



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

**APLICACIÓN DE TÉCNICAS DE MINERÍA GENÓMICA PARA LA DETECCIÓN Y
EXPRESIÓN DE UN CLÚSTER QUE CODIFIQUE PARA LA SÍNTESIS DE UN
COMPUESTO CON ACTIVIDAD ANTICANCERÍGENA**

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
Maestro en Ciencias**

**PRESENTA:
MARIA LUISA CASANOVA RIVERO**

**TUTOR PRINCIPAL
DR. SERGIO SÁNCHEZ ESQUIVEL
[Instituto de Investigaciones Biomédicas](#)**

**MIEMBROS DEL COMITÉ TUTOR
DRA. MARICARMEN QUIRASCO BARUCH
[Facultad de Química](#)**

**DRA. LETICIA ROCHA ZAVALETA
[Instituto de Investigaciones Biomédicas](#)**

Ciudad de México. Diciembre, 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Al CONACYT (Consejo Nacional de Ciencia y Tecnología) por la beca otorgada para la realización de la maestría.

Al la DGAPA y al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) UNAM por financiar el proyecto IN205922.

Al Dr. Sergio Sánchez Esquivel por aceptarme como parte de su equipo de trabajo, por su confianza, consejos y apoyo en la realización del presente proyecto.

A la Dra. Beatriz Ruiz Villafán por su asesoría técnica, paciencia y comentarios que fueron fundamentales para el término de este proyecto.

A la Dra. Silvia Ivonne Mora Herrera por su ayuda con los ensayos de UPLC-MS, desarrollados en la Unidad de Metabolómica y Proteómica del Instituto de Investigaciones Biomédicas, así como por la ayuda en el análisis de los resultados obtenidos.

A Karen González Sánchez por su amistad y trabajo realizado durante su servicio social que fue clave para poder culminar este proyecto.

A mi comité tutor, la Dra. Leticia Rocha Zavaleta y la Dra. Maricarmen Quirasco Baruch por sus comentarios y sugerencias durante el desarrollo de este proyecto.

A la M. en C. María Paula Gómez Román por su asesoría, consejos y apoyo en el manejo de la biblioteca genómica.

A mis compañeros del laboratorio C-106, Karen, Susana, Valerie, Onely, Carlos, Dulce, Ximena, Rodrigo, Alma, Fernando, Berenice por sus consejos técnicos y porque de todos he aprendido algo que me llevaré conmigo por siempre.

Dedicatoria

Al Dr. Sergio Sánchez Esquivel por su orientación, comprensión y apoyo en cada situación no prevista ocurrida en el transcurso de este proyecto, así como por permitirme seguir con mi desarrollo profesional formando parte del laboratorio C-106.

A mis padres, Martha y Julián y a mi hermana Mariana porque de ellos he aprendido que el esfuerzo es lo que hará posible alcanzar mis metas y por apoyarme durante cada etapa de mi vida, y ante cada dificultad. Por impulsarme a seguir con mi desarrollo académico y por las enseñanzas de vida.

A mi tía Isabel, porque siempre estuvo presente en cada momento importante de mi vida y sé que estaría presente y orgullosa de verme culminar este proyecto.

A Luis, por ser mi amigo y compañero estos años y recordarme que la solución siempre es más simple de lo que parece y que puedo lograr lo que me proponga.

Contenido

Contenido	4
Índice de figuras	7
Índice de tablas	9
RESUMEN	13
ABSTRACT	14
1. Introducción.....	15
2. Antecedentes	17
2.1 Productos naturales microbianos como anticancerígenos.....	17
2.1 Antraciclinas y su modo de acción.....	18
2.2 Actinobacterias	20
2.3 Microorganismos endófitos	21
2.4 Aislamiento y actividad biológica de microorganismos endófitos de <i>Amphipterygium adstringens</i>	21
2.5 La familia de las estefimicinas y sus derivados	22
2.6 Policétidos	24
2.6.1 Policétido sintasas tipo I.....	25
2.6.2 Policétido sintasas tipo II.....	26
2.6.3 Policétido sintasas tipo III.....	27
2.7 Minería genómica	28
2.8 Herramientas bioinformáticas	28
2.9 Expresión heteróloga de clústeres biosintéticos.....	29
3. JUSTIFICACIÓN	31
4. HIPÓTESIS	31
5. OBJETIVO GENERAL.....	32
6. OBJETIVOS PARTICULARES.....	32
7. METODOLOGÍA.....	33
7.1 Bioinformática	33
7.1.1 Selección de una actinobacteria para la búsqueda de PKS	33
7.1.2 Análisis <i>in silico</i> de los BGCs de PKS.....	34
7.1.2.1 Determinación de módulos de la PKS y la función sugerida	34

7.1.2.2 Búsqueda del sitio catalítico	34
7.1.3 Selección de BGC para su estudio	34
7.2 Experimental	35
7.2.1 Expresión heteróloga del BGC	35
7.2.1.1 Diseño de primers y estrategia de expresión	35
7.2.1.2 Escaneo por PCR de una biblioteca genómica en la búsqueda del BGC	36
7.2.1.3 Secuenciación del fósido de la biblioteca genómica	38
7.2.2 Producción, purificación y evaluación de la actividad antimicrobiana	39
7.2.2.1 Fermentación	39
7.2.2.2 Ensayos de antibiosis y bioautografía	39
7.2.2.3 Purificación y caracterización del policétido	40
7.2.2.4 Evaluación de la actividad del policétido	42
8. RESULTADOS Y DISCUSIÓN	43
8.1 Minería genómica	43
8.2 Análisis bioinformático de la región 1.4.....	45
8.2.1 Dominio KS	46
8.2.2 Dominio AT	47
8.2.3 Dominio ACP.....	49
8.2.4 Dominio DH.....	50
8.2.5 Dominio ER.....	50
8.2.6 Dominio KR.....	51
8.2.7 Dominio TE	52
8.2.8 Función putativa de la región 1.4	53
8.2.9 Expresión heteróloga de la región 1.4.....	54
8.3 Análisis bioinformático de la región 1.12	56
8.3.1 Expresión heteróloga del BGC 1.12	59
8.4 Escaneo de la biblioteca genómica de <i>Embleya</i> sp. NF3.....	59
8.4.1 PCR de colonia para la búsqueda del BGC 1.12.....	61
8.5 Vector de expresión PIJ10257	65
8.6 Producción y purificación del policétido.....	65
8.7 Mejora en la solubilidad del policétido y evaluación de su actividad.....	72
9. CONCLUSIÓN	74
10. PERSPECTIVAS.....	75

11. REFERENCIAS.....	75
12. ANEXOS	84
12.1 Anexo 1. Medios de cultivo y soluciones	84
12.2 Anexo 2. Secuencia de primers.	86
12.3 Anexo 3. Vectores utilizados.....	87
12.4 Anexo 4. Protocolo para transformación de células electrocompetentes. .	88
12.5 Anexo 5. Aislamiento de plásmido de bajo número de copias.....	88
12.6 Anexo 6. Minipreparación de plásmido de <i>E. coli</i> de la biblioteca genómica de <i>Embleya</i> sp. NF3.	89
12.7 Anexo 7. Protocolo para limpieza de ARN.....	89
12.8 Anexo 8. Protocolo PCR de colonia.....	90

Índice de figuras

Figura 1. Ranking nacional de cáncer como causa de muerte en edades <70 años en 2019..	15
Figura 2. Clasificación de las proteínas y péptidos anticancerígenos que han sido descritos.....	18
Figura 3. Fármacos anticancerígenos descritos de 1981-2019.....	20
Figura 4. Estructura de la doxorubicina (1) y daunorrubicina (2).....	19
Figura 5. Dominios relevantes de las antraciclinas para la unión al ADN y topoisomerasa II.....	19
Figura 6. Diferencias en la estructura química de estefimicina y estefimicina B...	23
Figura 7. Esquema de los módulos en una PKS del tipo I.....	25
Figura 8. Esquema del funcionamiento de una PKS del tipo II.....	26
Figura 9. Estrategia experimental.....	33
Figura 10. Región 1.4 correspondiente a PKS del tipo I en <i>Embleya</i> sp. NF3.	45
Figura 11. Dominios presentes en los módulos de la PKS I de la región 1.4.	46
Figura 12. Estructuras previstas por SeMPI 2.0 para la región 1.4.	46
Figura 13. Alineamientos múltiples del dominio KS de la PKS 1.4.	47
Figura 14. Modelado por homología del dominio KS por SWISS-MODEL.	47
Figura 15. Alineamientos múltiples del dominio AT de la PKS 1.4.	48
Figura 16. Modelado por homología del dominio KS por SWISS-MODEL.	48
Figura 17. Alineamientos múltiples del dominio ACP de la PKS 1.4..	49
Figura 18. Modelado por homología del dominio ACP por SWISS-MODEL.....	49
Figura 19. Alineamientos múltiples del dominio DH de la PKS 1.4.....	50
Figura 20. Modelado por homología del dominio DH por SWISS-MODEL.....	50
Figura 21. Alineamientos múltiples del dominio ER de la PKS 1.4.....	51
Figura 22. Alineamientos múltiples del dominio KR de la PKS 1.4.....	51
Figura 23. Modelado por homología del dominio KR por SWISS-MODEL.	52
Figura 24. Alineamientos múltiples del dominio TE de la PKS 1.4.	52
Figura 25. Modelado por homología del dominio KR por SWISS-MODEL.	53
Figura 26. Región 1.12 correspondiente a PKS del tipo II en <i>Embleya</i> sp. NF3...	56
Figura 27. Alineamientos múltiples del dominio KS de la PKS 1.12.	57
Figura 28. Alineamientos múltiples del dominio ACP de la PKS 1.12.	58
Figura 29. Regiones amplificadas por los primers diseñados.....	60
Figura 30. Identificación de stefFinal en pools de la biblioteca genómica.	60
Figura 31. Identificación de stefInicio en pools de la biblioteca genómica.....	60
Figura 32. Identificación de stefMedio en pools de la biblioteca genómica.	61
Figura 33. Identificación de stefInicio en las colonias del pool 21A1 de la biblioteca genómica.....	62

Figura 34. Identificación de <i>stefMedio</i> en las colonias del <i>pool</i> 21A1 de la biblioteca genómica.....	62
Figura 35. Identificación de <i>stefFinal</i> en las colonias del <i>pool</i> 21A1 de la biblioteca genómica.....	62
Figura 36. Fósido pCC2FOS de la biblioteca de <i>Embleya</i> sp. NF3 de las colonias A8, A9 y A12	63
Figura 37. Identificación de <i>stefInicio</i> para cinco colonias y <i>stefMedio</i> para cuatro colonias	64
Figura 38. Identificación de <i>stefFinal</i> para cinco colonias y <i>stefMedio</i> de una colonia.....	64
Figura 39. Resultado de la purificación del vector de expresión PIJ10257..	65
Figura 40. CCF de las fracciones 50:50, 30:70 y metanol.....	68
Figura 41. Bioautografía.	69
Figura 42. CCF del compuesto recuperado por CCF preparativa.	70
Figura 43. Masas presentes en el tiempo de retención 1.20 min.	71
Figura 44. Composición elemental de la muestra obtenida mediante LC- MS.	71
Figura 45. Dimensiones y estructura de la β -ciclodextrina	72

Índice de tablas

Tabla 1. Volúmenes de reactivo para la PCR.	36
Tabla 2. Condiciones de la PCR para la amplificación de las regiones del BGC 1.12	37
Tabla 3. Gradiente de solventes agua-metanol.	40
Tabla 4. Metabolitos secundarios identificados por antiSMASH 6.0.	43
Tabla 5. Regiones identificadas en <i>Embleya</i> sp. NF3.....	44
Tabla 6. Nomenclatura para identificar los dominios en PKS.	45
Tabla 7. Actividades biológicas previstas por PASS.....	53
Tabla 8. Resultado de predicción de líneas celulares cancerosas.....	54
Tabla 9. Comparación de los genes dentro del BGC 1.12 de <i>Embleya</i> sp. NF3 y <i>S. steffisburgensis</i>	56
Tabla 10. Regiones encontradas en los pooles evaluados de la biblioteca genómica de <i>Embleya</i> sp. NF3.....	61
Tabla 11. Resultados antibiosis del medio de cultivo.....	66
Tabla 12. Resultados antibiosis primera columna con amberlita.	66
Tabla 13. Resultados antibiosis segunda columna con amberlita.....	67
Tabla 14. Resultados antibiosis complejos con estefimicina B.	72
Tabla 15. Resultados CMI complejos con estefimicina B.	73
Tabla 16. Medio Terrific-Broth.....	84
Tabla 17. Medio LB.....	84
Tabla 18. Medio YMG líquido suplementado con CaCO ₃	84
Tabla 19. Medio YMG líquido.....	85
Tabla 20. Solución I.	85
Tabla 21. Solución II.	85
Tabla 22. Solución III.	85
Tabla 23. Buffer TE (Tris-EDTA).....	86
Tabla 24. Solución 0.5 de McFarland.	86
Tabla 25. Secuencias de primers diseñados para el escaneo de la biblioteca genómica.....	86
Tabla 26. Lista de reactivos y concentración para la PCR de colonia.	90

Lista de abreviaturas

μg	Microgramo
μL	Microlitro
ACP	Proteína acarreadora de acilo
ADN	Ácido desoxirribonucleico
antiSMASH	Antibiotics and Secondary Metabolite Analysis Shell
AR	Aromatasa
ARN	Ácido ribonucleico
AT	Aciltransferasa
BGC	Clúster de genes biosintéticos
CCF	Cromatografía en capa fina
CD	Ciclodextrina
CLC-Pred	Cell Line Cytotoxicity Predictor
CLF	Factor de elongación
CLF	Factor de elongación de cadena
CYC	Ciclasa
DH	Deshidratasa
DMSO	Dimetilsulfóxido
dNTP	Desoxinucleósido trifosfato
ER	Enoil reductasa
G+C	Guanina y Citosina
K_2HPO_4	Fosfato dipotásico
KH_2PO_4	Fosfato monopotásico
KR	Cetoreductasa

KS	Cetosintasa
MAT	Metiltransferasa
Mb	10 ⁶ pares de bases
Mg	Magnesio
MgCl ₂	Cloruro de magnesio
MH	Muller-Hinton
MIC	Concentración mínima inhibitoria
min	Minutos
mL	Mililitros
mM	Milimolar
ng	Nanogramo
nm	Nanómetro
nt	Nucleótidos
OMS	Organización mundial de la salud
PASS	Prediction of Activity Spectra for Substances
pb	Pares de bases
PCR	Reacción en cadena de la polimerasa
PK	Policétido
PKS	Policétido sintasas
pM	Picomolar
Rf	Factor de retención
rpm	Revoluciones por minuto
s	Segundos
SMILES	Simplified Molecular Input Line Entry Specification (especificación de introducción linear molecular simplificada)

TE Tioesterasa

ORF Open Reading Frame (Marco de lectura abierto)

RESUMEN

De acuerdo con la Organización Mundial de la Salud el cáncer es la segunda causa de muerte en las Américas y una de las principales a nivel mundial. Debido a esto, tanto la industria farmacéutica como las instituciones académicas se encuentran en la búsqueda constante de compuestos activos que ayuden al tratamiento de esta enfermedad. Los microorganismos endófitos han sido estudiados como una fuente de búsqueda de compuestos bioactivos y gracias al avance de la tecnología ahora se cuenta con técnicas como la minería genómica que implica el estudio del genoma de microorganismos secuenciados para la identificación de clústeres de genes biosintéticos (BGC) de productos naturales.

Al ser la minería genómica una metodología que presenta ventajas sobre el método tradicional de búsqueda de metabolitos nuevos en el presente trabajo se realizó un análisis bioinformático del genoma de tres actinobacterias endófitas, eligiendo a *Embleya* sp. NF3 como la más prometedora. Se analizaron los BGC 1.4 y 1.12 de esta bacteria que se encuentran posiblemente asociados a la producción de policétidos con actividad anticancerígena, y se hizo su caracterización *in silico* para determinar la posible funcionalidad de cada uno al realizar su expresión heteróloga.

De acuerdo con los antecedentes y análisis realizados es muy probable que el BGC 1.12 seleccionado sea el responsable de producir a la estefimicina B, antibiótico de la familia de las estefimicinas con capacidad anticancerígena. Con el objetivo de aumentar la producción del metabolito potencialmente asociado al BGC 1.12 mediante expresión heteróloga se exploró una biblioteca genómica de *Embleya* sp. NF3 y se identificó una colonia positiva que contiene los genes de este clúster.

Además, se optimizó el método de purificación de la estefimicina B probablemente asociado al BGC 1.12, utilizando la resina de intercambio iónico AmberLite™ XAD™ 7HP como auxiliar para la recuperación del compuesto en medio de fermentación. Se obtuvo un rendimiento de 99.3 mg/mL, superior a los rendimientos reportados previamente y se verificó su identidad mediante LC-MS. Por último, con el fin de solventar el problema de solubilidad del compuesto se hicieron complejos ciclodextrina β - estefimicina B y se evaluó su actividad antimicrobiana *in vitro* utilizando DMSO como control. El halo de inhibición generado por el complejo con ciclodextrina fue 0.61 mm mayor que el del control en cultivos de *Micrococcus luteus*. Las concentraciones mínimas inhibitorias con *M. luteus* y *Bacillus subtilis* fueron menores para el complejo con ciclodextrina.

ABSTRACT

According to the World Health Organization, cancer is the second leading cause of death in the Americas and one of the leading causes worldwide. Due to this, both the pharmaceutical industry and academic institutions are constantly searching for active compounds that help treat this disease. Endophytic microorganisms have been studied as a search source for bioactive compounds and, thanks to advances in technology, techniques such as genomic mining are now available, which involves the study of the genome of sequenced microorganisms for the identification of biosynthetic gene clusters (BGC) of natural products.

As genomic mining is a methodology that has advantages over the traditional method of searching for new metabolites, in this work a bioinformatic analysis of the genome of three endophytic actinobacteria was carried out, choosing *Embleya* sp. NF3 is the most promising. The BGC 1.4 and 1.12 of this bacterium, which is possibly associated with the production of polyketides with anticancer activity, were analyzed and their *in silico* characterization was carried out to determine the possible functionality of each one when carrying out their heterologous expression.

