Vietnam Journal of Biotechnology **22**(4): 602-615, 2024. DOI: 10.15625/vjbt-21610

IDENTIFICATION AND CHARACTERIZATION OF THE CYTOCHROME P450 COMPLEMENT IN *STREPTOMYCES CAVOURENSIS* YBQ59

Ngoc Tung Quach[1], Thi Hanh Nguyen Vu[1], Thi Mai Phuong Nguyen[1], Quyet Tien Phi[1] and Thi Bich Thuy Ly[1,✉]

[1]*Institute of Biotechnology- Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Ha Noi,10000, Vietnam.*

✉To whom correspondence should be addressed. E-mail: ltb.thuy@ibt.ac.vn

Received: 27.09.2024
Accepted: 18.11.2024

ABSTRACT

Cytochrome P450 enzymes (CYPs) are regarded as some of the most versatile biocatalysts. They are attractive candidates for natural product development because of their ability to selectively oxidize a broad range of substrates. *Streptomyces* spp. are not only producers of biologically active secondary metabolites but also a rich source of P450 enzymes. However, only a limited number of studies have explored the function and potential of P450 enzymes encoded in the *Streptomyces* genomes. In this study, the endophytic *Streptomyces cavourensis* YBQ59 isolated from *Cinnamomum cassia* J. Presl was sequenced using the Illumina sequencing platform to identify its P450 enzymes. The genome of YBQ59 was approximately 8,126,002 bp in size, with a G + C content of 72.1% and contained 7,020 genes. Genome annotation identified 21 CYP genes, distributed across 10 CYP families and 17 subfamilies. The possible role of these P450 enzymes in the synthesis of secondary metabolites was discussed. Since CYPs often require electron transport proteins to function, we analyzed the physical map of the genes encoding ferredoxins and ferredoxin reductases found in the genome of *S. cavourensis* YBQ59. Additionally, a phylogenetic tree was constructed to compare the P450 enzyme system from *S. cavourensis* YBQ59 with those of closely related and well-studied *Streptomyces* species, including *Streptomyces* sp. CFMR7, *S. fulvissimus* DSM 40593, *S. griseus* IFO 13350, and *S. globisporus* 1912. These results provide a basis for exploiting potential P450 enzymes from *S. cavourensis* YBQ59 for agricultural and medicinal applications.

Keywords: *cytochrome P450, ferredoxin, ferredoxin reductase, genome, streptomyces cavourensis*

INTRODUCTION

Cytochrome P450s are monooxygenases that contain heme-thiolate in their active sites. These enzymes can catalyze regio- and stereo-selective oxidation of a wide range of substrates at allylic positions as well as non-activated C-H bonds (Bernhardt, 2006). This characteristic makes them valuable candidates for biotransformation of natural compounds, thereby being particularly appealing for the development

and exploitation of bioactive natural products.

Streptomyces, a genus of actinobacteria, is recognized for its prolific production of important antibiotics and other bioactive molecules. Notable metabolites produced by this genus include the antibiotics streptomycin and tetracycline, the immunosuppressants rapamycin and tacrolimus, the antihelminthic avermectin, the antifungal amphotericin B, and the antitumor mitomycin C (Watve et al., 2001). Genomes of *Streptomyces* are quite large, ranging from 8 to 10 Mb, with high G+C content, and often encode numerous biosynthetic gene clusters for secondary metabolites, including polyketides and non-ribosomal peptides (Jackson et al., 2018). These bacteria possess an unusually large number of CYP enzyme-encoding genes in their genomes. These enzymes are involved not only in macrolide biosynthesis but also in the formation of other natural compounds. To date, only about 0.4 percent of *Streptomyces* CYPs have been found to have unique chemical properties essential for the synthesis of secondary metabolites (Rudolf et al., 2017).

S. cavourensis YBQ59 was isolated as an endophyte from the medicinal plant *Cinnamomum cassia* J. Presl. This plant is used in traditional medicine as a tonic, antibiotic, antidiabetic and anti-inflammatory agent, and its essential oil composition has been extensively studied for medicinal and food additive applications (Huang et al., 2014). Notably, eight bioactive metabolites were identified in the ethyl acetate extract of the YBQ59, including 1-monolinolein, bafilomycin D, nonactic acid, daidzein, 3′-hydroxydaidzein, 5,11-epoxy-10-cadinanol, prelactone B, and daucosterol (Vu et al., 2018). The endophytic strain YBQ59 was initially sequenced using Ion Torrent PGM technology (Nguyen et al., 2018). However, low-quality genome assembly hindered further genomic analysis. In this study, the genome of *S. cavourensis* YBQ59 was re-sequenced using the Illumina technology. The cytochrome P450 complement (CYPome) of *S. cavourensis* was identified for the first time to explore the role of P450 enzymes in synthesizing both known and putative secondary metabolites.

MATERIAL AND METHODS

Bacterial strain and DNA manipulation

The endophytic *S. cavourensis* YBQ59 was provided by VAST-Culture Collection of Microorganisms, Institute of Biotechnology, Vietnam Academy of Science and Technology. Genomic DNA was extracted using the G-spin™ Total DNA Extraction Mini Kit (Intron Bio, Korea).

Whole-genome sequencing of *S. cavourensis* YBQ59, *de novo* assembly and annotation

Genomic DNA was fragmented and purified with AMPure XP magnetic beads to prepare the sequencing library. The sample was analyzed using an Agilent 2100 Bioanalyzer and sequenced on the Illumina platform (San Diego, CA, USA). Quality control and read trimming were conducted using FastQC and Trimmomatic tools.

The draft genome of the endophytic *S. cavourensis* YBQ59 was assembled using the SPAdes 3.15 with default parameters. The completeness of the resulting genome was assessed using Benchmarking Universal Single-Copy Orthologous (BUSCO) 3 (https://gitlab.com/ezlab/busco).

