shaaban, K. A., Vargas, N., Ponomareva, L. V., Possiede, Y. M., Thorson, J. S., Glienke, C., and Rohr, J., 2015, *Microbispora* sp. LGMB259 endophytic actinomycete isolated from *Vochysia divergens* (Pantanal, Brazil) producing  $\beta$ -carboline and indoles with biological activity, Curr. Microbiol., **70**, pp. 345–354.
82. Zhang, J., Wang, J.-D., Liu, C.-X., Yuan, J.-H., Wang, X.-J., and Xiang, W.-S., 2014, A new prenylated indole derivative from endophytic actinobacteria *Streptomyces* sp. neau-D50, Nat. Prod. Res., **28**(7), pp. 431–437.
83. Lanoot, B., Vancanneyt, M., Cleenwerck, I., Wang, L., Li, W., Liu, Z., and Swings, J., 2002, The search for synonyms among streptomycetes by using SDS-PAGE of whole-cell proteins. Emendation of the species *Streptomyces aurantiacus*, *Streptomyces cacaoi* subsp. *cacaoi*, *Streptomyces caeruleus* and *Streptomyces violaceus*., Int. J. Syst. Evol. Microbiol., **52**(3), pp. 823–829.
84. Kaweewan, I., Hemmi, H., Komaki, H., and Kodani, S., 2020, Isolation and structure determination of a new antibacterial peptide pentaminomycin C from *Streptomyces cacaoi* subsp. *cacaoi*, J. Antibiot. (Tokyo), **73**(4), pp. 224–229.
85. Endo, A., and Misato, T., 1969, Polyoxin D, a competitive inhibitor of UDP-N-acetylglucosamine: chitin N-acetylglucosaminyltransferase in *Neurospora crassa*, Biochem. Biophys. Res. Commun., **37**(4), pp. 718–722.
86. Khan, N., Yilmaz, S., Aksoy, S., Uzel, A., Tosun, Ç., Kirmizibayrak, P. B., and Bedir, E., 2019, Polyethers isolated from the marine actinobacterium *Streptomyces cacaoi* inhibit autophagy and induce apoptosis in cancer cells, Chem. Biol. Interact., **307**, pp. 167–178.
87. Quach, N. T., Bui, T. L., Vu, T. H. N., Nguyen, T. T. A., Chu, H. H., and Phi, Q. T., 2022, Endophytic actinomycetes from mangrove plant

- Avicennia marina* in Quang Ninh province, Vietnam: distribution, cytotoxicity, and antioxidant activities, *Acad. J. Biol.*, **44**(3), pp. 87–98.
88. Espejo, R. T., and Plaza, N., 2018, Multiple ribosomal RNA operons in bacteria; their concerted evolution and potential consequences on the rate of evolution of their 16S rRNA, *Front. Microbiol.*, **9**, p. 1232.
89. Eftekharivash, L., and Hamed, J., 2020, Genome sequence and annotation of *Streptomyces tendae* UTM 3329, acid and alkaline tolerant actinobacterium, Iran. *J. Microbiol.*, **12**(4), p. 343.
90. Yi, J. S., Kim, J. M., Kang, M.-K., Kim, J. H., Cho, H. S., Ban, Y. H., Song, M. C., Son, K.-H., and Yoon, Y. J., 2022, Whole-genome sequencing and analysis of *Streptomyces* strains producing multiple antinematode drugs, *BMC Genomics*, **23**(1), pp. 1–18.
91. Ward, A. C., and Allenby, N. E., 2018, Genome mining for the search and discovery of bioactive compounds: the *Streptomyces* paradigm, *FEMS Microbiol. Lett.*, **365**(24), p. fny240.
92. Pita-Grisanti, V., Chasser, K., Sobol, T., and Cruz-Monserrate, Z., 2022, Understanding the potential and risk of bacterial siderophores in cancer, *Front. Oncol.*, **12**, p. 1885.
93. Salis, O., Bedir, A., Kilinc, V., Alacam, H., Gulden, S., and Okuyucu, A., 2014, The anticancer effects of desferrioxamine on human breast adenocarcinoma and hepatocellular carcinoma cells, *Cancer Biomark.*, **14**(6), pp. 419–426.
94. Jiang, Z., Tuo, L., Huang, D., Osterman, I. A., Tyurin, A. P., Liu, S., Lukyanov, D. A., Sergiev, P. V., Dontsova, O. A., and Korshun, V. A., 2018, Diversity, novelty, and antimicrobial activity of endophytic actinobacteria from mangrove plants in Beilun Estuary National Nature Reserve of Guangxi, China, *Front. Microbiol.*, **9**, p. 868.
95. Patzer, S. I., and Braun, V., 2010, Gene cluster involved in the biosynthesis of griseobactin, a catechol-peptide siderophore of *Streptomyces* sp. ATCC 700974, *J. Bacteriol.*, **192**(2), pp. 426–435.
96. Pinkerton, D. M., Banwell, M. G., Garson, M. J., Kumar, N., de Moraes, M. O., Cavalcanti, B. C., Barros, F. W., and Pessoa, C., 2010, Antimicrobial and cytotoxic activities of synthetically derived tambjamins C and E–J, BE-18591, and a related alkaloid from the marine bacterium *Pseudoalteromonas tunicata*, *Chem. Biodivers.*, **7**(5), pp. 1311–1324.