According to the background and analyses carried out, it is very likely that the selected BGC 1.12 is responsible for producing steffimycin B, an antibiotic of the steffimycin family with anticancer capacity. To increase the production of the metabolite associated with BGC 1.12 through heterologous expression, a genomic library of *Embleya* sp. NF3 and a positive colony containing genes from this cluster were identified.

In addition, the purification method for steffimycin B, probably associated with BGC 1.12, was optimized, using the AmberLite™ XAD™ 7HP ion exchange resin as an aid for the recovery of the compound in the fermentation medium. A yield of 99.3 mg/mL was obtained, higher than previously reported yields, and its identity was verified by LC-MS. Finally, to solve the compound solubility problem, β - steffimycin B cyclodextrin complexes were made and their *in vitro* antimicrobial activity was evaluated using DMSO as a control. The inhibition zone generated by the complex with cyclodextrin was 0.61 mm greater than that of the control in cultures of *Micrococcus luteus*. The minimum inhibitory concentrations with *M. luteus* and *Bacillus subtilis* were lower for the complex with cyclodextrin.

1. Introducción

El cáncer se posiciona como la causa principal de muertes y una barrera muy importante para el aumento de la expectativa de vida mundial. Según estimaciones de la OMS, en 2019 fue la primera o segunda causa de muerte en personas menores de 70 años en 112 de las 183 naciones miembro, y en las demás se coloca entre la tercera y cuarta causa principal de muerte (Sung et al., 2021) (Figura 1).

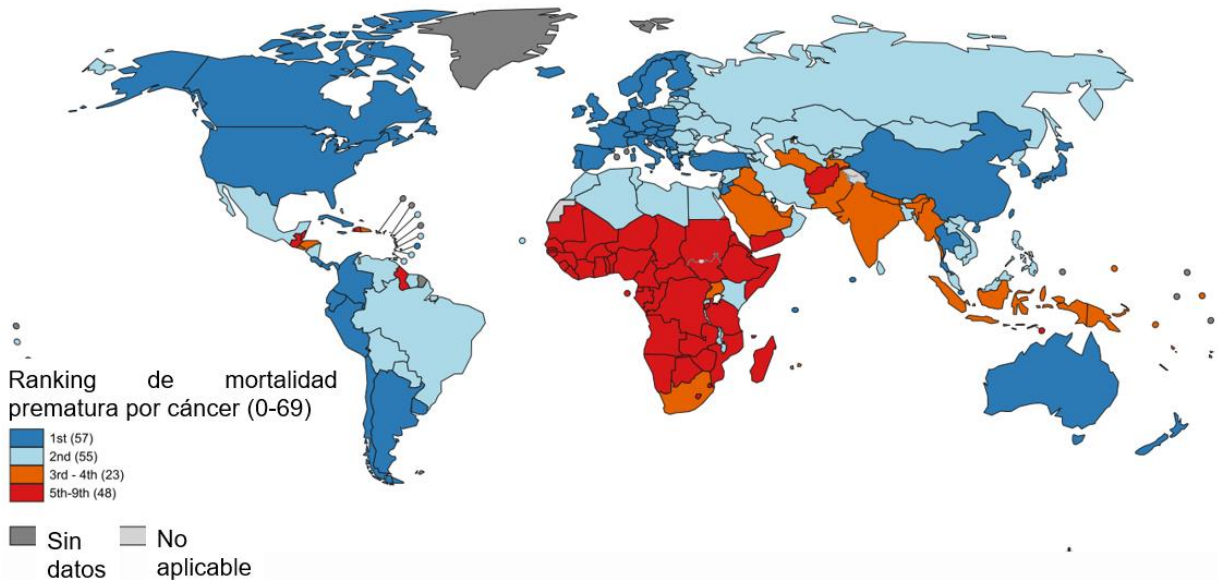


Figura 1. Ranking nacional de cáncer como causa de muerte en edades <70 años en 2019. En la leyenda se incluye el número de países representados en cada grupo de ranking. (Modificado de: Sung et. al., 2020).

En México el cáncer es actualmente la tercera causa de muerte, el 45.3% de las muertes ocurren en la población económicamente activa. México es un país con alta prevalencia de factores de riesgo asociados a esta enfermedad como el tabaquismo y el uso de alcohol. De acuerdo con datos del 2011 sobre adicciones, la prevalencia del consumo activo de tabaco es del 21.7% en el rango de edad de 12-65 años y de 12.3% en adolescentes. La prevalencia del alcoholismo de 12-64 años es del 51.4% (Mohar-Betancourt et al., 2017).

De entre 36 tipos diferentes de cáncer, el cáncer de mama ocupa el primer lugar en muertes y número de casos nuevos con el 11.7% del total de estos. Le siguen el cáncer de pulmón, colorrectal, próstata y estómago. En cuanto al diagnóstico respecto a las mujeres, el cáncer de mama, seguido por el cáncer cervicouterino, son la causa principal de muerte en 110 y 26 países respectivamente, seguido por el cáncer de pulmón en 25 naciones (Sung et al., 2021).

Para el tratamiento de esta enfermedad, los métodos tradicionales incluyen cirugía, radioterapia y quimioterapia; dependiendo del tipo de cáncer se elige el tratamiento más adecuado o una combinación de varios de ellos (Ma & Kwok, 2020).

El término quimioterapia fue dado por el químico Paul Ehrlich a principios de 1900s, y se refiere al uso de sustancias químicas para tratar enfermedades (Sharifi-Rad et al., 2019). Estas moléculas bioactivas actúan como herramientas reguladoras para la represión de oncogenes y el ciclo celular (Guimarães et al., 2020) .

Es conveniente destacar que, independientemente de la numerosa cantidad de fármacos que han sido utilizados y el éxito obtenido en varios programas de manejo, las terapias no han dado el nivel esperado de resultados y esto se ve reflejando en la reincidencia tumoral y el inicio de metástasis (Sharifi-Rad et al., 2019). La eficiencia del tratamiento con quimioterapia es cuestionable a causa del incremento constante de la resistencia de los tumores malignos hacia los fármacos antineoplásicos. Esta resistencia se clasifica en dos categorías: la intrínseca, causada por alteraciones genéticas en los tumores que ayudan a la formación de diferentes factores de resistencia, conduciendo a una terapia ineficiente, incluso antes de la aplicación de medicamentos. La resistencia adquirida puede ocurrir durante el transcurso del tratamiento

Por otro lado, está la adquirida, donde estas alteraciones genéticas surgen como respuesta adaptativa al fármaco. Si bien los estudios se han concentrado en la genética y biología molecular del cáncer, la búsqueda y desarrollo de nuevos medicamentos sigue siendo un objetivo principal para el tratamiento de esta enfermedad (Guimarães et al., 2020). Estos compuestos novedosos deben ser más selectivos, con menos efectos secundarios, rentables, con mayores atributos medicinales y tener un nivel mínimo de resistencia a la enfermedad (Sharifi-Rad et al., 2019).

2. Antecedentes

2.1 Productos naturales microbianos como anticancerígenos

Al descubrir un nuevo compuesto, las primeras pruebas biológicas que se realizan son antimicrobianas y antiinflamatorias a fin de procesarlos para uso farmacéutico. Si el compuesto es fuertemente bioactivo en una prueba en particular se mide su toxicidad con diferentes líneas celulares para confirmar la seguridad farmacéutica (Bhattacharai et al., 2020). La mayoría de los agentes antitumorales clínicamente efectivos de microbios fueron desarrollados en los 60's mediante el análisis de compuestos activos contra líneas celulares tumorales que eran demasiado tóxicos para su uso como drogas anti infectivas (Lancini & Demain, 2013), es por esta razón que la mayoría de los compuestos importantes utilizados para la quimioterapia de tumores son antibióticos producidos por microorganismos y suelen tener una actividad dual, es decir tienen actividad antimicrobiana y antitumoral .

Entre los agentes antitumorales de microorganismos tenemos a la actinomicina D, las antraciclinas sobre las cuales se profundizara más adelante, los glicopéptidos (bleomicina), mitomicina C, entre otros (Demain & Vaishnav, 2011). En la Figura 2 se muestran algunos de los diferentes péptidos y proteínas bacterianas descritos como agentes anticancerígenos, algunos de ellos se encuentran aún en fase clínica para probar su eficacia terapéutica.

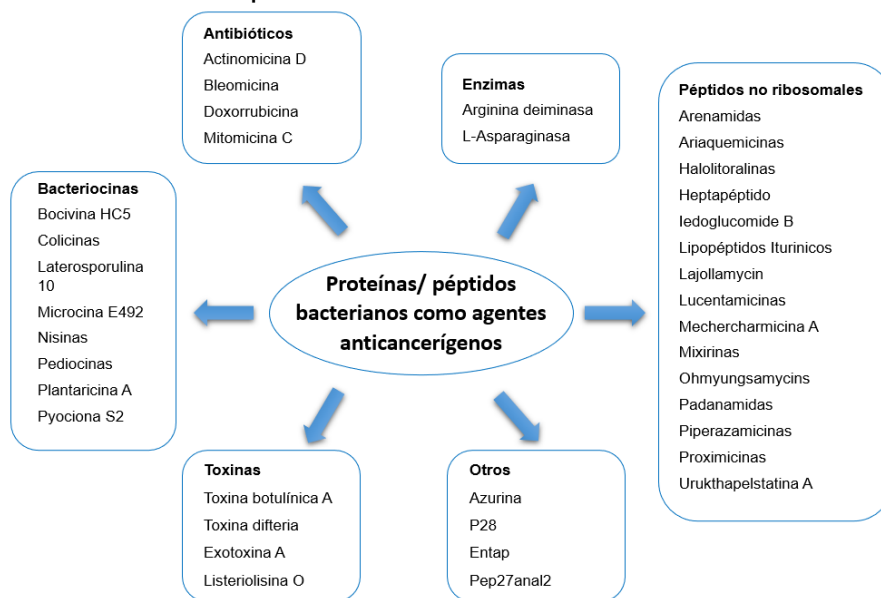


Figura 2. Clasificación de las proteínas y péptidos anticancerígenos que han sido descritos (Modificado de Karpiński & Adamczak, 2018).

Es evidente que los productos naturales son una fuente importante para el desarrollo de fármacos anticancerígenos, ya que de los 247 compuestos químicos

con actividad anticancerígena que han sido aprobados para su uso, una gran parte provienen de fuentes naturales o son derivados de estas, como se observa en la Figura 3 (Newman & Cragg, 2020).

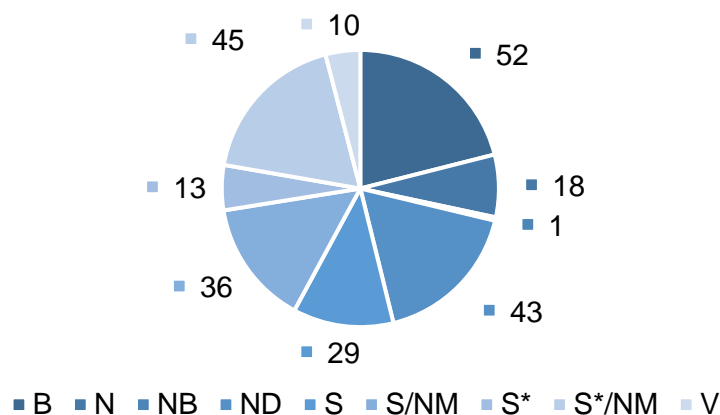


Figura 3. Fármacos anticancerígenos descritos de 1981-2019.

B: Biológico; generalmente un péptido o proteína ya sea aislado de un organismo / línea celular o producido por medios biotecnológicos en un huésped sustituto. N: Producto natural. NB: Producto natural “Botánico”. NM: imita a un producto natural. ND: Derivado de un producto natural y suele ser una modificación semisintética. S: droga totalmente sintética. S*: Elaborado por síntesis total, pero el farmacóforo es / era de un producto natural. V: vacuna. (Modificado de: Newman & Cragg, 2020).

2.1 Antraciclinas y su modo de acción

Las antraciclinas se posicionan entre los fármacos anticancerígenos más efectivos jamás desarrollados. La primera antraciclina descrita, la daunorrubicina, fue aislada de una actinobacteria obtenida de una muestra de suelo encontrada en Italia en 1960, es un antibiótico pigmentado producido por *Streptomyces peucetius*. En 1969 su homóloga, la doxorrubicina fue aislada de una mutante químicamente inducida de *S. peucetius* (Van der Zanden et al., 2021).

Estos fármacos desempeñan un papel significativo en el tratamiento de cáncer, ya que un 57-70% de pacientes con linfoma, 32% con cáncer de mama y 50-60% de niños con cáncer sobrevivientes, son tratados con un régimen de antraciclinas (Martins-Teixeira & Carvalho, 2020).

En la Figura 4 se muestra la estructura que presentan de la doxorrubicina (1) y la daunorrubicina (2). Estructuralmente son glucósidos de productos naturales derivados de poliketidos microbianos, comparten un esqueleto de polihidroxi-antraquinona (anillos) fusionado con un cuarto anillo saturado, que en conjunto corresponde al resto de la aglicona. La cadena lateral es diferente entre

ambas porque la primera tiene un grupo hidroxilo adicional (Martins-Teixeira & Carvalho, 2020).

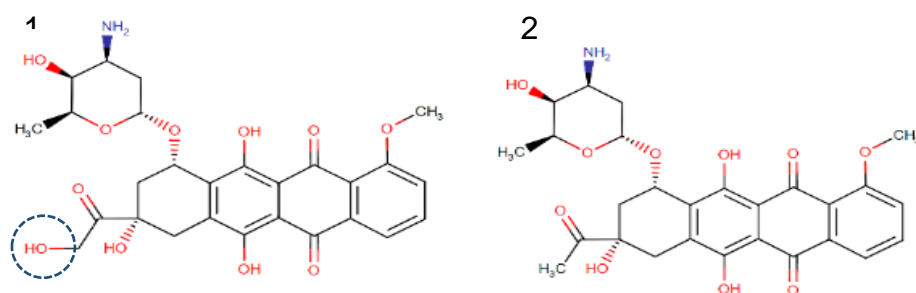


Figura 4. Estructura de la doxorubicina (1) y daunorubicina (2). Antraciclinas más representativas de este grupo de compuestos, su diferencia radica en el grupo OH en azul (Modificado de Sharif-Rad et al., 2019).

El mecanismo de acción de las antraciclinas varía de acuerdo con la naturaleza del tumor y la naturaleza estereoquímica de las drogas, pero de manera general se puede considerar que actúan de la siguiente manera:

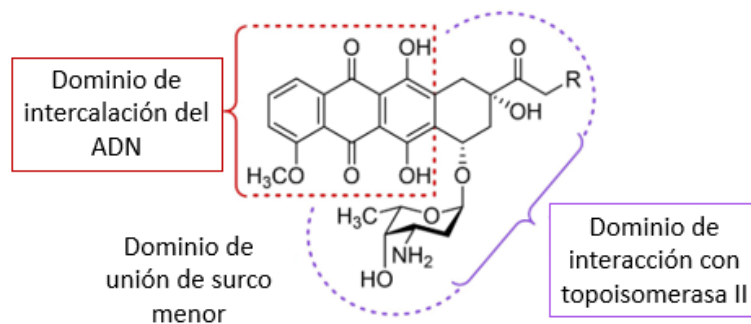


Figura 5. Dominios relevantes de las antraciclinas para la unión al ADN y topoisomerasa II, y consecuentemente a su actividad farmacológica (Modificado de Martins-Teixeira & Carvalho, 2020).

Inhibiendo a la topoisomerasa II, enzima esencial localizada en el núcleo y que juega un papel crítico en la replicación y la supervivencia de las células. También pueden intercalarse en el ADN debido a la inserción de la antraciclina entre los pares de bases de la doble hélice y la subsecuente inhibición de la síntesis de ADN y ARN (Figura 5). Y por último, generan especies altamente reactivas de radicales libres, lo que conlleva a la peroxidación de la membrana lipídica y el daño al ADN (Pashaei et al., 2020).

El uso de este tipo de fármacos en quimioterapia ha hecho evidente algunos efectos secundarios indeseables, un ejemplo de ellos es la cardiotoxicidad. Es una afección que puede llevar a complicaciones médicas. En especial, este efecto secundario se ha observado con el uso de la doxorubicina. La cardiotoxicidad de las antraciclinas,

al menos de cierto grado, está asociada con su capacidad de tomar parte en las reacciones redox que involucran al oxígeno molecular. Como subproducto de estas reacciones se forma un radical aniónico superóxido, lo que inicia una cascada de reacciones que involucran a las especies reactivas de oxígeno (ROS) y nitrógeno reactivo (RNS). Desafortunadamente el tejido cardíaco es particularmente sensible al estrés oxidativo, lo que ha demostrado que en el caso de la doxorubicina provoque daño al ADN, proteínas, lípidos y la estructura celular (Koss-Mikołajczyk et al., 2021).

La resistencia de las células tumorales a las antraciclinas ocurre principalmente por la alteración del mecanismo de transporte de membrana (Martins-Teixeira & Carvalho, 2020). Sin embargo, a pesar de estos efectos indeseables las antraciclinas siguen siendo una importante fuente de tratamiento para el cáncer.

2.2 Actinobacterias

Los compuestos con actividad anticancerígena antes mencionados corresponden en algunos casos a metabolitos secundarios que son producidos por microorganismos en la idiofase subsecuente al crecimiento. Este tipo de metabolitos no tienen una función en el crecimiento, sus estructuras químicas son variadas e inusuales y con frecuencia forman una mezcla de miembros cercanos en una familia química. En la naturaleza, su función es la de dar una ventaja para la supervivencia de la bacteria productora. Parece que los metabolitos secundarios le pueden servir a los organismos que los producen como armas competitivas usadas en contra de otros microorganismos, plantas, insectos o animales, como transportadores de metales, entre otras actividades (Sanchez & Demain, 2019).

Del total de metabolitos secundarios bioactivos que han sido descritos, aproximadamente el 45% son producidos por actinobacterias (Olano et al., 2009), lo que les confiere una gran importancia farmacéutica.