The whole genome sequence of the YBQ59 was annotated using Rapid Annotation using Subsystem Technology (RAST). Additionally, the genomic characteristics of YBQ59 were analyzed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). The draft genome sequence of *S. cavourensis* YBQ59 was deposited at the GenBank dataset (NCBI) under accession number JASEND000000000.

In-silico identification of CYPs

The CYPs of *S. cavourensis* YBQ59 were initially identified using Motif Scan (https://myhits.isb-sib.ch/cgi-bin/motif_scan) and Conserved Domain Database (CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The selected proteins were screened for the presence of characteristic CYP motifs, such as EXXR and GXXXCXG. The candidates were further annotated using the Nelson's International P450 Nomenclature Committee Database (https://drnelson.uthsc.edu) with an *E*-value threshold of 10^{-4}. Additionally, a sequence similarity cutoff of over 40% was used to identify known families, while CYPs with sequence similarity below 40% were classified as new families. The potential functions of all putative CYPs in *S. cavourensis* YBQ59 were predicted by comparing them to homologues CYPs characterized in other organisms, using Blast against the Protein Data Bank database (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). Finally, the CYP families and subfamilies were assigned and confirmed by Dr. David Nelson from the University of Tennessee Health Science Center, USA.

Phylogenetic reconstruction and physical maps of CYPs

The CYP protein sequences of *S. cavourensis* YBQ59 were aligned with those collected from *Streptomyces* sp. CFMR 7, *S. fulvissimus* DSM 40593, *S. griseus* IFO 13350, and *S. globisporus* 1912 using ClustalW. The bootstrap consensus tree inferred from 1000 replicates was used to represent the evolutionary history of *Streptomyces* genomes. The phylogenetic tree, based on the alignment of CYP protein sequences from *S. cavourensis* YBQ59 with related proteins, was constructed using MEGA 11.0.

Physical maps of *S. cavourensis* YBQ59, showing the distribution of cytochrome P450, ferredoxin, and ferredoxin reductase genes across the chromosome, were created using MG2C software (http://mg2c.iask.in/mg2c_v2.1/).

Identification of CYPs involved in secondary metabolite biosynthetic gene clusters

The antiSMASH bacterial version 7.0 was used to predict secondary metabolite biosynthetic gene clusters (smBGCs) present in the YBQ59 genome. Using default parameters, significant biosynthetic gene clusters were identified against MIBiG. CYPs associated with potential smBGCs of strain YBQ59 were identified and compared with reference CYPs using BLASTp and TBLASTN tools.

RESULTS AND DISCUSSION

Genomic characteristics of *S. cavourensis* YBQ59

The draft genome sequence of *S. cavourensis* YBQ59 consists of a circular chromosome of 8,126,002 bp, with an average G + C content of 72.1% (Table 1). A total of 7,020 genes was predicted, encoding 6,858 coding sequences. The chromosome contains 8 rRNAs, 65 tRNAs, and 86 pseudogenes. Notably, no plasmids were predicted. The draft genome sequence of *S. cavourensis* YBQ59 has been deposited in DDBJ/ENA/GenBank under the accession number JASEND000000000. The total genome sizes of previously reported *S. cavourensis*, such as *S. cavourensis* 1AS2a, *S. cavourensis* BUU135 and *S. cavourensis* 2BA6PG^T, were around 7.6 Mb and did not include a plasmid (Vargas et al., 2019; Tangwattanachuleeporn et al., 2021; Chong et al., 2023). This indicates that the genome size of YBQ59 is larger than that of the other strains.

Table 1. Genomic features of *S. cavourensis* YBQ59.

Features	Value
Genome size (bp)	8,126,002
G + C content (%)	72.1
Genes (total)	7,020
CDSs (coding)	6,858
rRNAs	8
tRNAs	65
ncRNAs	5
Pseudogenes	86

Cytochrome P450 complement (CYPome) of the *S. cavourensis* YBQ59

The genome analysis revealed that *S. cavourensis* YBQ59 contains 21 putative CYPs, covering approximately 0.33% of all coding sequences. According to the international P450 Nomenclature Committee Rules, these identified CYPs belong to 10 CYP families and 17 CYP subfamilies. Among them, there are 6 CYPs belonging to CYP107, 4 to CYP157, 3 to CYP154, 2 to CYP159, and 1 each in family (CYP105, CYP124, CYP156, CYP1035, CYP1046, CYP1047). CYP1046 and CYP1047 are rarely found in *Streptomyces* spp. Notably, the start codons in seven of these genes are not the typical methionine: six genes have valine triplets (gtg), including CYP1035A9, CYP157A6, CYP157C13, CYP154C3, CYP107U8 and CYP107F4, while one gene has a leucine triplet (ttg) (CYP107L28) at the 5' end of the putative open reading frame.

Table 2. Three conserved motifs of CYPs of *S. cavourensis* YBQ59: I-helix, K-helix and heme binding motif.

No.	Family	CYP names	I-helix	K-helix	Heme binding motif
1	105	CYP105D20	G^{259}HETT^{263}	E^{297}LLR^{300}	F^{363}GFGVHQCLG^{372}