97. Cerdeño, A. M., Bibb, M. J., and Challis, G. L., 2001, Analysis of the prodiginine biosynthesis gene cluster of *Streptomyces coelicolor* A3 (2): new mechanisms for chain initiation and termination in modular multienzymes, *Chem. Biol.*, **8**(8), pp. 817–829.
98. Grenade, N. L., Chiriac, D. S., Howe, G. W., and Ross, A., 2020, Identification of a tambjamine gene cluster in *Streptomyces* reveals convergent evolution of the biosynthetic pathway.
99. Zhang, W., Wang, L., Kong, L., Wang, T., Chu, Y., Deng, Z., and You, D., 2012, Unveiling the post-PKS redox tailoring steps in biosynthesis of the type II polyketide antitumor antibiotic xantholipin, *Chem. Biol.*, **19**(3), pp. 422–432.
100. Liu, L.-L., Chen, Z.-F., Liu, Y., Tang, D., Gao, H.-H., Zhang, Q., and Gao, J.-M., 2020, Molecular networking-based for the target discovery of potent antiproliferative polycyclic macrolactam ansamycins from *Streptomyces cacaoi* subsp. *asoensis*, *Org. Chem. Front.*, **7**(24), pp. 4008–4018.
101. Tang, D., Liu, L.-L., He, Q.-R., Yan, W., Li, D., and Gao, J.-M., 2018, Ansamycins with antiproliferative and antineuroinflammatory activity from moss-soil-derived *Streptomyces cacaoi* subsp. *asoensis* H2S5, *J. Nat. Prod.*, **81**(9), pp. 1984–1991.
102. Sosio, M., Gaspari, E., Iorio, M., Pessina, S., Medema, M. H., Bernasconi, A., Simone, M., Maffioli, S. I., Ebricht, R. H., and Donadio, S., 2018, Analysis of the pseudouridimycin biosynthetic pathway provides insights into the formation of C-nucleoside antibiotics, *Cell Chem. Biol.*, **25**(5), pp. 540-549. e4.
103. Rosenqvist, P., Palmu, K., Prajapati, R. K., Yamada, K., Niemi, J., Belogurov, G. A., Metsä-Ketelä, M., and Virta, P., 2019, Characterization of C-nucleoside antimicrobials from *Streptomyces albus* DSM 40763: strepturidin is pseudouridimycin, *Sci. Rep.*, **9**(1), p. 8935.
104. Inahashi, Y., Zhou, S., Bibb, M. J., Song, L., Al-Bassam, M. M., Bibb, M. J., and Challis, G. L., 2017, Watasemycin biosynthesis in *Streptomyces venezuelae*: thiazoline C-methylation by a type B radical-SAM methylase homologue, *Chem. Sci.*, **8**(4), pp. 2823–2831.
105. Robbins, N., Spitzer, M., Wang, W., Waglechner, N., Patel, D. J., O'Brien, J. S., Ejim, L., Ejim, O., Tyers, M., and Wright, G. D., 2016,

Discovery of ibomycin, a complex macrolactone that exerts antifungal activity by impeding endocytic trafficking and membrane function, *Cell Chem. Biol.*, **23**(11), pp. 1383–1394.

106. Van Arnam, E. B., Ruzzini, A. C., Sit, C. S., Horn, H., Pinto-Tomás, A. A., Currie, C. R., and Clardy, J., 2016, Selvamycin, an atypical antifungal polyene from two alternative genomic contexts, *Proc. Natl. Acad. Sci.*, **113**(46), pp. 12940–12945.

## PHỤ LỤC

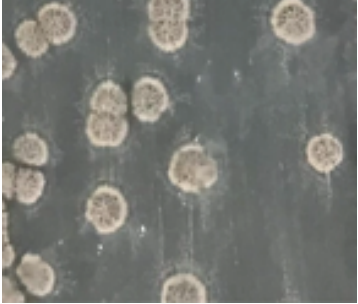
### Phụ lục 1. Trình tự 16S rRNA của chủng xạ khuẩn VCCM 22255.