Las actinobacterias son bacterias filamentosas Gram-positivas pertenecientes al filo de las *Actinobacteria*, y se caracterizan por la formación de micelio aéreo, por la presencia de esporas con diferentes superficies, así como un alto contenido de Guanina y Citosina >70% en su ADN (Cragg et al., 2009). Además, su genoma se considera grande comparado con el de otras bacterias, ya que va de 1 a 12 Mb, con un promedio de 5 Mb. Sintetizan un gran número de productos naturales con propiedades anticancerígenas, muchos de los cuales también tienen actividad antimicrobiana (Demain & Vaishnav, 2011). Se encuentran en una diversidad de ambientes, ya que se conoce que son habitantes del suelo normal e hipersalino,

también varios grupos de actinobacterias son estables en el suelo de la rizosfera. Son muy importantes para muchas plantas, ya que puede proteger sus raíces inhibiendo el crecimiento de hongos patógenos.

En las dos décadas anteriores ha habido una disminución en el descubrimiento de nuevos compuestos derivados de actinobacterias del suelo, de donde anteriormente se han producido un gran número de metabolitos secundarios. En consecuencia, hábitats poco explorados o inexplorados como el agua de los ríos y mares, así como cavernas volcánicas, desiertos, insectos y como endófitos de plantas se han considerado como fuentes importantes para el descubrimiento de nuevos compuestos bioactivos (Selim et al., 2021). Particularmente, estamos interesados en los microorganismos endófitos debido a la biodiversidad tan grande con la que cuenta nuestro país, México.

2.3 Microorganismos endófitos

Las preparaciones de plantas han sido utilizadas para el descubrimiento de moléculas novedosas con actividades biológicas relevantes debido a los metabolitos secundarios que producen. Sin embargo, los microorganismos que viven en simbiosis con estas plantas también producen este tipo de compuestos. Por ejemplo, el taxol y la camptotecina son elementos anticancerígenos sintetizados tanto por plantas como por microorganismos endófitos (Miller et al., 2012).

Un endófito es un organismo que vive dentro de la planta, comúnmente son hongos y bacterias. Las plantas que albergan endófitos no presentan síntomas de enfermedad, al menos durante la fase endofítica de su ciclo de vida (Martinez-Klimova et al., 2017). Pueden ser encontrados en prácticamente todas las plantas del planeta, residiendo en los tejidos de su huésped, en una variedad de relaciones oportunistas o mutualistas (Guzmán-Trampe et al., 2015). Muchos endófitos producen metabolitos secundarios que confieren beneficios ecológicos al anfitrión, incluyendo la promoción del crecimiento, resistencia a predadores y patógenos, así como a sequías (Miller et al., 2012). Con el paso de los años, los microorganismos endófitos se han vuelto una fuente importante para el descubrimiento de moléculas con características biológicas interesantes (Ludwig-Müller, 2015).

2.4 Aislamiento y actividad biológica de microorganismos endófitos de *Amphipterygium adstringens*.

Debido a lo anterior, se realizó la búsqueda de microorganismos endófitos de la planta medicinal mexicana *Amphipterygium adstringens*, una planta de particular

interés tradicionalmente conocida como “Cuachalalate”. Pertenece a la familia Anacardiaceae, es un árbol endémico de México de alrededor de 10 m de altura y se localiza en regiones que van desde los 100 a los 300 msnm (Cuevas, 2005). En México es usado para el tratamiento de una variedad de afecciones a la salud humana como problemas circulatorios, cáncer de estómago, trastornos urinarios y gastrointestinales, enfermedades respiratorias y de los pulmones, entre otras (Rodríguez-Peña et al., 2018). El té hecho de la corteza del árbol es la forma de consumo más común de la planta. Contiene ácidos anacárdicos, triterpenos y esteroides como componentes principales (Quiñonez-Bastidas & Navarrete, 2021).

De esta planta se logró aislar varios microorganismos endófitos y el genoma de algunos de ellos fue secuenciado. De estos, el genoma de *Actinoplanes* sp. TFC3 tiene una secuencia que comprende 7,752,284 pb con un contenido de G+C del 68.9% (Centeno-Leija et al., 2016). Los *Actinoplanes* son un grupo de actinobacterias conocidas por producir importantes componentes farmacéuticos como antimicrobianos, anticancerígenos y antidiabéticos, por lo que se postula como un importante género para la búsqueda de metabolitos especializados novedosos. Otro microorganismo secuenciado fue *Streptomyces* sp. L06 que tiene un tamaño de genoma de 6,887,193 pb y un contenido de G+C de 73.05% (Vázquez-Martínez, 2017).

Por último, se aisló a *Embleya* sp. NF3 (inicialmente descrito como *Streptomyces scabrisporus*). La secuencia de este microorganismo contiene 8 contigios con un tamaño de 10,968,352 pb (Vázquez-Martínez, 2017).

De este último microorganismo se analizaron los extractos orgánicos producto de su fermentación y posteriormente se purificaron e identificaron antraciclinas pertenecientes a la familia de las estefimicinas.

2.5 La familia de las estefimicinas y sus derivados

Las estefimicinas son parte de las antraciclinas antibióticas descritas hasta el momento. La primera en ser descubierta fue la estefimicina de *Streptomyces steffisburgensis* como un antibiótico de tipo antraciclina que mostró tener actividad inhibitoria contra organismos Gram- positivos. Posteriormente, en 1974 se aisló al segundo miembro de la familia, la estefimicina B, obtenida de una nueva especie designada como *Streptomyces elgreteus*, que además produce estefimicina (Figura 6) (Brodasky & Reusser, 1974).

La estefimicina B inhibió significativamente el crecimiento de las células de leucemia de ratón L-1210. En este estudio se hizo evidente un problema de solubilidad del compuesto por lo que para intentar darle una solución, prepararon una solución de sal de sodio de estefimicina B; sin embargo, mostró menor actividad *in vitro* que el

compuesto original que fue diluido en buffer de fosfatos 0.1 M (Brodasky & Reusser, 1974).

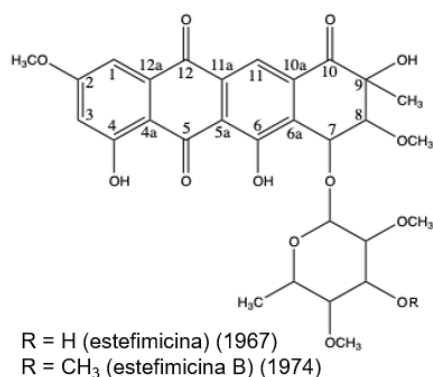


Figura 6. Diferencias en la estructura química de estefimicina y estefimicina B.

De los extractos orgánicos producto de la fermentación de *Embleya* sp. NF3, actinobacteria previamente aislada en el laboratorio, se caracterizaron químicamente 8 compuestos pertenecientes a la familia de las estefimicinas, estefimicinonas, estefimicina B, estefimicina C, la 8-demetoxi-10-deoxiestefimicinona y la 7-deoxiestefimicinona. La estefimicina B resultó ser el compuesto mayoritario (Rodríguez-Peña, 2020). Se probó la actividad citotóxica de los compuestos aislados en la línea celular MCF-7 de cáncer de mama con 6.25 $\mu\text{g/mL}$ de cada compuesto utilizando doxorrubicina como control positivo. De los resultados obtenidos, la estefimicina B mostró actividad citotóxica incluso mayor que la doxorrubicina ($p < 0.05$). Esto indica que este compuesto tiene el potencial para ser un agente quimioterapéutico. Sin embargo, al evaluar el estado fisiológico y muerte de las células mediante citometría de flujo, se observó que la doxorrubicina tuvo el mayor número de células en necrosis comparado con la estefimicina B. Nuevamente el problema de la solubilidad estuvo presente, ya que todos los compuestos aislados tuvieron una baja solubilidad y la formación de cristales fue observada 24 horas después de iniciado el experimento, lo que conduce a un tiempo menor de contacto entre las células y los compuestos. Por lo que se postuló que al aumentar la solubilidad del compuesto se incrementaría la muerte celular (Rodríguez-Peña et al., 2022). Respecto a los rendimientos de producción se obtuvieron 1.8 mg/L de estefimicina B en medio YMG (Rodríguez-Peña, 2020) y al suplementarlo con CaCO_3 este rendimiento aumenta a 70 mg/L (Trenado-Uribe, 2016); sin embargo, ambos valores están lejos de lo reportado para productos anticancerígenos comerciales. Por ejemplo, para la actinomicina D, se reportó una cepa de *Streptomyces sindenensis* super productora, con rendimientos de 850 mg/L (Praveen et al., 2008), así como otra de *Streptomyces parvus* que fue capaz de producir 1530 mg/L (Sousa et al., 2002). En el caso de la doxorrubicina, se han

logrado obtener rendimientos de hasta 1100 mg/L en fermentadores de 5 litros, lo que ha hecho que se considere a *S. peucetius* 33-24 como una cepa potencial para la producción de este fármaco directamente por fermentación (Wang et al., 2018). Industrialmente, a medida que aumenta la producción de un producto su costo disminuye, lo cual es particularmente importante para compuestos que tienen actividad anticancerígena lo que representa un área de oportunidad para este compuesto.

Además de la actividad anticancerígena, también se ha probado su actividad contra *Mycobacterium tuberculosis* H37Rv ATCC 27294 utilizando estefimicina B producida también por *Embleya* sp. NF3, pero en este caso, para solucionar el problema de solubilidad, se utilizó DMSO 100% (Trenado-Urbe et al., 2018).

Recientemente se describió a la estefimicina F, así como a la estefimicina y estefimicina D de *Streptomyces* sp. 0630c. Los tres compuestos mostraron actividad citotóxica en células humanas de cáncer de mama MCF-7, hígado HepG-2 y ovario A2780 así como actividad antibacteriana en el crecimiento de *Staphylococcus aureus* (Liu et al., 2021). Hasta el momento se ha descrito la actividad *in vitro* de estos compuestos, pero ninguno ha demostrado actividad *in vivo* y se cree que la falta de actividad puede ser en parte por la poca solubilidad en sistemas acuosos, lo que ha sido evidente en los resultados de los estudios ya realizados. Sin embargo, es importante resaltar que los resultados obtenidos hasta el momento hacen de la familia de las estefimicinas unos compuestos prometedores para ser utilizados en el tratamiento del cáncer (Liu et al., 2021; Olano et al., 2008).

2.6 Policétidos

Las antraciclinas, por mencionar un ejemplo de compuestos con actividad anticancerígena, pertenecen a una clase de metabolitos secundarios conocida como policétidos. Se estima que los policétidos tienen 5 veces más probabilidades de poseer propiedades como medicamentos en comparación con otras familias de productos de origen natural y cubren aproximadamente el 20% de los fármacos más vendidos con una estructura molecular pequeña (Dechert-Schmitt et al., 2014). Son producidos por hongos y cierto tipo de microorganismos, como las actinobacterias. Su estructura química consiste de grupos carbonilo y metileno alternados entre sí y son sintetizados por proteínas llamadas policétido sintasas (PKS) (Baerson & Rimando, 2007). Pueden tener una actividad biológica diversa y selectiva, por lo que obtener nuevos policétidos de la naturaleza para su evaluación como nuevos terapéuticos es un área de búsqueda activa (Ridley & Khosla, 2009), así como

también se han hecho esfuerzos por modificar la estructura química de policétidos ya descritos o aumentar su producción (Gao et al., 2021).

Los policétidos son sintetizados por las policétido sintasas (PKS). Las PKS son enzimas multifuncionales homólogas a las sintasas de ácidos grasos, aunque las PKS han reportado poseer una ruta biosintética más compleja y están presentes en grupos taxonómicos selectos (Singh et al., 2018). Son responsables de la biosíntesis de numerosos productos naturales que se utilizan como antibióticos (eritromicina), antiparasitarios (ivermectina) y para la quimioterapia de cáncer (epothilone) y doxorubicina (Nivina et al., 2019).

Se conocen tres tipos de PKS. El tipo I son enzimas multifuncionales organizadas en módulos. El tipo II son complejos multienzimáticos que llevan a cabo un único conjunto de actividades y por último, las tipo III son similares a la chalcona sintasa y actúan de forma iterativa como enzimas de condensación (Gullón et al., 2006). A continuación, se describe con mayor detalle a cada una de ellas.

2.6.1 Policétido sintasas tipo I

Para las PKS del tipo I, cada módulo de PKS consiste por lo menos de tres dominios proteicos absolutamente necesarios para el alargamiento de un policétido, una aciltransferasa (AT), la cetosintasa (KS) y la proteína acarreadora de acilo (ACP) (Figura 7).

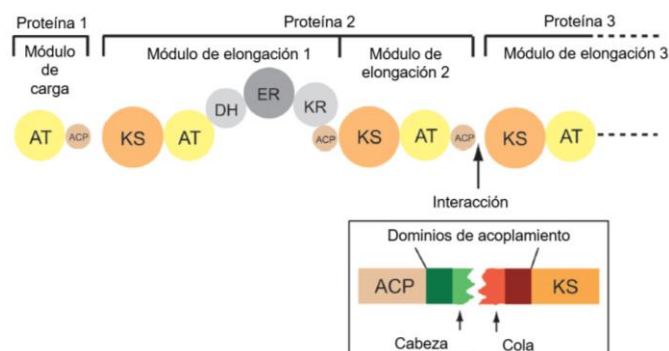


Figura 7. Esquema de los módulos en una PKS del tipo I. Los módulos de alargamiento mínimo de PKS comprenden los dominios KS, AT y ACP, mientras que los dominios (gris) DH (Deshidratasa), ER (Enoil reductasa), KR (Cetoreductasa) son opcionales. Las interacciones se basan en el acoplamiento de contactos de dominio entre “cabezas” C-terminal (verde) siguiendo el dominio ACP y “colas” N-terminal (rojo) frente al dominio KS (Modificado de Jenke-Kodama & Dittmann, 2009).

El proceso inicia cuando el dominio AT selecciona específicamente al sustrato que es un derivado de la acil-coenzima A, generalmente malonil-CoA o metilmalonil-CoA. Una vez seleccionado el sustrato, el dominio ACP es el responsable de anclar el metabolito intermedio del policétido en crecimiento y enviarlo a los diversos sitios activos dentro del módulo de elongación con ayuda de un brazo de fosfopanteteína.

A continuación, el dominio AT acila al dominio ACP y entonces el dominio KS puede catalizar una condensación descarboxilativa de tipo Claisen entre la unidad de extensión unida a ACP con la cadena del policétido en crecimiento que se encuentra en el módulo anterior. La catálisis de la KS da como resultado la generación de un producto / β - cetoacilo unido a ACP, que se convierte en el sustrato para el módulo de alargamiento siguiente. Además de estos tres dominios centrales, cada módulo de alargamiento puede contener dominios adicionales que pueden llevar al grupo ceto correspondiente a diferentes niveles de reducción. Es importante recalcar que cada dominio tiene participación solo una vez en la construcción del policétido, por lo que solo los dominios activos contribuyen en su formación.

Finalmente, la terminación de la biosíntesis del policétido es típicamente catalizada por un dominio tioesterasa (TE) ubicado en el extremo C-terminal del módulo de elongación final (Rath et al., 2010).

2.6.2 Policétido sintasas tipo II

Las PKSs tipo II consisten en una aciltransferasa (AT), dos unidades de cetosintasa una de ellas activa y otra inactiva a la que también se le conoce como factor de elongación de cadena (CLF) y de una proteína acarreadora de acilo (ACP). La biosíntesis de los policétidos inicia con la carga de la unidad de acilo, como en el caso de las PKS tipo I (Hu et al., 2020).

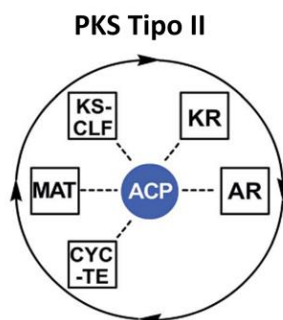


Figura 8. Esquema del funcionamiento de una PKS del tipo II. La actividad de las enzimas ocurre de manera iterativa. KS-CLF: cetosintasa- factor de elongación de cadena. KR: cetoreductasa. AR: aromatasa. CYC-TE: ciclasa-tioesterasa. MAT: metiltransferasa (Modificado de Chen et al., 2018).

Subsecuentemente el domino KS elonga la cadena de carbonos por condensaciones de Claisen descarboxilativas, pero a diferencia de los sistemas PKS I, las del tipo II únicamente utilizan malonil-CoA como unidad de extensión. Mientras la cadena de policétido está en crecimiento, la hendidura de la proteína KS-CLF mantiene a la cadena extendida permitiendo la actividad de las demás proteínas modificadoras incluyendo la KR, DH, ER, MAT por mencionar algunas

(Rohr & Hertweck, 2010). Después de la condensación de Claisen, la cadena poli β cetona es catalizada por una aromatasas y ciclasas para obtener un policétido aromático definido con grupos hidroxilos fenólicos. En algunos casos, KR cataliza la reducción del policétido naciente después de la cadena de ensamblaje, resultando en menos grupos hidroxilo, lo que contribuye a promover la ciclación de la cadena (Figura 8) (Hu et al., 2020).

Durante las últimas dos décadas, cerca de 40 clústeres biosintéticos que codifican para PKS del tipo II fueron localizados, clonados y en la mayoría de los casos, secuenciados. La identidad de los clústeres ha sido verificada casi exclusivamente mediante la inactivación y/o expresión heteróloga y los compuestos producidos pertenecen a una diversidad de familias, entre ellas las antraciclinas como la daunorrubicina, doxorubicina y nogalamicina, por mencionar algunos (Rohr & Hertweck, 2010).

2.6.3 Policétido sintasas tipo III

Las policétido sintasas del tipo III son homodímeros simples de cetosintasas las cuales catalizan una serie de reacciones que pueden dividirse en dos pasos. Primero, iterativamente condensan una unidad de sustrato inicial para generar un policétido lineal intermedio. Este intermedio es entonces ciclado intramolecularmente para producir un producto PKS ciclado. Hasta el momento, todas las PKS del tipo III conocidas utilizan malonil-CoA como unidad de extensión, y algunas de ellas pueden utilizar derivados como metilmalonil-, etilmalonil- y acetoacetil-CoA (Shimizu et al., 2016).

En las reacciones catalizadas por las PKS del tipo III, los intermedios de policétido son posteriormente ciclados por condensaciones de aldol, Claisen o lactonización. Los diferentes tipos difieren en la especificidad del sustrato inicial, la especificidad de los sustratos de extensión, el número de sustratos que son condensados y las reacciones de ciclación. Comparadas con las PKS I y II, las del tipo III tiene una estructura más simple y catalizan varias reacciones en un único centro catalítico.