2	107	CYP107F4	$G^{254}QDTT^{258}$	$E^{292}LLR^{295}$	$F^{358}GWGAHHCLG^{367}$
3		CYP107L16	$G^{230}HETT^{234}$	$E^{268}MLR^{271}$	$F^{333}GHGVHYCLG^{342}$
4		CYP107L28	$G^{242}HETT^{246}$	$E^{280}ILR^{283}$	$F^{346}GHGIHYCLG^{355}$
5		CYP107P7	$A^{247}TVNTT^{252}$	$E^{282}LLR^{285}$	$F^{346}GAGIHYCLG^{355}$
6		CYP107U8	$G^{262}FETT^{266}$	$E^{305}LLR^{308}$	$Y^{370}GHGIHYCLG^{379}$
7		CYP107BX7	$G^{229}HDST^{233}$	$E^{260}AAGR^{264}$	$L^{335}GAGAHYCLG^{344}$
8	124	CYP124G2	$G^{262}VETT^{266}$	$E^{301}IVR^{304}$	$Y^{363}GGGGPHHCLG^{373}$
9	154	CYP154A18	$G^{243}YETT^{247}$	$E^{281}TLR^{284}$	$F^{347}GHGIHFCLG^{356}$
10		CYP154C3	$G^{256}HETT^{260}$	$E^{294}TLR^{297}$	$F^{361}GHGPHICPG^{370}$
11		CYP154C4	$G^{239}HETT^{243}$	$E^{277}TLR^{280}$	$F^{343}GHGPHVCPG^{352}$
12	156	CYP156B5	Unidentified	$E^{285}DAL^{288}$	$W^{351}GAGPHVCPA^{360}$
13	157	CYP157A6	$G^{309}HQPT^{313}$	$E^{347}VLW^{350}$	$F^{410}GHGEHRCPF^{419}$
14		CYP157A7	$G^{257}HLPT^{261}$	$E^{295}VLW^{298}$	$F^{363}SHGEYRCPF^{372}$
15		CYP157B14	$A^{255}QQPT^{259}$	$E^{293}VLW^{296}$	$F^{355}SNGEHRCPY^{364}$
16		CYP157C13	$A^{260}YEST^{264}$	$E^{297}QTLW^{301}$	$F^{361}SGGPHECPG^{370}$
17	159	CYP159A5	$G^{233}GETT^{237}$	$E^{271}TLR^{274}$	$F^{345}ALGRHFCVG^{354}$
18		CYP159A7	Unidentified	$E^{241}TLR^{244}$	$F^{304}SGPADCPA^{312}$
19	1035	CYP1035A9	$A^{248}SLETT^{253}$	$E^{287}VLR^{290}$	$F^{250}GGGVHYCLG^{359}$
20	1046	CYP1046A5	$G^{253}IEAT^{257}$	$E^{309}VTR^{312}$	$Y^{381}GLGPRYCPG^{390}$
21	1047	CYP1047A3	$G^{269}HETT^{273}$	$E^{324}TLR^{327}$	$F^{397}GGGPRVCLG^{406}$

Although the primary amino acid sequences of cytochrome P450s are highly variable, their secondary and tertiary structures are generally conserved across all family members, featuring three sequence motifs. Table 2 presents the three conserved motifs of the CYPs from *S. cavourensis* YBQ59. The critical residues in the I-helix motif (T), K-helix motif (E and R), and the heme-binding domain signature (C) are highlighted in bold.

Among the three conserved motifs in CYPs, only the cysteine residue in the heme-binding motif is absolutely conserved across all CYP structures. The cysteine forms two hydrogen bonds with the neighboring backbone amide and acts as the fifth thiolate ligand to the CYP heme iron. This ligand gives CYPs their distinctive properties, as evidenced by the Soret absorbance at 450 nm observed for the ferrous-CO complex (Dawson et al., 1976). The most common amino acid sequence of the heme-binding signature is GXXXCXG, which is conserved in 16 of the 21 CYPs of *S. cavourensis* YBQ59. However, the heme-binding motif in the CYP157 family does not strictly follow the rule: GXXXCXG appears only in CYP157C13, while the last glycine is replaced by phenylalanine (F) in CYP157A6 and

CYP157A7, and by tyrosine (Y) in CYP157B14.

Like all CYPs, *Streptomyces* CYPs have an I-helix that extends along the protein molecule. The heme group, situated between the distal I- helix and proximal L-helix, is bound to the adjacent cysteine of the heme-domains (Dawson et al., 1976). The I-helix GXXTT motif, which is believed to be involved in oxygen activation, contains a highly conserved threonine (Martinis et al., 1989). This motif is present in most CYPs of *S. cavourensis* YBQ59, except CYP156B5 and CYP159A7. This characteristic is also observed in the same CYP families in *Streptomyces sp.* CFMR 7, *S. fulvissimus*, *S. griseus*, and *S. globisporus* (Senate et al., 2019).

The most common conserved sequence in the K-helix is the EXXR motif, which is present in 15 out of the 21 CYPs of *S. cavourensis* YBQ59. The remaining CYPs, including the 4 members of the CYP157 family - CYP157A6 (**E**^{347}VLW^{350}), CYP157A (**E**^{295}VLW^{298}), CYP157B1 (**E**^{293}VLW^{296}), CYP157C13 (**E**^{297}QTLW^{301}) and 1 member of the CYP156 family CYP156B5 (**E**^{285}DAL^{288}), lack the arginine residue in the K-helix. This same feature has been observed in the CYP157 family from *S. coelicolor* A3(2) (Rupasinghe et al., 2006). Glutamic acid and arginine in the K-helix motif (EXXR) were originally believed to form salt bridge interactions with the region in front of the helix containing the cysteine ligand, helping maintain the correct tertiary structure of CYP (Hasemann et al., 1995). However, later site-directed mutagenesis studies on CYP157C1 from *S. coelicolor* A3(2), which modified it to the typical EXXR motif, showed that this motif is not required for the tertiary structure of all CYPs (Rupasinghe et al., 2006).