LOCUS	VCCM 22255	742 bp	DNA	linear		
ORIGIN						
1	GAGTAACACG	TGGGCAATCT	GCCCTGCACT	CTGGGACAAG	CCCTGGAAAC	GGGGTCTAAT
61	ACCGGATATG	ACCACCGGCC	GCATGGTCTG	GTGGTGGAAA	GCTCCGGCGG	TGCAGGATGA
121	GCCCGCGGCC	TATCAGCTTG	TTGGTGGGGT	GATGGCCTAC	CAAGGCGACG	ACGGGTAGCC
181	GGCCTGAGAG	GGCGACCGGC	CACACTGGGA	CTGAGACACG	GCCCAGACTC	CTACGGGAGG
241	CAGCAGTGGG	GAATATTGCA	CAATGGGCGC	AAGCCTGATG	CAGCGACGCC	GCGTGAGGGA
301	TGACGGCCTT	CGGGTTGTAA	ACCTCTTTCA	GCAGGGAAGA	AGCGCGAGTG	ACGGTACCTG
361	CAGAAGAAGC	ACCGGCTAAC	TACGTGCCAG	CAGCCGCGGT	AATACGTAGG	GTGCGAGCGT
421	TGTCCGGAAT	TATTGGGCGT	AAAGAGCTCG	TAGGCGGCCT	GTCGCGTCGG	ATGTGAAAGC
481	CCGGGGCTTA	ACCCCGGGTC	TGCATTTCGAT	ACGGGCAGGC	TAGAGTTCGG	CAGGGGAGAT
541	TGGAATTCCT	GGTGTAGCGG	TGAAATGCGC	AGATATCAGG	AGGAACACCG	GTGGCGAAGG
601	CGGATCTCTG	GGCCGATACT	GACGCTGAGG	AGCGAAAGCG	TGGGGAGCGA	ACAGGATTAG
661	ATACCCTGGT	AGTCCACGCC	GTAAACGTTG	GGCACTAGGT	GTGGGCGGCA	TTCCACGTCG
721	TCCGTGCCGC	AGCTAACGCA	TT			

//



**Phụ lục 2.** Đặc điểm hình thái khuẩn lạc của chủng xạ khuẩn VCCM 22255.

<b>Đặc điểm khuẩn lạc</b>	<b>Hình ảnh khuẩn lạc</b>
Tròn không đều, màu xám trắng, khô, dạng phấn, có phóng xạ	

**Phụ lục 3.** Độ tương đồng của chủng xạ khuẩn VCCM 22255 với các chủng *Streptomyces* khác theo công cụ BLAST.

<b>STT</b>	<b>Tên chủng</b>	<b>Độ tương đồng (%)</b>	<b>Mã số GenBank</b>
1	<i>Streptomyces cacaoi</i> NBRC 12748	100	NR_041061.1
2	<i>Streptomyces cacaoi</i> A-87	100	LR702037.1
3	<i>Streptomyces cacaoi</i> ASC3-2	100	LC506284.1
4	<i>Streptomyces cacaoi</i> ICMP 643	100	MK356400.1
5	<i>Streptomyces cacaoi</i> MAP1	100	KU550188.1
6	<i>Streptomyces cacaoi</i> GACK1	100	KP970679.1
7	<i>Streptomyces smyrnaeus</i> SM3501	99,60	NR_134201.1
8	<i>Streptomyces violaceoruber</i> CSSP679	99,19	NR_115407.1
9	<i>Streptomyces qinglanensis</i> 172205	99,19	NR_109303.1
10	<i>Streptomyces albus</i> CSSP726	98,92	NR_115440.1

**Phụ lục 4.** Các nhóm gen CAZyme của chủng *S. cacaoi* VCCM 22255.

<b>Nhóm glycoside hydrolase (GH)</b>	<b>Số lượng gen</b>	<b>Nhóm glycosyl transferase (GT)</b>	<b>Số lượng gen</b>
GH 1	2	GT 1	11
GH 2	1	GT 2	16
GH 3	4	GT 4	13
GH 4	3	GT 9	2
GH 5	2	GT 20	1
GH 6	1	GT 28	1
GH 13	3	GT 39	1
GH 15	4	GT 51	3
GH 18	4	GT 81	1
GH 19	1	GT 83	1
GH 20	4	GT 87	3
GH 23	2	GT 89	1
GH 25	3	<b>Nhóm auxiliary activitie (AA)</b>	<b>Số lượng gen</b>
GH 29	1	AA 3	1
GH 31	1	AA 10	2
GH 33	2	<b>Nhóm carbohydrate esterase (CE)</b>	<b>Số lượng gen</b>
GH 35	1	CE 4	4
GH 38	1	CE 9	1
GH 46	1	CE 14	2
GH 63	1	<b>Nhóm carbohydrate-binding module (CBM)</b>	<b>Số lượng gen</b>
GH 64	1	CBM 6	1
GH 65	1	CBM 12	1
GH 77	1	CBM 13	1
GH 85	1	<b>Nhóm polysaccharide lyase (PL)</b>	<b>Số lượng gen</b>
GH 87	1	PL 8	2
GH 89	1	PL 12	1

GH 92	1		
GH 101	1		
GH 109	1		
GH 114	2		
GH 141	1		
GH 154	1		
GH 171	1		