Algunos ejemplos de compuestos catalizadas por estas PKS están involucrados en la biosíntesis de compuestos biológicamente importantes, por ejemplo, ArsB y ArsC de *Azotobacter vinelandii* que están involucrados en la biosíntesis de alquilresorcinoles y alquilpironas, respectivamente, que son importantes componentes de la cápsula del quiste. SrsA está involucrado en la formación de alquilquinona, que confiere resistencia a la penicilina a *S. griseus* (Katsuyama & Ohnishi, 2012).

2.7 Minería genómica

Los genes responsables de codificar las proteínas como las PKS están agrupados físicamente en el genoma de los microorganismos en grupos llamados clústeres de genes biosintéticos (BGC) (Belknap et al., 2020). Estos BGC pueden dividirse en varias clases dependiendo del tipo de genes biosintéticos y el tipo de metabolito producido. Los BGC también incluyen genes reguladores así como genes que le confieren resistencia al microorganismo frente al compuesto producido, normalmente identificados como transportadores (Meyer & Nodwell, 2021).

Teniendo esto en cuenta, el concepto de minería genómica fue desarrollado después de que se encontrara que el genoma de *Streptomyces coelicolor* contiene 22 clústeres de genes que codifican para la producción de metabolitos secundarios, pero solo cuatro de ellos eran expresados en productos bien caracterizados (Demain & Vaishnav, 2011).

Por lo tanto la minería genómica surge como una estrategia novedosa utilizada para el descubrimiento de nuevos compuestos basada en secuencias de grupos de genes y rutas biosintéticas (Li et al., 2018), así mismo es útil para identificar los BGCs responsables de sintetizar algún compuesto previamente caracterizado y así poder realizar mejoras o modificaciones en su biosíntesis.

El proceso consiste en la aplicación de plataformas para el análisis del ADN, que permita identificar a grupos de genes como por ejemplo, de PKS (Jenke-Kodama & Dittmann, 2009), entre otros. La identificación del clúster sugerirá también la novedad potencial del compuesto producido. En su caso, luego se procede al aislamiento y la expresión del clúster seleccionado, dando como resultado la síntesis de cantidades suficientes de producto para su posterior caracterización (Centeno-Leija et al., 2016).

2.7 Herramientas bioinformáticas

Para realizar la búsqueda de estos BGCs se cuenta con varias herramientas bioinformáticas, un ejemplo de estas son SEMPI 2.0 (Zierp et al., 2017) y antiSMASH (Weber et al., 2019).

AntiSMASH, (Antibiotics and Secondary Metabolite Analysis Shell) lanzada en el año 2011, identifica 52 diferentes tipos de BGCs (Blin et al., 2019) en una base de datos. Por ejemplo, policétidos, péptidos no ribosomales, terpenos, bacteriocinas, y sideróforos, entre otros. Realiza un alineamiento de las regiones identificadas con los genes homólogos más cercanos desde una base de datos que contiene todos

los otros grupos de genes conocidos, e integra o enlaza todos los métodos de análisis de genes específicos de metabolitos secundarios (Medema et al., 2011).

Una herramienta especializada en la predicción e identificación de productos naturales sintetizados por PKS del tipo I modulares es SEMPI 2.0. Esta plataforma utiliza la información proporcionada por antiSMASH y con el fin de limitar las posibles estructuras de los productos de las PKS e incluir reacciones de adaptación putativas, hace una comparación estructural con productos naturales previamente descritos (Zierep et al., 2017).

También existen herramientas como *PASS* (Prediction of Activity Spectra for Substances) (Lagunin et al., 2000), la cual permite predecir el tipo de actividad biológica que podría tener el compuesto de interés con base en su estructura y la comparación con sustancias ya reportadas. Predice cerca de 3,500 tipos de actividades biológicas. Las actividades biológicas son descritas cualitativamente (activa – inactiva) y son hechas con modelos de Naive Bayes (Filimonov et al., 2014).

Por último, CLC-Pred (Cell Line Cytotoxicity Predictor) es un servicio web para la predicción *in silico* de la citotoxicidad de compuestos químicos en líneas celulares no transformadas y cancerígenas basándose en la fórmula estructural. Brinda una predicción de la citotoxicidad de un compuesto químico para evaluar la relevancia de la inclusión de la sustancia en el escaneo experimental (Lagunin et al., 2018).

2.9 Expresión heteróloga de clústeres biosintéticos

La manipulación genética de nuevas cepas es a menudo difícil, y a veces es acompañado de un crecimiento lento y escasa esporulación. Esto puede llegar a dificultar la exploración y producción del compuesto buscado, es por esto que la expresión heteróloga de los clústeres biosintéticos de metabolitos secundarios es uno de los métodos clave en la minería genómica (Bekiesch et al., 2016).

Esta técnica ayuda a determinar el alcance y tamaño de este, ya que es importante conocer que genes dentro del clúster participan en la síntesis del compuesto. De esta manera, si se desea, poder incrementar la expresión de los genes biosintéticos o crear nuevos productos por biosíntesis combinatoria (Bekiesch et al., 2016).

Por ejemplo, para la estefimicina se realizó la expresión heteróloga de una región de aproximadamente 15 kb en *S. albus* que contenía 26 ORFs, donde 24 de estos probablemente estén involucrados en la biosíntesis de la estefimicina (Gullón et al., 2006). De esta manera al conocer los genes involucrados en su síntesis podría aumentarse su producción pensando en su potencial como medicamento antineoplásico.

Para lograr la expresión heteróloga, un método para la clonación de un BGC involucra la construcción de una librería de cósmidos o fósidos mediante una digestión parcial o aleatoria del ADN cromosomal. El tamaño típico de un BGC que codifica para un producto natural del tipo policétido es usualmente de más de 20 kb (algunas veces más de 100 kb) y un vector de este tipo puede aceptar tamaños de hasta 40 kb, lo que correspondería a un BGC pequeño o solo una parte de una BGC grande. La ventaja de este tipo de estrategia es que no tiene el problema de reconocimientos indeseados o alineamientos incorrectos en la PCR en el momento de la clonación, lo que es especialmente útil para los BGC de policétidos debido a su alto contenido de G+C y secuencias repetitivas.

Sin embargo, la clonación y la sobreexpresión eficiente de un BGC completo todavía sigue siendo un desafío debido a la ineficiencia de las células huésped actuales, incluidas las características genéticas y metabólicas en la manipulación de los BGC grandes para la expresión heteróloga (Nah et al., 2017).

3. JUSTIFICACIÓN

El cáncer ha sido destacado como uno de los mayores temas de preocupación para la salud pública mundial. A pesar de la numerosa cantidad de drogas sintéticas que han sido utilizadas no se han obtenido los resultados esperados debido a la resistencia de las células cancerígenas frente a los tratamientos por lo que es necesario desarrollar nuevos compuestos. Dado que los microorganismos son una fuente importante en la producción de sustancias bioactivas y gracias a la disminución en el costo de secuenciación de genomas, se han revelado numerosas rutas biosintéticas para metabolitos secundarios microbianos. Dentro de estos, los policétidos tienen cinco veces más probabilidades de poseer propiedades como medicamentos por su amplia diversidad estructural química y constituyen aproximadamente el 20% de los fármacos más vendidos con una estructura molecular pequeña, por lo que la minería genómica puede conducir al análisis de los clústeres de genes biosintéticos responsables de la síntesis de policétidos con actividad bioactiva útil y poder aumentar su producción con el fin de tener una aplicación médica en el tratamiento del cáncer.

4. HIPÓTESIS

El análisis del genoma de las actinobacterias endófitas *Embleya* sp. NF3, *Streptomyces* sp. L06 y *Actinoplanes* sp. TFC3 mediante la técnica de minería genómica permitirá la identificación de clústeres de genes biosintéticos de policétidos con actividad anticancerígena potencial.

5. OBJETIVO GENERAL

Identificar y caracterizar un clúster biosintético potencialmente asociado a la producción de un policétido con actividad anticancerígena en una actinobacteria endófito de *Amphipterygium adstringens*.

6. OBJETIVOS PARTICULARES

1. Explorar los genomas de *Embleya* sp. NF3, *Streptomyces* sp. L06 y *Actinoplanes* sp. TFC3 y seleccionar los clústeres que potencialmente codifiquen para la síntesis de un policétido con probable actividad anticancerígena.
2. Analizar *in silico* los BGC de policétidos identificados para determinar su posible funcionalidad al ser expresados de manera heteróloga.
3. Estandarizar las condiciones para aumentar el rendimiento en la producción y purificación del policétido con actividad anticancerígena seleccionado.
4. Evaluar la actividad antimicrobiana del policétido expresado y determinar algunas de sus propiedades.

7. METODOLOGÍA

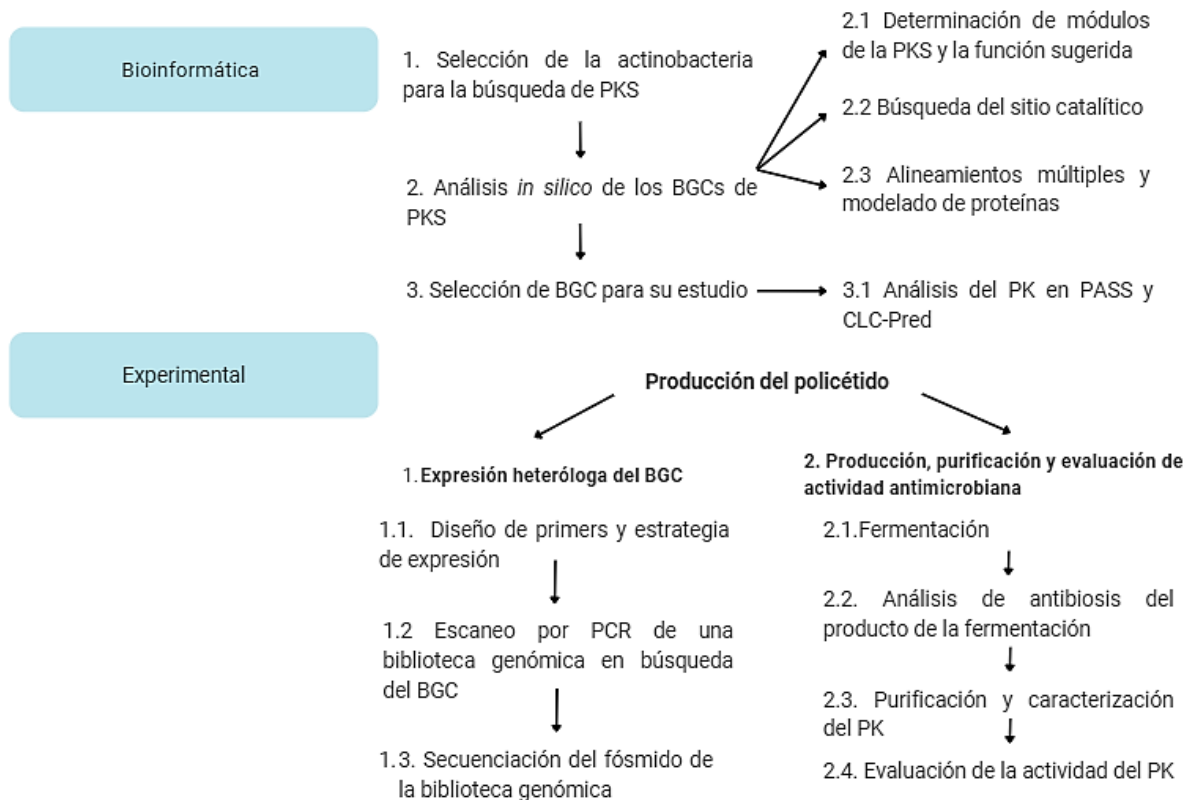


Figura 9. Estrategia experimental.

7.1 Bioinformática

7.1.1 Selección de una actinobacteria para la búsqueda de PKS

En el grupo de trabajo se tiene el genoma secuenciado de tres microorganismos endófitos de interés: *Actinoplanes* sp. TFC3, *Streptomyces* sp. L06 y *Embleya* sp. NF3. Se utilizó el archivo de la secuencia en formato FASTA y el programa antiSMASH v.6.0 (Blin et al., 2021). Para el rigor de detección se seleccionó relajado, lo que permite detectar clústeres bien definidos que contengan todas las partes requeridas, así como aquellos que tengan una o más partes faltantes, pero siguen perteneciendo a la misma clase de BGC. Respecto a las características adicionales se seleccionaron todas para poder obtener información adicional respecto a los BGCs identificados. La bacteria endófito para trabajar se eligió con base en el número de PKS novedosas identificadas, así como por el número de módulos que contienen.

7.1.2 Análisis *in silico* de los BGCs de PKS

7.1.2.1 Determinación de módulos de la PKS y la función sugerida

Para determinar el número de módulos que tiene cada PKS del tipo I se utilizaron los resultados del programa antiSMASH v. 6.0 (Blin et al., 2021). En el caso de PKS del tipo II únicamente se verificó la presencia de las enzimas mínimas necesarias para la biosíntesis del policétido. Para las PKS del tipo III se verificó la presencia de los genes necesarios.

Para conocer la función putativa de los genes dentro de la región identificada se hizo un análisis individual mediante protein BLAST utilizando bases de datos de proteínas no redundantes y como algoritmo de búsqueda Blastp. También se utilizó Uniprot e Interprot, así como Pfam.

7.1.2.2 Búsqueda del sitio catalítico

Se realizaron alineamientos múltiples de las secuencias de aminoácidos de cada proteína con secuencias de PKS ya reportadas previamente, para facilitar la visualización de los motivos conservados. Estos alineamientos se realizaron en Jalview (Waterhouse et al., 2009). La búsqueda del sitio catalítico se hizo considerando lo reportado en la literatura para cada tipo de proteína.

7.1.2.3 Alineamientos múltiples y modelado de proteínas

Así mismo, ya que existen algunas PKS del tipo I que han logrado ser cristalizadas y de las que se conoce el tipo de estructura que presentan, se hizo un modelado por homología con SWISS-MODEL (Waterhouse et al., 2018). El modelado se realizó únicamente de las proteínas del primer módulo tomando como templado lo recomendado por el programa en cada caso, de esta manera se verificó que cada dominio tenga el tipo de estructura previamente reportada.

7.1.3 Selección de BGC para su estudio

Para el clúster del tipo PKS I se utilizaron herramientas bioinformáticas especializadas en su análisis como es SeMPI 2.0 (Zierrep et al., 2020). Como archivo de entrada se cargó el genoma *Embleya* sp. NF3 en formato FASTA. En características adicionales se eligió que hubiera una predicción de genes, en el parámetro de la máxima distancia del clúster se dejó la cantidad predeterminada que es 20,000 pb y para las bases de datos se seleccionaron todas.

El resultado de la posible estructura química del policétido se introdujo en formato SMILES a la plataforma PASS (Filimonov et al., 2014) para predecir la actividad

biológica del policétido y en CLC-Pred (Lagunin et al., 2018) para conocer las líneas celulares donde podría tener acción antineoplásica.

Para la PKS del tipo II, se realizó el análisis de todos los genes que componen el clúster para delimitarlo a aquellos involucrados en la síntesis del compuesto.

7.2 Experimental

La preparación de las soluciones, medios de cultivo y buffers utilizados en la parte experimental se encuentran en el Anexo 1.

7.2.1 Expresión heteróloga del BGC

7.2.1.1 Diseño de primers y estrategia de expresión

Con el objetivo de realizar un escaneo de la biblioteca genómica de *Embleya* sp. NF3 y buscar el BGC 1.12 completo se realizó el diseño de tres pares de primers denominados steffInicio, que considera los genes *ctg_1768* al *ctg_1771*, steffMedio genes de *ctg_1781* al *ctg_1783* y steffFinal *ctg_1788* al *ctg_1790* (Figura 29, Anexo 2). Las características de los primers diseñados se evaluaron con la herramienta OligoAnalyzer Tool de Integrated DNA technologies (www.idtdna.com/calc/analyzer) y se realizaron PCR *in silico* en el programa SnapGene de Insightful Science (www.snapgene.com) para verificar que solo hubiera un sitio de unión en el genoma del microorganismo. Las secuencias finales de los primers se encuentran en el Anexo 2.

Para la expresión heteróloga se planeó utilizar a *Streptomyces lividans* ya que se ha reportado su uso como hospedero para la producción de PKS del tipo I y II (Nah et al., 2017). *S. lividans* TK24 es una cepa libre de plásmidos, es un pariente cercano de *S. coelicolor* A3(2). Produce pequeñas cantidades de antibióticos del tipo policétidos como actinorrodina y undecilprodigiosina durante la fase tardía de crecimiento y muestra resistencia a varios antibacterianos debido a sus sistemas de eflujo. Es un hospedero popular para la clonación y expresión heteróloga de varios antibióticos y otros metabolitos secundarios (Rückert et al., 2015).

En cuanto al vector de expresión a utilizar, se seleccionó el pIJ10257 (ver Anexo 3), el cual fue proporcionado por el Dr. Fernando Guzmán Chávez. El vector se recibió en seco en un papel filtro, se resuspendió el papel filtro con el vector en 50 μ L de buffer TE y se realizó la transformación en *Escherichia coli* DH5 α mediante electroporación como se especifica en el Anexo 4.

Se crecieron las colonias en medio LB con higromicina a una concentración de 80 µg/mL. Para la extracción del plásmido se siguió el protocolo especificado en el Anexo 5.

7.2.1.2 Escaneo por PCR de una biblioteca genómica en la búsqueda del BGC

Recientemente se construyó una biblioteca genómica de *Embleya* sp. NF3 que fue realizada en un fósido mediante el kit “CopyControl™ HTP Fosmid Library Production Kit” de Lucigen. Dicha biblioteca consta de 87 pooles correspondientes a dos ligaciones con dos vectores de fósido diferentes pCC1FOS™ y pCC2FOS™ (Anexo 3) con una capacidad de hasta 40 kpb contenidos en *E. coli*. Cada *pool* contiene aproximadamente 30 colonias positivas seleccionadas después de la ligación y están contenidos en placas de 24 pozos. Para la contención individual de las 1734 colonias positivas se utilizaron placas de 96 pozos, en todos los casos se encuentran en glicerol al 20% y -80°C. Se determinó la funcionalidad de dicha biblioteca genómica considerando la cobertura, los resultados sugieren que el genoma de la cepa NF3 está contenido aproximadamente 5.6 veces en las 1734 clonas que la conforman. Por tanto, era alta la probabilidad de encontrar una secuencia de interés en la biblioteca (Gómez, 2021).