Phyletic distribution of CYP families in representative *Streptomyces* genomes

The phylogenetic tree showed that the CYPs in *S. cavourensis* YBQ59 were highly similar to those in *Streptomyces* sp. CFMR7, a strain known for its natural rubber degradation and isolated from a rubber plantation (Nanthini et al., 2015). The genome size of *Streptomyces* sp. CFMR7 was 8,1731 bp, closely matching that of *S. cavourensis* YBQ59. Additionally, *Streptomyces* sp. CFMR7 had three more CYP enzymes compared to YBQ59, including CYP125A61, CYP134A4, and CYP177A2. Thirteen out of the seventeen CYP subfamilies from *S. cavourensis* YBQ59 - CYP105D, CYP107F, CYP107L, CYP107P, CYP107U, CYP107BX, CYP124G, CYP154C, CYP156B, CYP157A, CYP157C, CYP159A and CYP1047A - were also found in *S. fulvissimus*, *S. griseus* and *S. globisporus*, indicating that these strains may share significant similarities in the production of secondary metabolites. *S. fulvissimus* is known to produce the ionophore antibiotic valinomycin and an antibacterial protein that inhibits many bacterial strains (Myronovskyi et al., 2013). *S. griseus* is recognized for producing streptomycin, a broad-spectrum antibiotic and anticancer drug (Ohnishi et al., 2008). *S. globisporus* is characterized as a producer of the antitumor angucyclines landomycin A and landomycin E, as well as enediyne antitumor antibiotic C-1027 (Li et al., 2016). Among these *Streptomyces*, *S. griseus* has the largest genome (8,545,929 bp) and the highest number of P450 enzymes (28). *S. globisporus* has the smallest genome

(7,693,617 bp), while *S.fulvissimus* has the fewest P450 enzymes (19). A phylogenetic tree was constructed for the CYPome of *S. cavourensis* YBQ59, *Streptomyces* sp. CFMR7, *S. fulvissimus*, *S. griseus* and *S. globisporus* (Fig. 1).

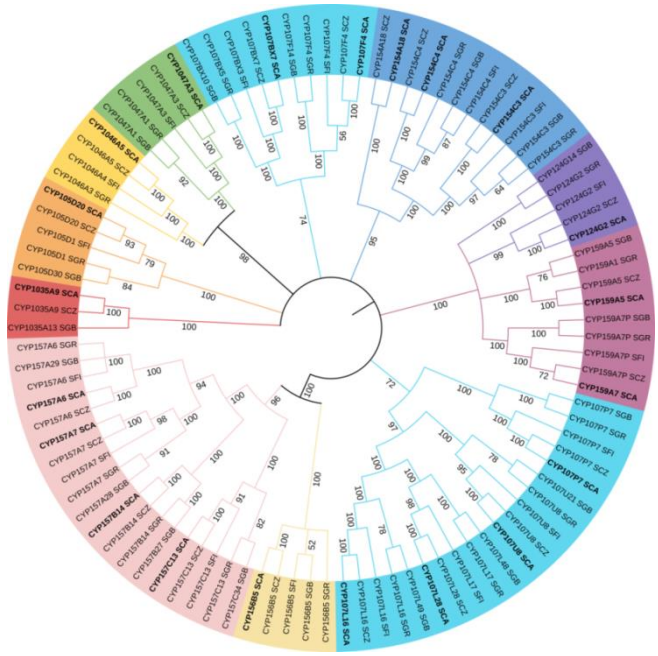

Figure 1. Phylogenetic tree of the CYPome from *S. cavourensis* YBQ59 (SCA), *Streptomyces* sp. CFMR7 (SCZ), *S. fulvissimus* (SFI), *S. griseus* (SRG), and *S. globisporus* (SRB). Numbers indicate the bootstrap probability values for the branch topology are shown.

CYP1046A is present in four *Streptomyces* but absent in *S. globisporus*. While CYP1035A9 is found in *S. cavourensis* YBQ59, *Streptomyces* sp. CFMR7 and *S. griseus*, but is absent *in S. fulvissimus* and *S. globisporus*. Additionally, three CYPs are present in *Streptomyces* sp. CFMR7 but not in *S. cavourensis* YBQ59, suggesting that these CYP families may assist *Streptomyces* sp. CFMR7 in adapting to its distinct environment. *S. avermitilis* has a notably high number of CYPs (53 enzymes) compared to other species. However, *S. avermitilis* does not have CYP1035, CYP1046 and CYP1047 in its genomes. To adapt to their environment and thrive in their ecological niche, *Streptomyces* spp. must produce a variety of metabolites, often toxic to other bacteria, allowing them to survive and make use of available carbon sources. As a result, different lifestyles and environments can alter their gene pool, influencing their CYP profiles, including variations in the number of CYPs, subfamilies, and those associated with secondary metabolite biosynthetic gene clusters.

Identification of CYP redox partners

To carry out their catalytic function, cytochrome P450 enzymes must receive electrons from NAD(P)H through an electron transfer chain (Hannemann et al., 2007). In most prokaryotic CYPs, including those found in *Streptomyces* species, the

electron transport proteins typically consist of a ferredoxin and ferredoxin reductase. The distribution of CYPs, ferredoxins and ferredoxin reductases in *Streptomyces* species varies, reflecting their roles in secondary metabolism. The genome of *S. cavourensis* YBQ59 contains six ferredoxin reductase genes and five ferredoxin genes. Similarly, *S. avermitilis* MA4680 also has six ferredoxin reductase genes but possesses four additional ferredoxin genes, totaling nine (Pandey et al., 2014). In contrast, *S. coelicolor* A3(2) has only two ferredoxin reductase genes and six ferredoxin genes (Lei et al., 2004). Unlike these species, none of the CYPs in *S. cavourensis* YBQ59 are clustered with ferredoxin reductase. The physical maps of *S. cavourensis* YBQ59, showing the distribution of cytochrome P450, ferredoxin, and ferredoxin reductase genes are presented in Fig. 2.

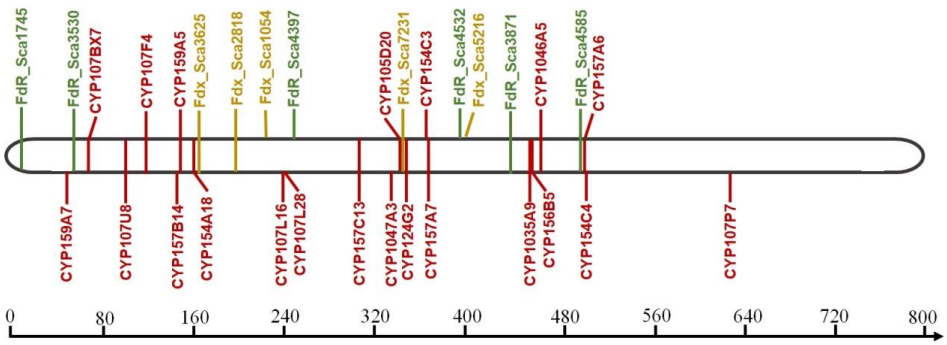

Figure 2. Physical maps of *S. cavourensis* YBQ59 illustrate the distribution of cytochrome P450, ferredoxin (Fdx), and ferredoxin reductase (FdR) genes.

In addition to autologous redox partners, P450 enzymes can also be supported by redox proteins from other sources, such as those from the human adrenal glands, including adrenodoxin (Adx) and adrenodoxin reductase (AdR). These heterologous redox partners are considered highly versatile, as they can interact with a variety of cytochrome P450 enzymes across species, including bacterial P450s (Nguyen et al., 2012; Khatri et al., 2016), making them valuable tools in heterologous systems.

CYPs involved in the production of secondary metabolites

In silico genome analysis revealed that four CYPs in *S. cavourensis* YBQ59 are located within known predicted gene clusters, suggesting their involvement in secondary metabolite production (Fig. 3). Specifically, CYP107BX7, CYP124G2, CYP154A18, and CYP105D20 belong to biosynthetic gene cluster for polycyclic tetramate macrolactams, melanin, cadaside, and o-dialkylbenzene, with homology similarities of 100%, 100%, 19% and 12%, respectively.

Besides, several CYPs are clustered together: CYP154C3 and CYP157A7 are located adjacent to each other on the same strand, with only a 3-bp separation. The genomic organization of this cluster resembles that of CYP157A1 and CYP154C1 in S. *coelicolor* A3(2) (Lei et al., 2004) as well as CYP157A2 and CYP154C2 in *S. avermitilis* MA4680 (Pandey et al., 2014). A second cluster of CYP157 and CYP154 family members in *S. cavourensis* YBQ59 is found between

CYP157A6 and CYP154C4, with an overlap of 3 bp. The gene encoding CYP1035A9 is located adjacent to the gene for CYP156B5, separated by just 164 base pairs. In *S. coelicolor* A3(2), the CYP156 family member CYP156A1 is also found in a cluster with another CYP (CYP154A1). The first reported cluster between two CYP107 family members is observed in *S. cavourensis* YBQ59, where CYP107L16 and CYP107L28 are positioned only 38 nucleotides apart. The close proximity of these two CYP suggests that they may be part of an operon and could play a role in the biosynthesis of secondary metabolites for inter-microbial interaction or detoxification. It seems that the CYP105 family in *Streptomyces* is the most frequently clustered with ferredoxins (Parajuli et al., 2004). CYP105D20 from *S. cavourensis* YBQ59 shares a high degree of homology with CYP105D20 from *Streptomyces* sp. CFMR 7, exhibiting 97.3% identity, and is arranged in an operon with a ferredoxin gene, which is likely its natural redox partner. A similar arrangement has also been observed with CYP105D5 in *S. coelicolor* A3 (2), CYP105D6 in *S. avermitilis*, CYP105D1 in *S. griseus* and CYP105D4 in *Streptomyces lividans* (Lamb et al., 2003).

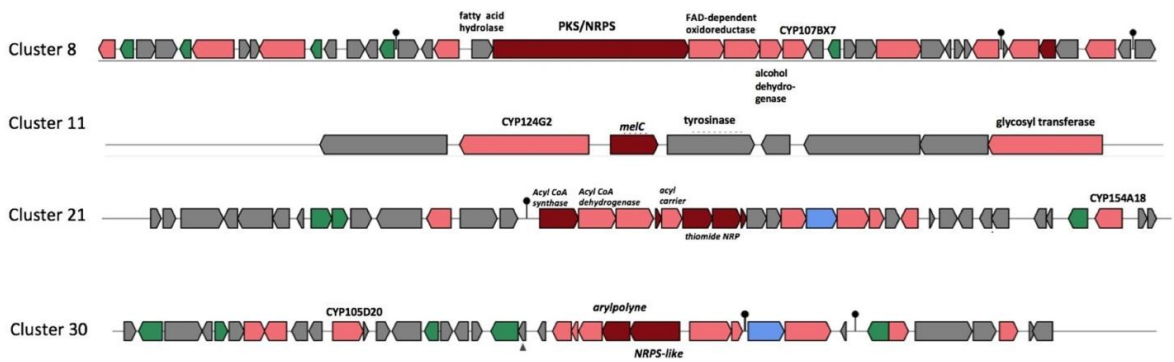

Figure 3. Secondary metabolite biosynthetic gene clusters and P450 analysis in *S. cavourensis* YBQ59.

The most numerous members belong to the CYP107 family, recognized as the dominant family across all *Streptomyces* strains. In *S. cavourensis* YBQ59, the CYP107 family comprises 28.5% of the total, featuring six members: CYP107P7, CYP107BX7, CYP107U8, CYP107L16, CYP107L28, and CYP107F4. Given their abundance among *Streptomyces* species, CYP107 enzymes are thought to play an important role in the synthesis of secondary metabolites (Mnguni et al., 2020). CYP107U1 from *S. coelicolor* A3(2) plays a crucial role in sporulation and antibiotic production in *Streptomyces coelicolor* (Tian et al., 2013). CYP107DY1 from *Bacillus megaterium* converts mevastatin to pravastatin, an important therapeutic drug for treating hypercholesterolemia (Milhim et al., 2016). CYP107X1 from *Streptomyces avermitilis* catalyzes the stereoselective hydroxylation of progesterone at the 16α position (Lin et al., 2022). In addition, a CYP107 from *Sebekia benihana* has been identified as a carbon-25 vitamin D hydroxylase (Jensen and Estrada, 2023).