Con el objetivo de realizar el escaneo de la biblioteca por PCR con los primers diseñados se extrajo el ADN plasmídico de los 87 *pooles* que la conforman. La extracción se hizo siguiendo el protocolo de mini preparación de plásmido de *E. coli* que se encuentra en el Anexo 6. Una vez obtenido el plásmido, se realizó la limpieza de ARN (Anexo 7). Para visualizar la presencia del fósido se hicieron geles con 0.7% (p/v) de agarosa y 0.01% de bromuro de etidio. La electroforesis se corrió a 90 V por 60 minutos en una cámara de electroforesis utilizando TAE 1x. Además, se hizo la cuantificación de ADN presente en las muestras de los 87 *pooles* por NanoDrop™ One/OneC espectrofotómetro de micro volúmenes de UV-Vi y se hicieron diluciones con agua miliQ estéril para obtener una concentración final de ADN plasmídico de 40 µg/µL.

La PCR se realizó en el termociclador Veriti (Applied Biosystems ®) con la enzima Invitrogen™ Platinum™ Taq ADN Polymerase (ThermoFisher). Los volúmenes de reactivos utilizados para la reacción de PCR se especifican en la Tabla 1.

Tabla 1. Volúmenes de reactivo para la PCR.

Reactivo	Volumen
ADN fósido pool [40 µg/µL]	0.8 µL <i>pool</i> individual, 2.4 µL total
Buffer de PCR 10X, -Mg	6 µL

MgCl ₂ [50 mM]	1.8 µL
dNTP mix [10 mM]	1.2 µL
DMSO 100%	3 µL
Primer forward [10 pM/ µL]	1.2 µL
Primer reverse [10 pM/ µL]	1.2 µL
Platinum™ Taq ADN Polymerase	0.6 µL
Agua miliQ	46.2 µL
Volumen total de la reacción	60 µL

Las condiciones utilizadas en la reacción de PCR para el escaneo de los *pooles* con los tres pares de primers se muestran en la Tabla 2, en todos los casos se utilizó como control positivo ADN genómico de *Embleya* sp. NF3 y se hicieron 30 ciclos.

Tabla 2. Condiciones de la PCR para la amplificación de las regiones del BGC 1.12

Nombre de la región	Etapa		Temperatura	Tiempo
StefInicio	1ra etapa	Desnaturalización inicial	95° C	5 min
	2da etapa	Desnaturalización	95° C	45 s
		Alineamiento	60° C	45 s
		Extensión	72° C	2:45 min
	3ra etapa	Extensión final	72° C	5 min
StefMedio	1ra etapa	Desnaturalización inicial	95° C	5 min
	2da etapa	Desnaturalización	95° C	45 s
		Alineamiento	68°C	45 s
		Extensión	72° C	3 min
	3ra etapa	Extensión final	72° C	5 min

SteffFinal	1ra etapa	Desnaturalización inicial	95° C	5 min
	2da etapa	Desnaturalización	95° C	45 s
		Alineamiento	64°C	45 s
		Extensión	72° C	2:45 min
	3ra etapa	Extensión final	72° C	5 min

Las 27 colonias individuales que conforman el *pool* positivo a la presencia de las tres regiones del BGC fueron crecidas en medio LB sólido con cloranfenicol a un volumen de 12.5 µL durante 16 horas, adicionalmente estas colonias reactivadas se guardaron en placas de 24 pozos con glicerol al 20%. Es estas colonias se les denominó copia *pool* 21 y se guardaron a -80°C.

Posteriormente, se realizó el protocolo de PCR de colonia (Anexo 8). La PCR de colonia toma ventaja de la alta sensibilidad de la reacción, ya que una pequeña cantidad del ADN templado es requerida para visualizar una banda en el gel de agarosa después de la amplificación por PCR. Las células se lisan en agua por un corto periodo de calentamiento y se agregan directamente a la reacción de PCR. Este paso inicial de calentamiento causa la liberación del ADN plasmídico de las células y puede servir como templado para la amplificación por PCR (Bergkessel & Guthrie, 2013).

Los productos de todas las reacciones de PCR se visualizaron en gel con 0.7% (p/v) de agarosa y 0.01% de bromuro de etidio. La electroforesis se corrió a 90 V por 60 minutos en una cámara de electroforesis utilizando TAE 1x, como marcador de peso molecular de utilizaron el marcador de ADN de 1 kpb de New England Biolabs (NEB).

7.2.1.3 Secuenciación del fósido de la biblioteca genómica

Una vez identificadas las colonias que tienen el clúster se procedió a la extracción y purificación del plásmido siguiendo el protocolo de purificación de ADN plasmídico o de cósmidos del kit de QIAGEN Plasmid Midi and Maxi disponible en www.qiagen.com. También se hizo una PCR de confirmación con las condiciones ya descritas.

Para determinar la cantidad de ADN presente en las muestras que se enviaron a secuenciar se utilizó un método fluorométrico con bromuro de etidio. El volumen de reacción fue de 50 μ L, utilizando una concentración de bromuro de etidio de 10 μ g/mL. Para realizar la curva de calibración se utilizó ADN genómico de *Embleya* sp. NF3 en concentraciones de 20, 40, 60 y 80 ng/mL. Las muestras se dejaron incubar por 5 minutos en una placa negra de 96 pozos utilizando una longitud de onda de 546 nm para la excitación y 590 nm para la emisión. Las lecturas se realizaron en el equipo Agilent BioTek Synergy H4 lector híbrido de microplacas.

Las muestras de las colonias A8 y A12 se enviaron para su secuenciación a la unidad universitaria de secuenciación masiva y bioinformática del Instituto de Biotecnología de la UNAM mediante secuenciación Sanger en una unidad de secuenciación masiva Genome Analyzer II de Illumina y el ensamble de las secuencias se hizo en el programa Spades (v.3.12.0).

7.2.2 Producción, purificación y evaluación de la actividad antimicrobiana

7.2.2.1 Fermentación

Para realizar la fermentación se realizó inicialmente una recolecta de esporas de *Embleya* sp. NF3 y se determinó la cantidad de esporas viables presentes. Se utilizó como inóculo para la fermentación una concentración de esporas de 9.55×10^8 UFC/mL y se utilizaron 100 μ L de inóculo por cada 100 mL de medio de cultivo.

El medio de fermentación fue YMG+CaCO₃, ya que en las actinobacterias se ha establecido una relación cercana entre el metabolismo secundario y la diferenciación celular. Esto debido a que los metabolitos exógenos pueden regular positivamente la producción de antibióticos y los resultados confirman la necesidad de calcio para la inducción de la esporulación y actividad antibacteriana (El-Naggar & Yousry, 2001).

Se realizaron dos fermentaciones de manera simultánea. Para la primera se inocularon 300 mL de medio con 100 μ L de esporas. La segunda fue en 150 mL de medio con la misma cantidad de inóculo, pero con 5% (v/v) de AmberLite™ XAD™ 7HP. La fermentación se llevó a cabo por 21 días a una temperatura de 29°C y agitación de 180 rpm.

7.2.2.2 Ensayos de antibiosis y bioautografía

Los ensayos de antibiosis se hicieron por el método Kirby-Bauer colocando discos con 20 μ L de muestra en medio de cultivo agar Mueller-Hinton (Polvo de extracto de carne, 2.0 g, digerido ácido de caseína, 17.5 g, almidón, 1.5 g, y agar 17.0 g).

El microorganismo para determinar la sensibilidad fue *M. luteus*. Como control positivo se utilizó ampicilina a una concentración de 50 µg/mL y como control negativo medio YMG estéril. Las placas se incubaron a 37°C por 48 horas en todos los casos.

Para probar la actividad de los extractos obtenidos contra bacterias Gram-positivas por bioautografía, se inoculó a *M. luteus* en medio TSB durante 24 h y se centrifugó a 8,000 rpm a 4°C, se desechó el sobrenadante y se preservaron con glicerol al 20%.

Para tener un estándar de la concentración de microorganismos contenidos en cada una de las pruebas se preparó la solución 0.5 McFarland que equivale a un patrón de turbidez de 1.5×10^8 células de *E. coli* por ml³ si se lee a 0.08-0.10 de absorbancia a 625 nm. Se preparó una suspensión con solución salina estéril al 0.85% y *M. luteus* equivalente a la turbidez de McFarland. La placa de CCF de las muestras se colocó en una charola de aluminio estéril. Se agregaron 4 mL de esta solución a 350 mL de medio sólido TSB y se colocó en la charola. Se incubó a 37°C por 48 h y se observaron los halos de inhibición.

7.2.2.3 Purificación y caracterización del policétido

Al finalizar la primera fermentación, se colocó 10% (v/v) de AmberLite™ XAD™ 7HP. Para la segunda, se filtró la AmberLite™ XAD™ 7HP del medio una vez terminada la fermentación.

Como paso siguiente, la amberlita se colocó en una columna cromatográfica abierta y se realizó el proceso de elución de los compuestos empleando mezclas de agua y metanol en gradiente (0 a 1) como se especifica en la Tabla 3.

Tabla 3. Gradiente de solventes agua-metanol.

Fracciones	Solventes	Porcentaje de solventes (%)
1	agua-metanol	70:30
2	agua-metanol	50:50
3	agua-metanol	30:70
4	metanol	100
5	acetona	100

Para todas las eluciones se usó un volumen de dos camas para el solvente y las fracciones se secaron en desecador. Posteriormente se agregaron 5 mL de hexano para extracción de ácidos grasos y se desecharon por decantación.

Las fracciones activas obtenidas del extracto crudo se analizaron utilizando cromatografía en capa fina (CCF) disolviéndolas en metanol y como fase móvil se utilizó cloruro de metileno 30 partes, acetona 10 partes, hexano (95%) 3 partes y metanol 2 partes. Se dejó correr la CCF por una hora con 20 minutos y para visualizar la placa fue en luz UV con una lámpara UVGL-55 Hand-held UV lamp.

Posteriormente hizo una CCF preparativa, para recuperar el compuesto. La banda del compuesto se delimitó al observarla con luz UV y la recuperación se hizo mediante centrifugación a 12,000 rpm por 15 minutos utilizando metanol como disolvente, las muestras se secaron en desecador.

Para la caracterización del compuesto la muestra se analizó por cromatografía líquida- espectrometría de masas. Para esto se utilizó un cromatógrafo de líquidos de ultra alta eficiencia (UPLC) acoplado mediante ionización electrospray a un espectrómetro de masas Xevo G2 XS TOF (Ambos marca Waters). El estándar fue de estefimicina B de Enzo Life Sciences ALX-380-220-M001 (Ann Arbor, MI, USA) y la muestra, fueron disueltos en 50:50 de Acetonitrilo-Agua, llevándolas a una concentración aproximada de 1 µg/ml. Antes del análisis la muestra se filtró utilizando un filtro de PVDF 0.45 µm debido a la turbidez que presentaba. Las condiciones del UPLC y del equipo de masas fueron las siguientes.

UPLC Condiciones

LC-Sistema:	Acquity UPLC
Columna:	ACQUITY UPLC BEH C18 1.7µm (2.1mm x 50mm)
Temperatura de la columna:	30 °C
Temperatura de la muestra:	10 °C
Volumen de inyección:	5 µL
Velocidad de flujo:	300 µl/min
Fase móvil A:	Bicarbonato de Amonio 5 mM
Fase móvil B:	Acetonitrilo
Elución Isocrática:	5 min 50% A:50%B

MS Condiciones

MS sistema:	Xevo G2 XS TOF
Modo de adquisición:	Full scan MS
Modo de ionización:	ESI-
Adquisición:	5 min
Source temp. (°C):	100

Capillary (kV):	1.5
Sampling cone:	20
Source offset:	80
Desolvation temp.(°C):	450
Cone gas flow (L/Hr):	50.0
Desolvation gas flow (L/Hr):	800.0.

7.2.2.4 Evaluación de la actividad del policétido

Gracias a los antecedentes reportados, se conoce la poca solubilidad de la estefimicina B en medios acuosos. Para solventar este problema se hizo un complejo en proporción 1:3 p/p estefimicina: β -ciclodextrina. La preparación de la ciclodextrina β se hizo con agua estéril al 0.9% (p/v) y se calentó para facilitar la solubilidad. Para los ensayos de antibiosis se utilizaron 20 μ L de estefimicina B: ciclodextrina y como control el mismo volumen de estefimicina B: DMSO. El DMSO es un líquido orgánico altamente polar que se usa ampliamente como solvente. El control positivo fue ampicilina a una concentración de 50 μ g/mL. Se hicieron con *M. luteus* en medio MH y se dejaron incubando durante 24 horas a 37°C.

Los ensayos de concentración mínima inhibitoria (MIC) permiten evaluar la actividad antimicrobiana de un compuesto, por lo que se decidió hacer estos ensayos en medio MH líquido y como microorganismos para medir la susceptibilidad *M. luteus* y *Bacillus subtilis*. El procedimiento se hizo siguiendo el protocolo de métodos de dilución en agar y caldo para determinar la MIC de sustancia antimicrobianas de Nature Protocols (Wiegand et al., 2008). Se colocó medio MH líquido como control de esterilidad. Se evaluaron 10 concentraciones diferentes del complejo ciclodextrina β - estefimicina B y DMSO- estefimicina B, para *B. subtilis* del rango de 50 - 0.1 μ g/mL y para *M. luteus* 10 - 0.02 μ g/mL. Los ensayos se hicieron por triplicado, se colocó un control de crecimiento y como control de sensibilidad se utilizó ampicilina a diferentes concentraciones. Las placas se incubaron por 24 horas a 37° C y se leyeron a una longitud de onda de 595 nm en el equipo Thermo Scientific™ Fotómetro para microplacas Multiskan™ FC.

8. RESULTADOS Y DISCUSIÓN

8.1 Minería genómica

Se realizó el análisis del genoma de las tres actinobacterias endófitas en antiSMASH versión 6.0 para la identificación de metabolitos secundarios y se obtuvieron los resultados mostrados en la Tabla 4. Estos resultados son para BGCs no caracterizados previamente.

Tabla 4. Metabolitos secundarios identificados por antiSMASH 6.0.

Tipo de metabolito secundario	<i>Embleya</i> sp. NF3	<i>Actinoplanes</i> sp. TFC3	<i>Streptomyces</i> sp. L06
NAPAA (non-alpha poly-amino acids like e-Polylysin)	1	1	-
Lantipéptidos	6	-	1
NRPS	5	2	-
NRPS-otros	11	4	3
Lassopéptidos	1	-	1
PKS (Tipo I, II y III)	6	4	2
Terpenos	4	2	3
Sideróforos	3	1	2
RiPP	-	1	3
Otros	10	1	1

En *Actinoplanes* sp. TFC3, se identificaron 17 clústeres, en su mayoría son clústeres combinados del tipo NRPS, solo de NRPS y terpenos. Las PKS presentes cuentan con menos de 4 módulos completos (Tabla 4). Resalta la presencia de una trans aciltransferasa novedosa, que sugiere una vía biosintética diferente y compleja (Centeno-Leija et al., 2016). El tamaño de toda la región es de más de 120 kpb.

Streptomyces sp. L06 tiene 18 clústeres biosintéticos, donde 16 posiblemente codifican para compuestos novedosos. Estos son del tipo NRPS, RiPP y terpenos, también se detecta la presencia de sideróforos, lantipéptidos, y péptidos de lazo (Tabla 4). Se encontraron PKS del tipo I y III. En el caso de la PKS I detectada tiene

únicamente un módulo incompleto, por lo que no se consideró seguir analizándola. La PKS del tipo III presenta únicamente una cetosintasa, y varios genes a los que no se les pudo encontrar una identidad mediante alineamientos.

De los tres genomas analizados, el de *Embleya* sp. NF3 es el que presentó un mayor número de clústeres, con 49 identificados. Se encontró que 47 codificarían para compuestos inéditos y en su mayoría corresponden al tipo NRPS o una combinación con otro tipo de clúster (Tabla 4).

Observamos la presencia de PKS del Tipo I, II y III, así como su combinación. Lo reportado anteriormente respecto a compuestos con actividad anticancerígena se basa principalmente en los dos primeros tipos de PKS, por lo que se decidió continuar con un análisis más riguroso del genoma de este microorganismo.

Para poder seleccionar un clúster biosintético se procedió a analizar aquellos del tipo PKS buscando que contaran con los dominios mínimos requeridos para la síntesis de este tipo de compuestos, para después analizar los genes presentes en toda la región identificada. Otro aspecto que nos puede guiar para la selección es considerar el largo de cadena de compuestos reportados anteriormente. Se han observado diversos tamaños; sin embargo, podemos detectar una tendencia a favor de los compuestos de más de 5 carbonos de longitud (Belknap et al., 2020; Yan et al., 2016), por lo que se optó por buscar clústeres que tuvieran el mayor número de módulos completos.

A continuación, se enlistan las regiones identificadas por antiSMASH 6.0 que se consideran como candidatos para la posible producción de un metabolito con propiedades anticancerígenas (Tabla 5).

Tabla 5. Regiones identificadas en *Embleya* sp. NF3.

Región antiSMASH	Tipo de compuesto	Total, nt
1.4	T1PKS	70,745
1.12	T2PKS	72,489

Se decidió hacer un análisis bioinformático de cada región para determinar si cuenta con los elementos reportados para cada tipo de PKS, así como definir la factibilidad de realizar la expresión heteróloga del clúster.

8.2 Análisis bioinformático de la región 1.4.

En antiSMASH, una región es aquella que contiene uno o más clústeres candidatos (Blin et al., 2019). Se analizó cada gen con Pfam y BLAST para corroborar su función putativa (Figura 10).

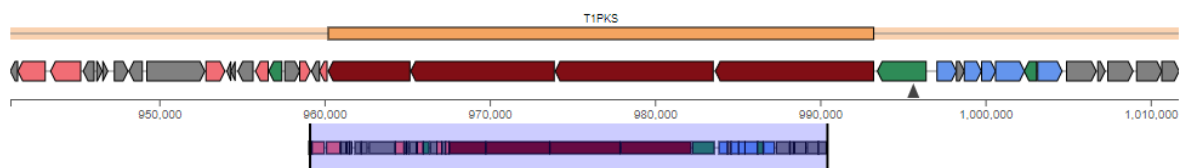


Figura 10. Región 1.4 correspondiente a PKS del tipo I en *Embleya* sp. NF3. Genes core (guinda), genes relacionados a reguladores (verde), genes adicionales en la biosíntesis (rosa), genes relacionados al transporte (azul).