The second most abundant family in the YBQ59 genome is CYP157, which consists of four enzymes: CYP157A6, CYP157A7, CYP157B14 and CYP157C13. This family is also found in the majority of the *Streptomyces* species, although reports on this enzyme family are quite limited. The third most abundant family in *S. cavourensis* YBQ59 is CYP154, which includes three members: CYP154C3, CYP154C4 and CYP154A18. Many members of the CYP154 family are known to function as steroid hydroxylases and contribute to the production of secondary metabolites (Makino et al., 2014). CYP154A1 is involved in di-pentaenone cyclization and plays a role in polyketide metabolism [Cheng et al, 2010]. CYP154C1 from *S. coelicolor* A3(2) functions as a macrolactone monooxygenase at the 12- and 14-carbon positions (Produst et al., 2003). Furthermore, CYP154E1 from *Thermobifida fusca* YX has been shown to selectively hydroxylate the allylic position of acyclic terpenoids (Bogazkaya et al., 2014).

S. cavourensis YBQ59 contains two CYP159 enzymes, but the function of this family has only few clues. Among the six remaining subfamilies, each has only one member. While CYP105, CYP124, and CYP156 are prevalent in the majority of the *Streptomyces* species, CYP1035, CYP1046 and CYP1047 are found in only few (Senate et al., 2019). CYP124 from *Mycobacterium tuberculos* is involved in the hydroxylation of cholesterol and lipids (Johnston etl al., 2009). CYP105 is one of the most abundant families in *Streptomyces* species and plays a crucial role in biotransformation and bioremediation. CYP105D6 and CYP105P1 are involved in the modification of filipin, a polyketide that is part of the 28-component polyene macrolide antifungal drug, which is used as a probe for cholesterol in biological membranes (Xu et al., 2010).

Analyzing similar cytochrome P450 enzymes from different sources can reveal subtle differences in substrate specificity and hydroxylation selectivity, often due to slight structural variations in their active sites. Such comparative studies help identify critical structural motifs that determine substrate compatibility and product formation, uncovering the evolution of enzyme function across species. These insights are valuable for understanding structure-function relationships and are especially useful in drug development and biocatalysis, where precise substrate recognition and selective hydroxylation are crucial.

CONCLUSION

The endophytic *S. cavourensis* YBQ59 contains 21 P450 enzymes, which belong to 10 CYP families and 17 CYP subfamilies. The CYPome of *S. cavourensis* YBQ59 is similar to that of *Streptomyces* sp. CFMR7, a naturally occurring rubber degrader isolated from a rubber plantation. It also shares many similarities with *S. fulvissimus*, known for producing the ionophore antibiotic valinomycin; *S. griseus*, a producer of broad-spectrum antibiotics and anticancer drugs; and *S. globisporus*, which produces the antitumor angucyclines landomycin A and landomycin E, as well as the antitumor antibiotic C-1027. Four CYPs have been identified as likely belonging to known predicted gene clusters. Furthermore, *S. cavourensis* YBQ59 contains six ferredoxin reductase genes, seven

ferredoxin genes, and three flavodoxin reductase genes. Analysis of the CYPome data from this *Streptomyces* strain has revealed numerous potential P450 enzymes that could be exploited for applications in agriculture and medicine.

ACKNOWLEDGEMENTS

This research is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 106.02-2020.29. We would like to express our gratitude to Dr. David Nelson from the University of Tennessee Health Science Center, USA, for his contributions to the nomenclature of cytochrome P450.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

Bernhardt R., 2006. Cytochromes P450 as versatile biocatalysts. *Journal of Biotechnology*, 124:128−145. https://doi.org/10.1016/j.jbiotec.2006.01.026

Bogazkaya A.M., von Bühler C.J., Kriening S., Busch A., Seifert A., Pleiss J., Laschat S., and Urlacher V.B., 2014. Selective allylic hydroxylation of acyclic terpenoids by CYP154E1 from *Thermobifida fusca* YX. *Beilstein Journal of Organic Chemistry*, 10: 1347−1353. https://doi.org/10.3762/bjoc.10.137

Cheng Q., Lamb D.C., Kelly S.L., Lei L., and Guengerich F.P., 2010. Cyclization of a cellular dipentaenone by *Streptomyces coelicolor* cytochrome P450 154A1 without oxidation/reduction. *Journal of the American Chemical Society*, 132: 15173−15175. https://doi.org/10.1021/ja107801v

Chong E.T.J., Chiang C., Png K.K., Abidueva E., Zaitseva S., Sun C., and Lee P.C., 2023. Dataset of the complete genome of *Streptomyces cavourensis* strain 2BA6PGT isolated from sediment from the bottom of the salt lake Verkhnee Beloe (Buryatia, Russia). *Data in Brief*, 46: 108877. https://doi.org/10.1016/j.dib.2022.108877

Dawson, J.H., Holm, R.H., Trudell, J.R., Barth, G., Linder, R.E., Bunnenberg, E., Djerassi, C., and Tang, S.C., 1976. Oxidized cytochrome P-450. Magnetic circular dichroism evidence for thiolate ligation in the substrate-bound form: Implications for the catalytic mechanism. *Journal of the American Chemical Society*, 98: 3707−3708. https://doi.org/10.1021/ja00428a054