Se analizaron primero los cuatro genes principales involucrados en la biosíntesis, identificados como “core”. En la Tabla 6 se coloca la nomenclatura utilizada en la identificación de dominios en cada módulo. El tamaño de los cuatro genes (Figura 10) es de 32,760 pb.

Tabla 6. Nomenclatura para identificar los dominios en PKS.

Icono	Descripción	Icono	Descripción
ACP	Dominio de la proteína portadora de acilo	KS	Dominio cetosintasa
AT	Dominio acetil transferasa	TE	Dominio tioesterasa
DH	Dominio deshidratasa	KR	Dominio cetoreductasa
ER	Dominio de enoil reductasa		

Observamos la presencia de los módulos mínimos necesarios para una PKS del tipo I (KS, AT, ACP y TE). Respecto a los módulos adicionales, se sabe que la presencia de un dominio cetoreductasa (KR) da lugar a la funcionalidad β -hidroxilo y la presencia tanto de un dominio KR como de un dominio deshidratasa (DH) genera un alqueno, mientras que la combinación de KR, DH y enoilreductasa (ER) da como resultado alcanos (Rath et al., 2010) (Figura 11).

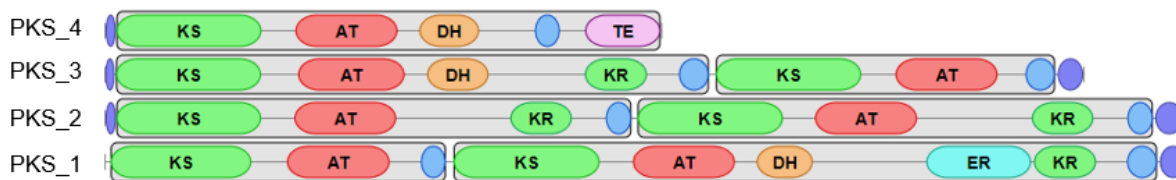


Figura 11. Dominios presentes en los módulos de la PKS I de la región 1.4.

Esta combinación de módulos es lo que permite predecir la posible estructura del compuesto, lo que se puede observar en el tipo de estructuras predichas por SeMPI 2.0 (Figura 12).

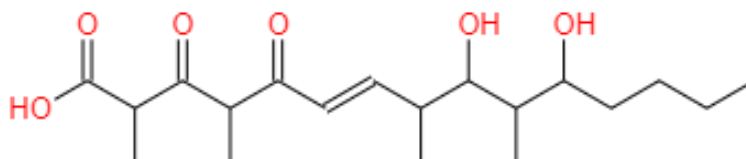


Figura 12. Estructuras previstas por SeMPI 2.0 para la región 1.4.

A continuación, se presentan los resultados para los análisis de cada uno de los dominios presentes en esta PKS I.

8.2.1 Dominio KS

El dominio cetosintasa es de aproximadamente 430 residuos de aminoácidos y cataliza la formación de enlaces carbono-carbono, estas enzimas hacen posible el crecimiento de las cadenas de acilos. Representa al dominio más conservado de todos. La maquinaria catalítica está comprendida por una cisteína encontrada en el motivo (TACSSS), y dos histidinas en los motivos (EAHG TG y KSNIGHT) (Keatinge-Clay, 2012). Estos motivos fueron encontrados en todas las KS de la PKS 1.4 como se muestra en la Figura 13.

Los dominios KS del tipo PKS I, tienen la triada catalítica cisteína, histidina, histidina, en donde la cisteína es el nucleófilo involucrado en las reacciones de trans-esterificación. En los alineamientos múltiples realizados se observan los aminoácidos catalíticos, indicando que todos los dominios son funcionales. Para la KS_1 se observa el cambio de cisteína por glutamina pero se ha reportado que este cambio sigue permitiendo que KS sea una buena descarboxilasa (Keatinge-Clay, 2012).

	170	180	313	348	358
KS_1	SLTVDTAQSSSLVVIQYV	ELHGTG	TAVG	SVKTNIGH	LÉGASGIVGL
KS_2	ALTIDTACSSSLVVDVAV	EAHGTG	TLVG	SVKSNLGH	TQAAAGVTGV
KS_3	AVTVDTACSSSLVVDVAV	EAHGTG	TWLG	SVKSNIGH	AQAAAGVAGV
KS_4	AVTVDTACSSSLVVDVV	EAHGTG	TWLG	SVKSNIGH	AQAAAGVAGV
KS_5	AVTVDTACSSSLVIDAV	EAHGTG	TWIG	SAKSNIGH	TQAAAGIVGV
KS_6	AVTIDTACSSSLVVDVAV	EAHGTG	TWLG	SVKSNIGH	AQAAAGVAGV
KS_7	AVTVDTACSSSLVVDVV	EAHGTG	TWLG	SVKSNIGH	AQAAAGVAGV
Chlorothricin	AVTIDTACSSSLVVDVAV	EAHGTG	TWLG	SVKSNIGH	AQAAAGVAGV
FD891	AVTVDTACSSSLVVDVV	EAHGTG	TWLG	SVKSNIGH	TQAAAGVGGV
EryAll	SISVDTACSSSLVVDVAV	EAHGTG	TWLG	SVKSNIGH	TQAAAGVTGL

Figura 13. Alineamientos múltiples del dominio KS de la PKS 1.4. PKS de chlorothricin de *Streptomyces antibioticus*, FD891 de *Streptomyces graminofaciens* y eritromicina de *Saccharopolyspora erythraea*.

Respecto a la estructura, el dominio KS es un homo-dímero, lo cual se comprobó con el modelado por homología (Figura 14). La cisteína está enterrada en el bolsillo de sustrato y solo se puede alcanzar con el brazo de fosfopanteteína adjunto al dominio ACP. Los dominios KS en las PKS han sido asociados con el control de la longitud de la cadena (Xu et al., 2013).



Figura 14. Modelado por homología del dominio KS por SWISS-MODEL. Templado de EryAll de la policétido sintasa de eritromicina con una identidad de 65.80% y cobertura del 99%.

8.2.2 Dominio AT

Las aciltransferasas utilizan una pareja conservada serina-histidina y un mecanismo de ping-pong para transferir el grupo acilo de CoA a los dominios ACP. La nucleoficidad del grupo hidroxilo de la serina catalítica es fortalecida mediante un enlace de hidrógeno con la histidina.

Alineamientos de todas las AT de la PKS 1.14 con otras reportadas previamente en actinobacterias permitieron identificar esta dupla lo que indicaría que todos los módulos son funcionales (Figura 15) (Xu et al., 2013).

	100	119	196	206
AT_1	RPDÁVIGHSQGEIÁADAARLVAR	SLLIPVDYASH	SAHVEPLF	
AT_2	RADHLIGHSIGELTAADAAALVVA	GRLLRVSHAF	SPHMDAVL	
AT_3	VPDALVGHSQGEIAAADGATVVAL	SRAIPVDYASH	SAQVEEIF	
AT_4	VPDALVGHSQGEIAAADGARVVAL	SRAIPVDYASH	SAQVEEIF	
AT_5	RPDFVTGHSIGELTAADAATLVAA	ARLLRVSHAF	SPLMDGML	
AT_6	TPDAVVGHSQGEIAAADGARVVAL	SRAIPVDYASH	SAQVEEIF	
AT_7	TPDAVVGHSQGEIAAADGARVVAL	SRAIPVDYASH	SAQVEEIF	
Neoaureothin	TPDLLLGHSIGELTAADACALVAA	GRLLRVSHAF	SPQMDGML	
Lasalocid	TADFLVGHSVGEIAAADACALVAA	GRLLHVSHAF	SPRMDAML	
Amphotericin	TPDYVGGHSIGEIAAADACALVAA	ARLLRVSHAF	SPLMDPML	
Fr008	TPDHLTGHSVGEIAAADACTLVTA	GRLLRVSHAF	SPLMDPML	
Chalcomycin	TPQVVVGHSIGELTAADAAALVAA	GRLLQVSHAF	SPRMEPML	
Herbimycin	RPAYVLGHSIGEVAAADACTLVRA	GRLLVVS HAF	SPHMDGML	
Nystatin	RPDHLAGHSVGEIVAADAATLVAA	GRLLRVSHAF	SPLMEPML	
Salinomycin	APAAVIGHSQGEIAADAARKVSAL	AKA INVDYASH	GPQVEHIF	

Figura 15. Alineamientos múltiples del dominio AT de la PKS 1.4. Los aminoácidos catalíticos son serina (verde) e histidina (azul). Arginina (rojo) probablemente imita al extremo carboxilo de un grupo malonil unido (Smith & Tsai, 2007).

Las AT catalizan la transferencia de una unidad extensora, como el malonil o metilmalonil-CoA a ACP. Los motivos GHSQG y YASH en AT significan que es específico para (2S)-metilmalonil-CoA (Keatinge-Clay, 2012), como se observa en la Figura 15 no todos los dominios AT tienen la misma especificidad y estas modificaciones son las que permiten que programas como antiSMASH y SEMPI 2.0 puedan predecir la estructura del policétido resultante (Figura 12).

De su estructura se sabe que es un monómero y tiene dos subdominios, uno de aproximadamente 250 residuos que es el catalítico y posee un pliegue α/β hidrolasa, así como un subdominio ferredoxina, al realizar el modelado por homología podemos corroborar esta estructura (Figura 16).

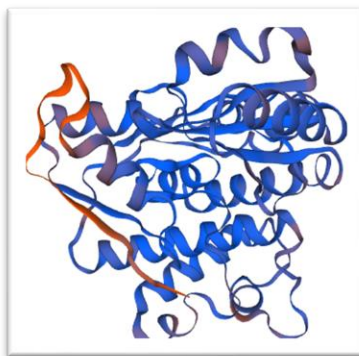


Figura 16. Modelado por homología del dominio KS por SWISS-MODEL. Templado PKS de salinomicina con una identidad de 59.46% y cobertura del 99%.

La serina catalítica se encuentra en el motivo conservado GHSXG, que se encuentra en un giro brusco entre una hoja β y una hélice α , una característica estructural comúnmente conocida como “codo nucleofílico” (Smith & Tsai, 2007).

8.2.3 Dominio ACP

El sitio de acoplamiento ACP es flexible para que permita el movimiento. El brazo de fosfopanteteína es atacado por una serina localizada en el N-terminal de ACP en un motivo conservado DSL, el aminoácido que debe ser invariable es la serina (Keatinge-Clay et al., 2003). Al realizar los alineamientos múltiples se observa este motivo en todas las ACP de la PKS (Figura 17).

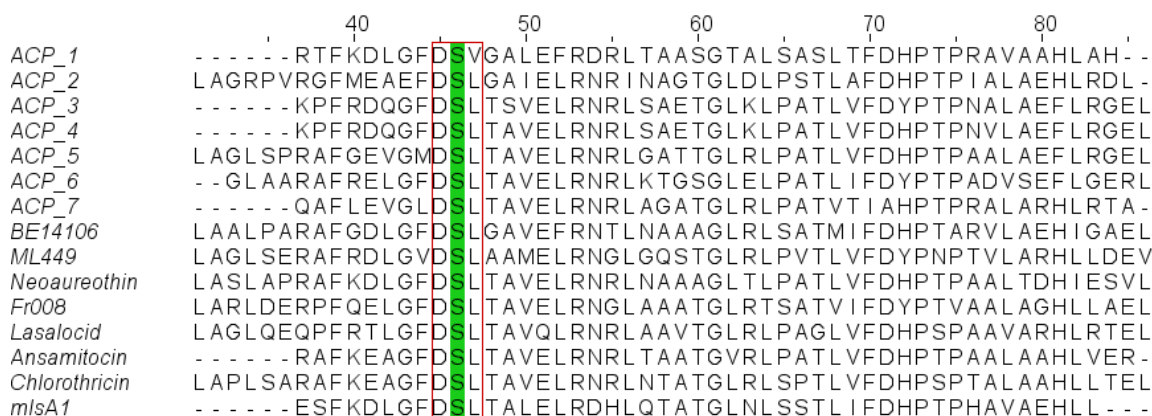


Figura 17. Alineamientos múltiples del dominio ACP de la PKS 1.4. Los aminoácidos catalíticos son serina (verde) dentro del motivo DSL.

Los dominios ACP son de aproximadamente 80 residuos y tienen cuatro hélices que están situadas en los extremos C-terminales de los módulos PKS, transportan policétidos entre los sitios activos de la sintasa (Figura 18).

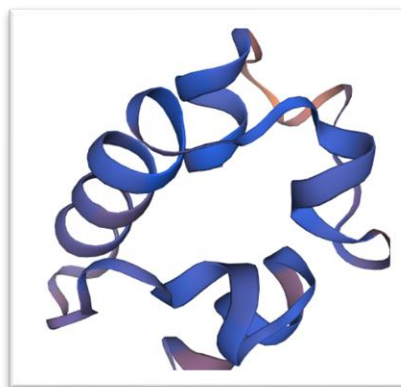


Figura 18. Modelado por homología del dominio ACP por SWISS-MODEL. Templado PKS de MLSA1 con una identidad de 56.92% y cobertura del 92%.

Una ACP lanza covalentemente unidades de extensión y policétidos intermedios a través de un enlace tioéster a un brazo de fosfopanteteína. Este grupo se añade

postraduccionalmente a la serina del motivo DSL mediante una fosfopanteteinil transferasa (PPTasa) utilizando CoA como sustrato (Xu et al., 2013).

8.2.4 Dominio DH

Los dominios deshidratasa catalizan reacciones de deshidratación reversibles en los intermedios β -hidroxiacilos. La diada catalítica del sitio activo en cada monómero está localizada en un túnel formado entre el C y N terminal de los pliegues de “hot-dog” que comprenden una histidina del primer pliegue y un aspartato del segundo (Figura 19).

	38	48	58	68
DH_1	DG - - - LAD	HAVQDTVLLP	GTAFVDLALYAA	ERTGAGGVGELVL
DH_2	GGAADGLAGH	VVFGSVLVLP	GTAFVEAAIRAGD	EVGLDTVEELVI
DH_3	DG - - - LAD	HAVRDTVLLP	PASAFLELAVRAGD	EVGLGVVDELVI
Herboxidiene	AG - - - LTD	HAVDGTVLLP	GTAFVELAVRAGD	EAGCDRVDELTL
Tautomycin	QT - - - PAE	HPVGGSAVLP	PATAYVELAVRAGD	SAGCDRLDELVV
Oleandomycin	DS - - - LAD	HEVLGSVLLP	GTAFVELAVQAAD	RAGYDVLDELTL
Apoptolidin	DS - - - LAG	HAVLGSVLFP	GTGFVELAVRAAD	QVGCAVLDELVL
Rifamycin	QS - - - LAD	HAVRDVVIVP	GTGLVELAVRAGD	EAGCPVLDDELVI

Figura 19. Alineamientos múltiples del dominio DH de la PKS 1.4. Los aminoácidos catalíticos son histidina (verde) dentro del motivo HAVLGSVLLP y aspartato (rosa).

Es un dominio de aproximadamente 280 residuos que cataliza la deshidratación de un intermedio policétido para hacer un doble enlace entre los carbonos α y β . Los monómeros del dímero DH hacen contacto a través del doble eje de la sintasa (Figura 20) (Keatinge-Clay, 2012).



Figura 20. Modelado por homología del dominio DH por SWISS-MODEL. Templado de rifamicina con una identidad de 44.44 % y cobertura del 100%.

8.2.5 Dominio ER

Aún no se cuenta con dominios PKS enoil reductasa activos descritos en su

totalidad y por lo tanto no hay una estructura cristalina de rayos X disponible, por lo que el rol de los residuos activos en la catálisis continua sin ser claro (Xu et al., 2013).

Mediante mutagénesis puntual se ha podido elucidar que los residuos catalíticos son el Aspartato y la Tirosina pero aún faltan estudios para tener una mayor claridad en estos cambios (Kwan & Leadlay, 2010). A pesar de esta limitante se hicieron los alineamientos múltiples y se encontraron los mismos residuos de aminoácidos con las secuencias reportadas para otras actinobacterias (Figura 21).

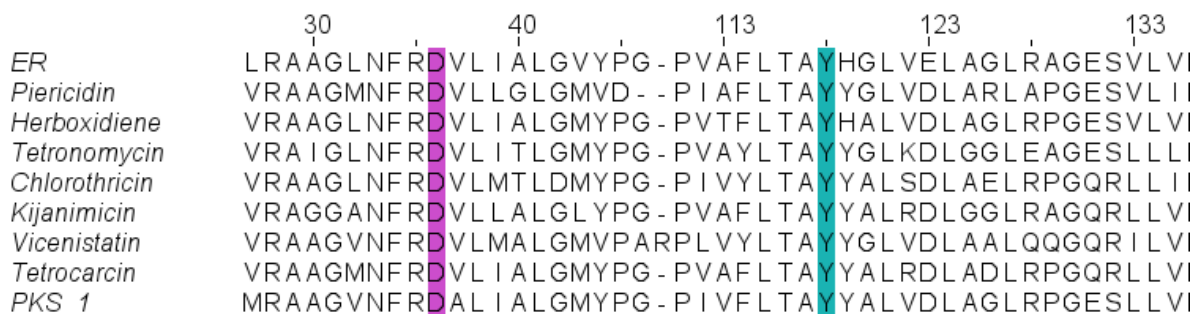


Figura 21. Alineamientos múltiples del dominio ER de la PKS 1.4. Los aminoácidos catalíticos son aspartato (rosa) y tirosina (verde).

8.2.6 Dominio KR

El dominio monomérico cetoreductasa de aproximadamente 460 residuos es uno de los más versátiles en las PKS del tipo I.

La tríada catalítica conservada en el subdominio catalítico serina-tirosina-lisina (SYK), es la responsable de la activación del grupo β -carbonilo objetivo en el sustrato y de estabilizar el oxianión resultante después de la adición del hidruro. El motivo "LDD" está localizado en el bucle que precede al sitio activo (Xu et al., 2013) (Figura 22).