Hannemann, F., Bichet, A., Ewen, K.M., and Bernhardt, R., 2007. Cytochrome P450 systems—biological variations of electron transport chains. *Biochimica et Biophysica Acta*, 1770: 330−344. https://doi.org/10.1016/j.bbagen.2006.07.017

Hasemann, C.A., Kurumbail, R.G., Boddupalli, S.S., Peterson, J.A., and Deisenhofer, J., 1995. Structure and function of cytochromes P450: a comparative analysis of three crystal structures. *Structure*, 3: 41−62. https://doi.org/10.1016/s0969-2126(01)00134-4

Huang, D.F., Xu, J.G., Liu, J.X., Zhang, H., and Hua, Q.P., 2014. Chemical constituents, antibacterial activity and mechanism of action of the essential oil from *Cinnamomum cassia* bark against four food-related bacteria. *Microbiology*, 83: 357−365. http://dx.doi.org/10.1134/S0026261714040067

Jackson, S.A., Crossman, L., Almeida, E.L., Margassery, L.M., Kennedy, J., and Dobson, A.D.W., 2018. Diverse and abundant secondary metabolism biosynthetic gene clusters in the genomes of marine sponge-derived *Streptomyces* spp. isolates. *Marine Drugs*, 16: 67. https://doi.org/10.3390/md16020067

Jensen, L. and Estrada, D.F., 2023. Characterization of a novel bifunctional cytochrome P450 enzyme CYP107 from bacterium *Sebekia benihana*. *Journal of*

Pharmacology and Experimental Therapeutics, 385(S3): 580. https://doi.org/10.1124/jpet.122.260550

Johnston, J.B., Kells, P.M., Podust, L.M., and Ortiz de Montellano, P.R., 2009. Biochemical and structural characterization of CYP124: a methyl-branched lipid omega-hydroxylase from *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences*, 106: 20687−20692. https://doi.org/10.1073/pnas.0907398106

Khatri, Y., Ringle, M., Lisurek, M., von Kries, J.P., Zapp, J., and Bernhardt, R., 2016. Substrate hunting for the myxobacterial CYP260A1 revealed new 1α-hydroxylated products from C-19 steroids. *ChemBioChem*, 17(1): 90−101. https://doi.org/10.1002/cbic.201500420

Lamb, D.C., Ikeda, H., Nelson, D.R., Ishikawa, J., Skaug, T., Jackson, C., Omura, S., Waterman, M.R., and Kelly, S.L., 2003. Cytochrome P450 complement (CYPome) of the avermectin-producer *Streptomyces avermitilis* and comparison to that of *Streptomyces coelicolor* A3(2). *Biochemical and Biophysical Research Communications*, 307: 610−619. https://doi.org/10.1016/s0006291x(03)01231-2

Lei, L., Waterman, M.R., Fulco, A.J., Kelly, S.L., and Lamb, D.C., 2004. Availability of specific reductases controls the temporal activity of the cytochrome P450 complement of *Streptomyces coelicolor* A3(2). *Proceedings of the National Academy of Sciences*, 101: 494−499. https://doi.org/10.1073/pnas.2435922100

Li, X., Lei, X., Zhang, C., Jiang, Z., Shi, Y., Wang, S., Wang, L., and Hong, B., 2016. Complete genome sequence of *Streptomyces globisporus* C-1027, the producer of an enediyne antibiotic lidamycin. *Journal of Biotechnology*, 222: 9−10. https://doi.org/10.1016/j.jbiotec.2016.02.004

Lin, S., Ma, B., Gao, Q., Yang, J., Lai, G., Lin, R., Yang, B., Han, B.N., and Xu, L.H., 2022. The 16α-hydroxylation of progesterone by cytochrome P450 107X1 from *Streptomyces avermitilis*. *ChemBioDivers*, 19: e202200177. https://doi.org/10.1002/cbdv.202200177

Makino, T., Katsuyama, Y., Otomatsu, T., Misawa, N., and Ohnishi, Y., 2014. Regio- and stereospecific hydroxylation of various steroids at the 16α position of the D ring by the *Streptomyces griseus* cytochrome P450 CYP154C3. *Applied and Environmental Microbiology*, 80: 1371−1379. https://doi.org/10.1128/aem.03504-13

Martinis, S.A., Atkins, W.M., Stayton, P.S., and Sligar, S., 1989. A conserved residue of cytochrome P-450 is involved in heme-oxygen stability and activation. *Journal of the American Chemical Society*, 111: 9252−9253. https://doi.org/10.1021/ja00208a031

Milhim, M., Putkaradze, N., Abdulmughni, A., Kern, F., Hartz, P., and Bernhardt, R., 2016. Identification of a new plasmid-encoded cytochrome P450 CYP107DY1 from *Bacillus megaterium* with a catalytic activity towards mevastatin. *Journal of Biotechnology*, 240: 68−75. https://doi.org/10.1016/j.jbiotec.2016.11.002

Mnguni, F.C., Padayachee, T., Chen, W., Gront, D., Yu, J.H., Nelson, D.R., and Syed, K., 2020. More P450s are involved in secondary metabolite biosynthesis in *Streptomyces* compared to *Bacillus*, *Cyanobacteria*, and *Mycobacterium*. *International Journal of Molecular Sciences*, 21: 4814. https://doi.org/10.3390/ijms21134814

Myronovskyi, M., Tokovenko, B., Manderscheid, N., Petzke, L., and Luzhetskyy, A., 2013. Complete genome sequence of *Streptomyces fulvissimus*. *Journal of Biotechnology*, 168: 117−118. https://doi.org/10.1016/j.jbiotec.2013.08.013

Nanthini, J., Chia, K.H., Thottathil, G.P., Taylor, T.D., Kondo, S., Najimudin, N., Baybayan, P., Singh, S., and Sudesh, K., 2015. Complete genome sequence of *Streptomyces* sp. strain CFMR 7, a natural rubber degrading

actinomycete isolated from Penang, Malaysia. *Journal of Biotechnology*, 214: 47−48. https://doi.org/10.1016/j.jbiotec.2015.09.007

Nguyen, K.T., Virus, C., Günnewich, N., Hannemann, F., and Bernhardt, R., 2012. Changing the regioselectivity of a P450 from C15 to C11 hydroxylation of progesterone. *ChemBioChem*, 13(8): 1161−1166. https://doi.org/10.1002/cbic.201100811

Nguyen, Q.H., Vu, N.T.H., Ha, H.C., Son, K.C., Ha, H., Thanh, T.T., Cuong, N., Linh, D.T.M., Hien, T.T.T., and Tien, Q.P., 2018. Draft genome sequence of *Streptomyces cavourensis* YBQ59, an endophytic producer of antibiotics bafilomycin D, nonactic acid, prelactone B, and 5,11-epoxy-10-cadinanol. *Microbiology Resource Announcements*, 7: e01056−18. https://doi.org/10.1128/mra.01056-18

Ohnishi, Y., Ishikawa, J., Hara, H., Suzuki, H., Ikenoya, M., Ikeda, H., Yamashita, A., Hattori, M., and Horinouchi, S., 2008. Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. *Journal of Bacteriology*, 190: 4050−4060. https://doi.org/10.1128/jb.00204-08

Pandey, B.P., Lee, N., Choi, K.Y., Kim, J.N., Kim, E.J., and Kim, B.G., 2014. Identification of the specific electron transfer proteins, ferredoxin, and ferredoxin reductase, for CYP105D7 in *Streptomyces avermitilis* MA4680. *Applied Microbiology and Biotechnology*, 98: 5009−5017. https://doi.org/10.1007/s00253-014-5525-x

Parajuli, N., Basnet, D.B., Chan Lee, H., Sohng, J.K., and Liou, K., 2004. Genome analyses of *Streptomyces peucetius* ATCC 27952 for the identification and comparison of cytochrome P450 complement with other *Streptomyces*. *Archives of Biochemistry and Biophysics*, 425: 233−241. https://doi.org/10.1016/j.abb.2004.03.011

Podust, L.M., Kim, Y., Arase, M., Neely, B.A., Beck, B.J., Bach, H., Sherman, D.H., Lamb, D.C., Kelly, S.L., and Waterman, M.R., 2003. The 1.92-Å structure of *Streptomyces coelicolor* A3(2) CYP154C1. A new monooxygenase that functionalizes macrolide ring systems. *Journal of Biological Chemistry*, 278: 12214−12221. https://doi.org/10.1074/jbc.M212210200

Rudolf, J.D., Chang, C.Y., Ma, M., and Shen, B., 2017. Cytochromes P450 for natural product biosynthesis in *Streptomyces*: sequence, structure, and function. *Nature Product Reports*, 34: 1141−1172. https://doi.org/10.1039/c7np00034k

Rupasinghe, S., Schuler, M.A., Kagawa, N., Yuan, H., Lei, L., Zhao, B., Kelly, S.L., Waterman, M.R., and Lamb, D.C., 2006. The cytochrome P450 gene family CYP157 does not contain EXXR in the K-helix reducing the absolute conserved P450 residues to a single cysteine. *FEBS Letters*, 580: 6338−6342. https://doi.org/10.1016/j.febslet.2006.10.043

Senate, L.M., Tjatji, M.P., Pillay, K., Chen, W., Zondo, N.M., Syed, P.R., Mnguni, F.C., Chiliza, Z.E., Bamal, H.D., Karpoormath, R., Khoza, T., Mashele, S.S., Blackburn, J.M., Yu, J.H., Nelson, D.R., and Syed, K., 2019. Similarities, variations, and evolution of cytochrome P450s in *Streptomyces* versus *Mycobacterium*. *Scientific Reports*, 9: 3962. https://doi.org/10.1038/s41598-019-40646-y

Tangwattanachuleeporn, M., Ruangsuj, P., Yamprayoonswat, W., Sittihan, S., Jumpathong, W., and Yasawong, M., 2021. Genome sequence of *Streptomyces cavourensis* BUU135, isolated from soil from a tropical fruit farm in Thailand. *Microbiology Resource Announcements*, 10: e01428−20. https://doi.org/10.1128/mra.01428-20

Tian, Z., Cheng, Q., Yoshimoto, F.K., Lei, L., Lamb, D.C., and Guengerich, F.P., 2013. Cytochrome P450 107U1 is required for sporulation and antibiotic production in *Streptomyces coelicolor*. *Archives of Biochemistry and Biophysics*, 530: 101−107. https://doi.org/10.1016/j.abb.2013.01.001

Vargas, H.H.A., Santos, S.N., Padilla, G., and Melo, I.S., 2019. Genome sequence of *Streptomyces cavourensis* 1AS2a, a rhizobacterium isolated from the Brazilian Cerrado Biome. *Microbiology Resource Announcements*, 8: e00065−19. https://doi.org/10.1128/mra.00065-19

Vu, H.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* Prels in Yen Bai province of Vietnam. *Current Microbiology*, 75: 1247−1255. https://doi.org/10.1007/s00284-018-1517-x

Watve, M.G., Tickoo, R., Jog, M.M., and Bhole, B.D., 2001. How many antibiotics are produced by the genus *Streptomyces*? *Archives of Microbiology*, 176: 386−390. https://doi.org/10.1007/s002030100345

Xu, L.H., Fushinobu, S., Takamatsu, S., Wakagi, T., Ikeda, H., and Shoun, H., 2010. Regio- and stereospecificity of filipin hydroxylation sites revealed by crystal structures of cytochrome P450 105P1 and 105D6 from *Streptomyces avermitilis*. *Journal of Biological Chemistry*, 285: 16844−16853. https://doi.org/10.1074/jbc.M109.092460