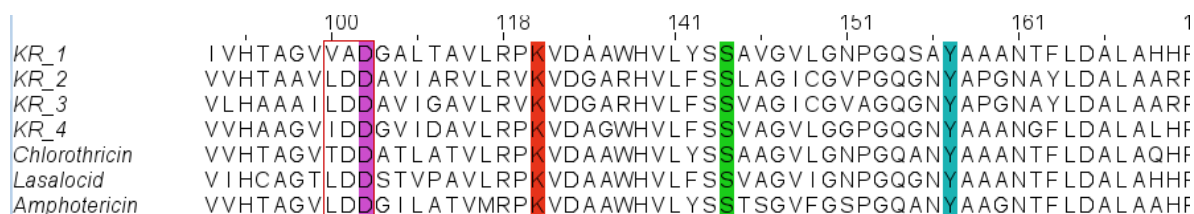


Figura 22. Alineamientos múltiples del dominio KR de la PKS 1.4 con PKS de diferentes actinobacterias. Los aminoácidos catalíticos son lisina (rojo), serina (verde) y tirosina (azul). El motivo LDD muestra al aspartato conservado (rosa).

Este tipo de oxidorreductasas son dependientes de NADPH, y los dominios conservados crean un surco donde se une el cofactor NADPH (Figura 23). En el modelado por homología fue posible encontrar este tipo de estructura.

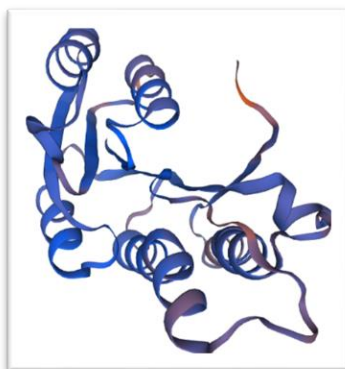


Figura 23. Modelado por homología del dominio KR por SWISS-MODEL. Templado de amfotericina B con una identidad de 56.18 % y cobertura del 99%.

8.2.7 Dominio TE

La terminación de la biosíntesis de policétidos típicamente involucra la terminación mediada por una tioesterasa que cataliza una hidrólisis simple del tioéster para generar un producto linear libre (Buchholz et al., 2012).

El modo en el que actúan es mediante una triada catalítica con histidina aproximadamente 20 residuos antes de C-terminal, el aspartato está aproximadamente 27 residuos separado de la serina (Keatinge-Clay, 2012). En los alineamientos de la TE se encontró esta característica (Figura 24).

	75	85	107	199	209
<i>TE</i>	K I V C F P S F V V V G R S M G G C I A V Y P V G L G L V D T F P I D T F W P Q P H D L I E I R G D H F				
<i>Piericidin</i>	- - F C F P S F V L L G H S A G G W I A T A P V S V V L L D S Y R P G S F W P V P H T E D E V P G D H F				
<i>Herboxidiene</i>	- - F C F P S Y A L L G H S S G A M I A A P P R A V V L L D V V P T D V F W P H P H T A V D A P G D H F				
<i>FD891</i>	- - V C F P S F V L V G Y S S G G W L A A A P A G V V L L D T Y L P E G F W D L P H T A V D V P G N H Y				
<i>Oleandomycin</i>	- - I C C S G F V L V G Y S A G G L M A T P P S G E V L V D V Y P P G R M W E Y D H T A V D M P G N H F				
<i>Neoauerothin</i>	- - I C F P T V V L L G Y S A G G W I A I A P A A V V L L D S H R P D S F W P L P H T A V D T A G T H F				
<i>Halstoctacosanolide</i>	- - I C F P A F A L L G H S S G G M V A V V P A A V V P M D V Y L P E Q F W P L P H T P V E V S G N H F				
<i>Megalomycin</i>	- - I C C A G F V L V G H S A G A L M A H P P R G V V L L D V Y P P G H T W P F G H D R V T V P G D H F				
<i>Fostriecin</i>	- - V C F P P Y A L L G Y S S S G W L A T G P V G V V L L D T Y P P D S F W P L D H E E T Q V P G D H C				
<i>Pikromycin</i>	- L V G C T G V V L L G H S S G G A L L A A P P A G I V L V D P Y P P G H L W D L P H T V A D V P G D H F				

Figura 24 Alineamientos múltiples del dominio TE de la PKS 1.4. Los aminoácidos catalíticos son serina (verde), aspartato (rosa) e histidina (azul).

Respecto a la estructura, el dominio TE (240-290 residuos) contiene un núcleo catalítico α/β hidrolasa. Las TEs de las PKS modulares del tipo 1 son diméricas. Al

hacer el modelo por homología se obtuvo una baja identidad, por lo que esto puede explicar que no sea tan claro identificar al dímero (Figura 25).

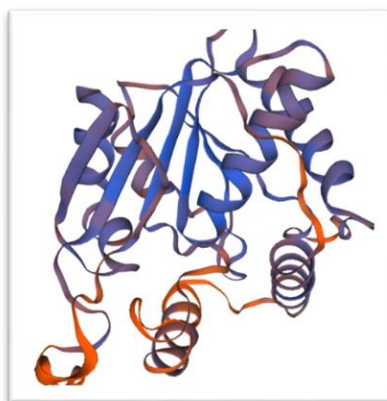


Figura 25. Modelado por homología del dominio KR por SWISS-MODEL. Templado de pikromicina con una identidad de 34.29 % y cobertura del 96%.

8.2.8 Función putativa de la región 1.4

Se utilizó PASS, plataforma que es capaz de proveer predicciones para muchos tipos de actividades biológicas basándose en la estructura de componentes farmacológicos descritos previamente. Sin embargo, dentro del algoritmo del programa no se toma en consideración las peculiaridades de la molécula, ya que sustancias estereoquímicamente diferentes en estructura son consideradas como iguales.

Los resultados en la Tabla 7 muestran las posibles actividades que puede tener el compuesto.

Tabla 7. Actividades biológicas previstas por PASS. Pa (activa) Pi (Inactiva).

Pa	Pi	Nombre de la actividad
0,871	0,008	Inhibidor de quimosina
0,871	0,008	Inhibidor de acrocilindropepsina
0,871	0,008	Inhibidor de saccharopepsina
0,837	0,008	Antineoplásico
0,835	0,014	Inhibidor de poliporopepsina

0,815	0,000	Inhibidor de metilmalonil-CoA carboxitransferasa
0,826	0,024	Inhibidor de unibiquinol-citrocromo-c-reductasa
0,770	0,005	Antineoplásico (cáncer de mama)

Resalta la presencia de actividad antineoplásica en estos resultados. Al profundizar en el análisis del compuesto, CLC-Pred indicó las posibles líneas celulares, mostradas en la Tabla 8, sobre las que tendría actividad antineoplásica el policétido.

Tabla 8. Resultado de predicción de líneas celulares cancerosas. Pa (activa) Pi (Inactiva).

Pa	Pi	Línea celular	Tejido	Tipo de tumor
0.890	0.005	MCF7	Mama	Carcinoma
0.658	0.011	PC-3	Próstata	Carcinoma
0.525	0.056	A549	Pulmón	Carcinoma

Sería importante considerar estas líneas celulares para evaluar la actividad terapéutica del compuesto, ya que se han reportado casos de éxito mediante el uso de esta plataforma. Tal es el caso del Sorafenib, donde se obtuvo una predicción para carcinoma de hígado y la predicción coincidió con la aplicación terapéutica (Lagunin et al., 2018).

8.2.9 Expresión heteróloga de la región 1.4

La región 1.4 del genoma de *Embleya* sp. NF3 es un posible candidato para la producción de un compuesto con actividad anticancerígena. El análisis bioinformático de cada uno de los módulos del BGC demuestran que su funcionalidad no se ve comprometida y la estructura química prevista por herramientas bioinformáticas indica que probablemente tenga la actividad biológica buscada.

Sin embargo, para realizar la expresión heteróloga de este tipo de BGC suelen necesitarse genes extra a los requeridos para la síntesis del compuesto. Es el caso de los genes relacionados con el transporte, ya que estos son los encargados de proporcionarle resistencia a la cepa frente al compuesto que produce. Además, otra manera de mejorar los niveles de producción es modificando los genes reguladores involucrados en la ruta de síntesis (Bekiesch et al., 2016).

Así que se identificaron los genes asociados a estas actividades dentro de la región 1.4 y se encontraron tanto transportadores como reguladores (Figura 10). Respecto a los transportadores, mediante BLAST y PFAM se encontró que pertenecen a familias de proteínas involucradas en el transporte extracelular, el gen *ctg1_1798* fue identificado como un transportador del tipo ABC. La presencia de este tipo de transportadores en los BGC juega un papel crucial previniendo la auto intoxicación del productor de los antibióticos mediante el bombeo hacia fuera del producto final activo antes de que llegue a niveles altos de toxicidad en el citoplasma. Si la desintoxicación de alguna manera se ve comprometida, la producción se puede ver afectada negativamente y la viabilidad celular puede caer drásticamente (Méndez & Salas, 2001). Un ejemplo de que los exportadores están acoplados a la biosíntesis de compuestos es el del Lasso péptido Mccj25. En este caso la disrupción del transportador ABC del BGC fue letal para la bacteria. Y como éste hay más ejemplos, como la daunorrubicina/doxorubicina, donde el transportador DrrAB de la familia ABC de *S. peucetius* se encontró que además de otorgar resistencia, contribuía en un mecanismo de regulación transcripcional por retroalimentación, donde la exportación deficiente en parte reguló a la baja a los genes involucrados en la biosíntesis (Severi & Thomas, 2019).

Respecto a los reguladores, siguiendo la misma metodología se identificó que pueden pertenecer a las familias de reguladores LuxR, SARP y TetR. Se sabe que una manera de mejorar los niveles de producción es modificando los genes reguladores involucrados en la ruta de síntesis (Bekiesch et al., 2016). Los reguladores multifuncionales globales no solo regulan la producción de múltiples metabolitos secundarios, sino que también participan en la diferenciación morfológica de los actinomicetos. Factores reguladores de rutas específicas están involucrados principalmente en la biosíntesis de metabolitos específicos, como los miembros de la familia de reguladores SARP (Proteína reguladora de antibióticos en *Streptomyces*) y LuxR. La familia de reguladores TetR (proteína represora de la tetraciclina) juega un rol importante en coordinar el crecimiento y desarrollo de actinobacterias en respuesta a factores ambientales, los miembros de la familia TetR están principalmente involucrados en la regulación de biosíntesis de antibióticos, bombas de eflujo y estrés osmótico. Son proteínas de doble dominio, un dominio se une y recibe señales y el otro se une al ADN y transduce señales (Li et al., 2018).

Con todo el análisis anterior, es muy probable que estos genes también sean necesarios para lograr la correcta expresión del clúster, por lo que el tamaño a considerar es de aproximadamente 44 kb, lo que excede el tamaño máximo de inserto que puede contener la biblioteca genómica de *Embleya* sp. NF3 construida previamente. Para la correcta producción de manera heteróloga del compuesto de la región 1.14 deberá considerarse por tanto una estrategia diferente.

8.3 Análisis bioinformático de la región 1.12

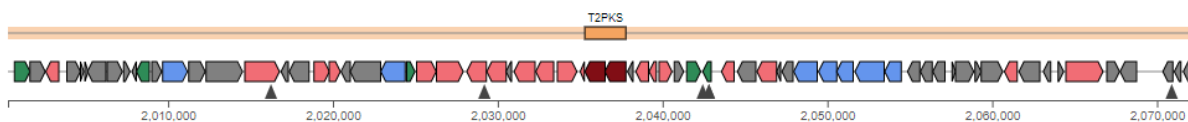


Figura 26. Región 1.12 correspondiente a PKS del tipo II en *Embleya* sp. NF3. Genes core (guinda), genes relacionados a reguladores (verde), genes adicionales en la biosíntesis (rosa), genes relacionados al transporte (azul).

El tamaño total de la región identificada es de 72.5 kpb, el programa antiSMASH 6.0 identificó que este clúster cumple con la estructura de una PKS del tipo II (Figura 26) así como que tiene un identidad del 63% con el clúster de la estefimicina reportado por Gullón et al. en 2006. Lo anterior apunta a que la región 1.12 probablemente sea la responsable de producir los compuestos identificados en *Embleya* sp. NF3 y de los cuales ya se ha probado su actividad (Rodríguez-Peña et al., 2018), como lo son la estefimicina B y derivados de ella.

Tomando en cuenta lo reportado, se hizo un análisis de los genes presentes en este clúster de *Embleya* sp. NF3 y el clúster de la estefimicina en *S. steffisburgensis* (Gullón et al., 2006), los resultados de esta comparación se presentan en la Tabla 9.

Tabla 9. Comparación de los genes dentro del BGC 1.12 de *Embleya* sp. NF3 y *S. steffisburgensis*.

Gen de <i>Embleya</i> sp. NF3	Tamaño (aa) Stef B	Tamaño (aa) Stef	Función putativa	Gen homólogo de <i>S. steffisburgensis</i>	% Identidad	% Similitud
<i>ctg1_1768</i>	311	310	Aromatasa	<i>stfQ</i>	73	82
<i>ctg1_1769</i>	226	225	O-Metiltransferasa	<i>stfMl</i>	81	89
<i>ctg1_1770</i>	187	186	Desconocida	<i>stfC</i>	69	80
<i>ctg1_1771</i>	619	618	Desconocida	<i>stfD</i>	79	87
<i>ctg1_1772</i>	493	492	Proteína de transporte MFS	<i>sfrB</i>	75	85
<i>ctg1_1773</i>	168	167	Regulador MarR	<i>stfRlll</i>	77	83
<i>ctg1_1774</i>	396	395	Citocromo P450	<i>stfPl</i>	79	87
<i>ctg1_1775</i>	541	540	Deshidrogenasa	<i>stfE</i>	80	88
<i>ctg1_1776</i>	393	392	Monooxigenasa	<i>stfOll</i>	77	87
<i>ctg1_1777</i>	392	391	O-Metiltransferasa	<i>stfMll</i>	78	88
<i>ctg1_1778</i>	113	112	Desconocida	<i>stfF</i>	79	85

<i>ctg1_1779</i>	421	420	Glicosiltransferasa	<i>stfG</i>	84	90
<i>ctg1_1780</i>	366	365	Citocromo P450	<i>stfPll</i>	56	65
<i>ctg1_1781</i>	393	392	O-Metiltransferasa	<i>stfMlll</i>	76	83
<i>ctg1_1782</i>	86	85	ACP	<i>stfS</i>	69	79
<i>ctg1_1783</i>	423	422	Cetoacil sintasa (CLF)	<i>stfP</i>	85	92
<i>ctg1_1784</i>	408	407	Cetoacil sintasa (KS)	<i>stfK</i>	80	87
<i>ctg1_1785</i>	112	111	Monooxigenasa	<i>stfOl</i>	76	86
<i>ctg1_1786</i>	256	255	Cetoreductasa	<i>stfT</i>	70	76
<i>ctg1_1787</i>	146	145	Ciclasa	<i>stfX</i>	79	87
<i>ctg1_1788</i>	260	259	Ciclasa	<i>stfY</i>	83	91
<i>ctg1_1789</i>	187	*	Desconocida	*	*	*
<i>ctg1_1790</i>	283	282	Regulador SARP	<i>stfRl</i>	75	83
<i>ctg1_1791</i>	172	171	Regulador LuxR	<i>stfRll</i>	54	66

*No presente en el BGC

Los genes core identificados son los genes *ctg1_1783* y *ctg1_1784*, estos corresponden a una cetosintasa (KS) y al CLF (factor de elongación). La diferencia entre ambos radica en que el CLF no contiene el residuo de cisteína catalítico.

Este homodímero es característico de las PKS del tipo II y catalizan otro paso de elongación, esta acción es iterativa ya que este complejo retiene y estabiliza a los intermedios del policétido (Chen et al., 2018). La longitud de la cadena está determinada por el número de condensaciones de Claisen que haga esta dupla y el resultado es una cadena de policétido muy reactiva que puede ser modificada por las enzimas adicionales (Wang et al., 2020).

Se realizaron alineamientos con diferentes cetosintasas de otras actinobacterias (Figura 27), y como se esperaba el gen *ctg1_1783* es el factor de elongación, mientras que la cetosintasa activa corresponde al gen *ctg1_1784*.

```

                150           160           170
ctg1_1784_Embleya_NF3 VPSS IAAEV AWEVAAEGPVAT I STG C TSGLD
ctg1_1783_Embleya_NF3 YAVN - TGQ I S I REGLR GPSGVL V TEQAGGLD
KS_S_resistomycificus VPSSLAAEV AWEVGAEGPVAT V STG C TSGLD
KS_S_steffisburgensis VPSSLAAEV AWDVGAEGPVAT V STG C TSGLD
KS_M_halophytica VPSS IAAEV AWEVGAEGPVAT V STG C TSGLD
KS_S_arenicola VPSSLATEV AWEVGVGEPVSVVSTG C TSGLD
KS_N_tenerifensis I PSSLATEV AWEVGAEGPVST V STG C TSGLD
KS_N_colli VPSSLATEV AWEV LGAEGPVAT V STG C TSGLD

```

Figura 27. Alineamientos múltiples del dominio KS de la PKS 1.12 con PKS de otras actinobacterias. La cisteína catalítica (rojo) se encuentra en todas las KS activas.

El gen *ctg1_1782* fue identificado como el dominio ACP. En este caso la funcionalidad y estructura de la proteína es la misma que para las PKS del tipo I, por lo que al encontrar la serina activa en este gen mediante alineamientos con otras ACP nos indicaría su funcionalidad dentro del BGC (Figura 28).

```

                26           36           46
ACP_EmbleyaNF3 MSEDÉTSNL - DGD I LDAEFGELGYD S I AMLETA
ACP_S_steffisburgensis - - EDENAGLRGGDVLDTTFTDLGYDSLAMLETA
ACP_S.resistomycificus - - EDETSALRGGDVLDTDFKDLGYDSLAMLETA
ACP_S.incarnatus_ - - EDEGVDL - DGD I LDTLFA DLGYDSLALLEAS
  
```

Figura 28. Alineamientos múltiples del dominio ACP de la PKS 1.12 con PKS de diferentes actinobacterias. El aminoácido catalítico es la serina (verde) dentro del motivo DSL.

Respecto a las proteínas adicionales identificadas se conoce que en el caso de policétidos aromáticos como en la estefimicina B, la ciclación y/o aromatización es catalizada por las aromatasas y las oxigenasas para generar el núcleo del policétido aromático. Luego siguen las post modificaciones por las oxigenasas, metiltransferasas y glicosiltransferasas que modifican el núcleo del policétido mediante catálisis de diferentes reacciones hasta el producto final. Adicionalmente se identificaron los genes *ctg1_1774* y *ctg1_1780* que tienen una función putativa como proteínas citocromo P-450 monooxigenasas. La presencia de este tipo de proteínas ha sido poco reportada. Sin embargo, un ejemplo conocido es DoxA que finaliza la síntesis de la doxorubicina mediante una triple hidroxilación (Rohr & Hertweck, 2010). Por lo tanto, se consideró relevante incluir estos genes en la delimitación del BGC.

La metilación es otro tipo de reacción que expande la diversidad de los policétidos aromáticos. Las metiltransferasas llevan a cabo reacciones donde transfieren un grupo metilo de un donador como S-adenosil-l-metionina (SAM) a átomos de carbono, nitrógeno u oxígeno receptores, en este caso, las encontradas en la región 1.12 son O-metiltransferasas, por lo que el átomo que pueden modificar es el oxígeno (Wang et al., 2020). Finalmente el policétido puede ser ciclado o terminado por alguna de las dos ciclasas identificadas dentro del BGC (Hu et al., 2020).

Los clústeres biosintéticos de antibióticos también suelen incluir genes de resistencia y reguladores. En el caso de la daunorrubicina se identificaron tres genes en el BGC de la daunorrubicina en *S. peucetius* que le confieren la resistencia cuando son expresados en *S. lividans* o en *E. coli*. Dos de estas proteínas actúan juntas como una bomba de eflujo y se dedujo que pertenecen a la familia de transportadores ABC (Méndez & Salas, 2001).

En el caso de la estefimicina, se deduce que antiportadores del tipo MLS que están codificados por *sfrB* en *S. steffisburgensis* (Gullón et al., 2006) son los que confieren la resistencia al microorganismo frente al compuesto producido. En la región 1.12 su homólogo corresponde al *ctg1_1772* y es por esto que se consideró como un factor determinante al momento de realizar la expresión heteróloga del clúster.

Por último, se encontraron tres genes reguladores importantes en el BGC. El primero, *ctg1_1773* pertenece a la familia MarR y se conoce que puede actuar como activador o represor. Su función en *Streptomyces* está asociada a regulación de la resistencia frente al estrés oxidativo y producción de antibióticos. El gen *ctg1_1790* se identificó como un regulador de la familia SARP que suelen funcionar como activadores y están involucrados en el metabolismo secundario. Por último, *ctg1_1791* coincide con un regulador de la familia LuxR cuya acción se le relaciona con la producción de metabolitos secundarios (Romero-Rodríguez et al., 2015).

Teniendo en cuenta todo lo anterior y considerando que se encontraron genes similares entre el clúster de la estefimicina previamente reportado y la región 1.12 se delimitó el BGC del gen *ctg1_1768* al *ctg1_1791*, lo que resulta en un clúster con tamaño final de 24 kpb y es posiblemente el responsable de producir a la estefimicina B en la cepa NF3.

8.3.1 Expresión heteróloga del BGC 1.12

Como ya se había mencionado en los antecedentes, la cepa NF3 inicialmente aislada produce 1.8 mg/L de estefimicina B en medio YMG (Rodríguez-Peña, 2020). Al suplementarlo con CaCO₃ este rendimiento aumenta a 70 mg/L del mismo compuesto (Trenado-Urbe, 2016), y ya que estos valores están lejos de lo reportado para productos anticancerígenos comerciales se decidió explorar la opción de realizar la expresión heteróloga de este BGC con el fin de aumentar su producción. Al buscar la técnica más adecuada y por el tamaño relativamente grande (~ 24 kpb) no es posible utilizar alguna técnica de clonación tradicional. Afortunadamente, se tiene construida una biblioteca genómica en fósmidos de este microorganismo y se decidió continuar con esta estrategia para lograr la producción heteróloga del compuesto.

8.4 Escaneo de la biblioteca genómica de *Embleya* sp. NF3.

Se diseñaron tres pares de primers (Figura 29, Anexo 2), para hacer la búsqueda del BGC completo en la biblioteca genómica mediante PCR. Previó a esto, se encontraron las condiciones óptimas de amplificación de los primers reportadas en la metodología.

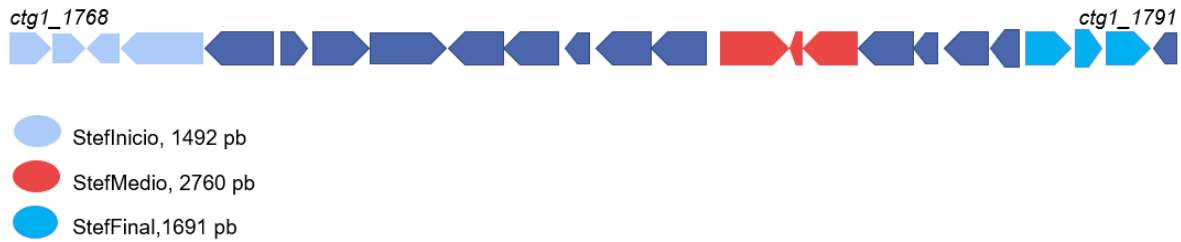


Figura 29. Regiones amplificadas por los primers diseñados. StefInicio (azul claro), stefMedio (rojo) y stefFinal (azul cielo).

Cuando se realizó la caracterización de la biblioteca genómica se utilizó el par de primers StefFinal, por lo que esta fue la primera región identificada en los *pooles* 21A1, 22B5 y 21D6 (Figura 30).

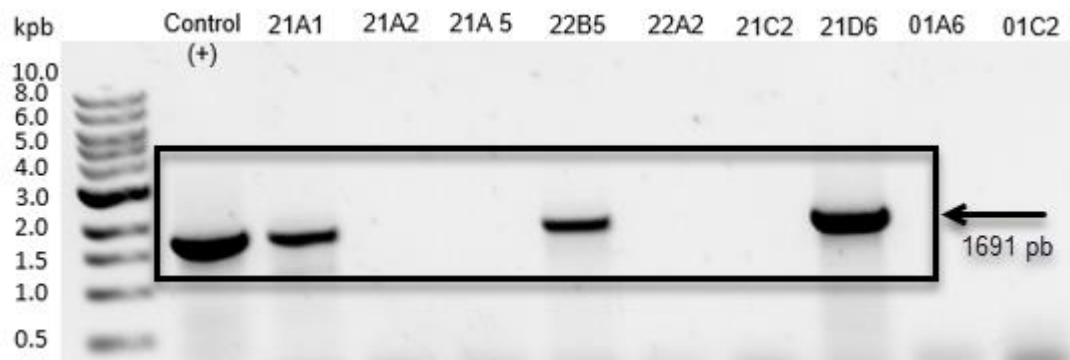


Figura 30. Identificación de stefFinal (1691 pb) en *pooles* de la biblioteca genómica (Rectángulo negro). Control (+) ADN genómico de *Embleya* sp. NF3.

Dado que tres *pooles* dieron positivo, se decidió buscar las demás regiones únicamente en estos tres. Para stefInicio (Figura 31), se encontró positivo únicamente el *pool* 21A1.

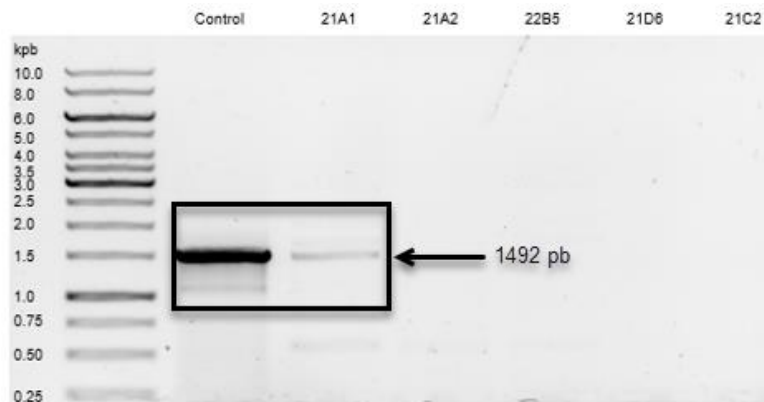


Figura 31. Identificación de stefInicio (1492 pb) en *pooles* de la biblioteca genómica (Rectángulo negro). Control (+) ADN genómico de *Embleya* sp. NF3.

Y finalmente la región stefMedio se encontró en los *pooles* 21A1, 22B5 y 21D6 (Figura 32).

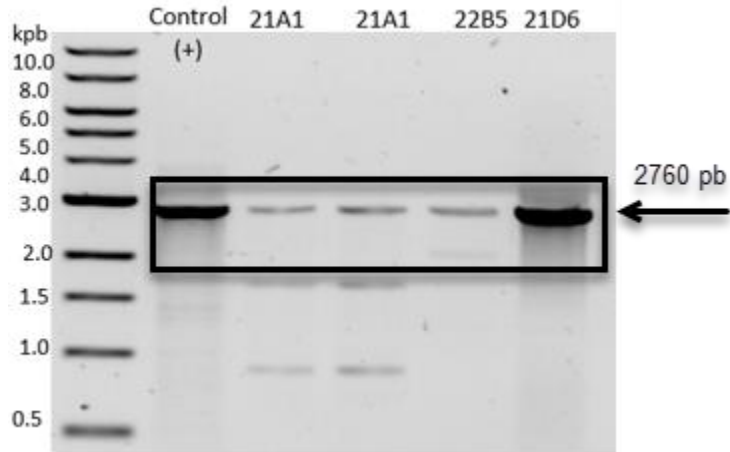


Figura 32. Identificación de stefMedio (2760 pb) en *pooles* de la biblioteca genómica (Rectángulo negro). Control (+) ADN genómico de *Embleya* sp. NF3.

El resumen de los resultados se presenta en la Tabla 10.

Tabla 10. Regiones encontradas en los *pooles* evaluados de la biblioteca genómica de *Embleya* sp. NF3.

No. De pool	StefInicio	StefMedio	StefFinal
21A1	✓	✓	✓
21A2		✓	✓
22B5		✓	✓
21D6		✓	✓

De los 87 *pooles*, se encontraron las tres regiones únicamente en el *pool* 21A1. Este *pool* está conformado por 27 colonias, por lo que se hizo la búsqueda con los tres pares de primers en cada una.

8.4.1 PCR de colonia para la búsqueda del BGC 1.12

El procedimiento seguido para realizar PCR de colonia se indica en el Anexo 8. Para stefInicio los resultados se presentan en la Figura 33. Además de las colonias A8, A9 y A12 se encontraron positivos en B3, B8, B10, B11, B12, C1, C2 y C3.

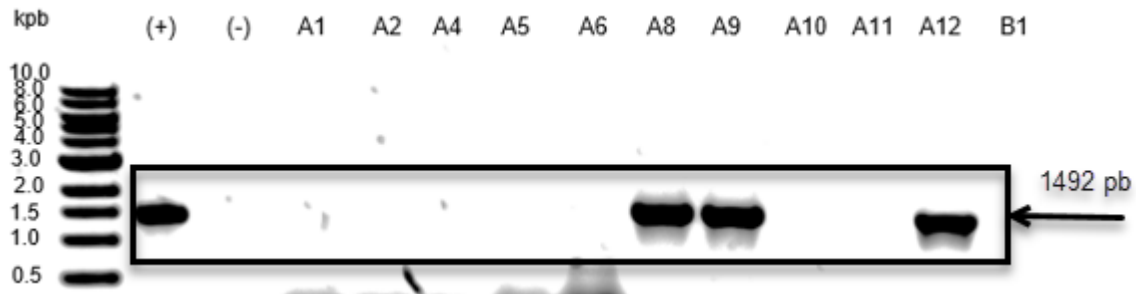


Figura 33. Identificación de *stefInicio* (1492 pb) en las colonias del pool 21A1 de la biblioteca genómica (Rectángulo negro). Control (+) ADN genómico de *Embleya* sp. NF3, (-) colonia H11 del pool 21A4.

Para el caso de *stefMedio*, los resultados se muestran en la Figura 34. En este caso las colonias positivas fueron únicamente A7, A8 y A12.

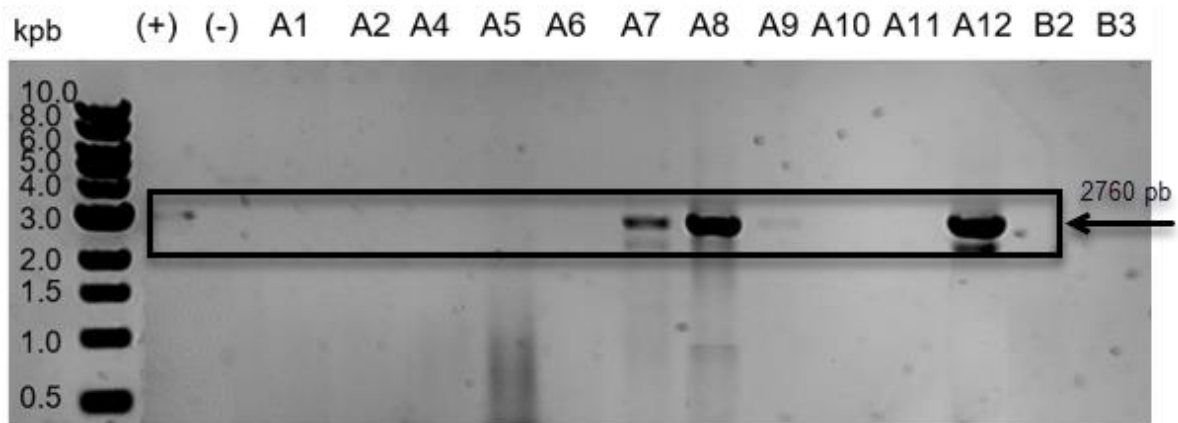


Figura 34. Identificación de *stefMedio* (2760 pb) en las colonias del pool 21A1 de la biblioteca genómica (Rectángulo negro). Control (+) ADN genómico de *Embleya* sp. NF3, (-) colonia H11 del pool 21A4.

Por lo último, los resultados para *stefFinal* se muestran en la Figura 35.

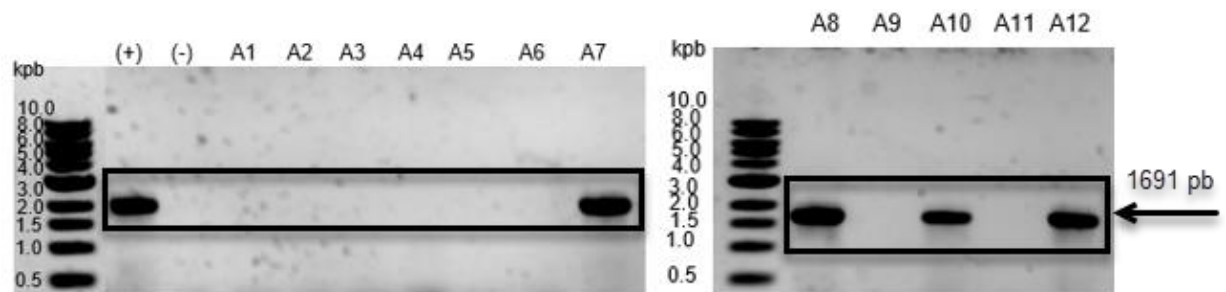


Figura 35. Identificación de *stefFinal* (1691 pb) en las colonias del *pool* 21A1 de la biblioteca genómica (Rectángulo negro). Control (+) ADN genómico de *Embleya* sp. NF3, (-) colonia H11 del *pool* 21A4.

Los datos nos indican que es probable que se encuentre el BGC completo en al menos tres colonias denominadas A8, A9 y A12. Por consiguiente, se procedió a la extracción del fósmino siguiendo el protocolo del Anexo 6, para la cuantificación se utilizó un método fluorométrico con bromuro de etidio y la cantidad obtenida de plásmido de las colonias A8 y A12 fue de 37 y 33 μg respectivamente.

Como se había mencionado, el tamaño aproximado de inserto que puede tener como máximo el fósmino es de 40 kbp, y esto se comprobó (Figura 36) al realizar la extracción de plásmido previo al envío para su secuenciación.

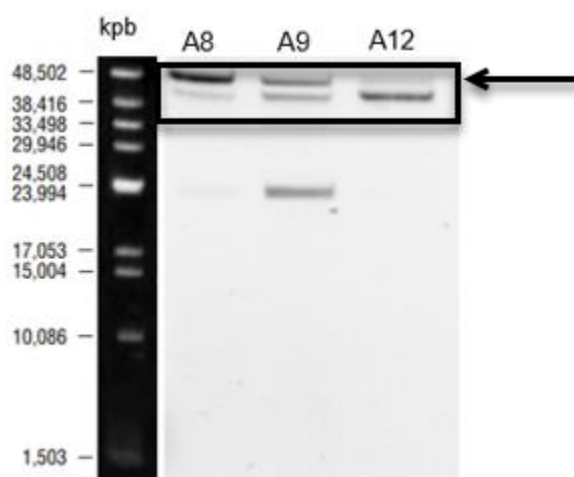


Figura 36. Fósmino pCC2FOS de la biblioteca de *Embleya* sp. NF3 de las colonias A8, A9 y A12 (Rectángulo negro).

Se analizaron los archivos entregados por la unidad de secuenciación y se encontró que para la colonia A8 hay dos secuencias presentes dentro de la muestra. La primera, denominada NODE41 es de 38,072 pb y la segunda NODE48 es de 29,850 pb. Al hacer el análisis de cada una de las secuencias se encontró que no pertenecen al BGC 1.12 pero son secuencias del genoma de *Embleya* sp. NF3 no cercanas ni adyacentes a la región buscada. Ya que en la amplificación de los tres pares de primers fue positiva en esta colonia se decidió utilizar el programa Velvet para buscar secuencias adicionales. Con este segundo ensamble fue posible corroborar la presencia de una secuencia adicional de 12,936 pb que corresponde con la región 1.12.

Para la colonia A12, el primer ensamble mostró la presencia de cuatro secuencias dentro de la muestra. NODE44 con 38,072 pb, NODE52 con 31,018 pb, NODE55 con 29,850 pb y por último NODE46 que corresponde a la región 1.12 con un tamaño de 37,428 pb. Las secuencias adicionales encontradas son del genoma de *Embleya* sp. NF3, pero tampoco son adyacentes o cercanas al clúster buscado.

Lo esperado en los resultados de la secuenciación era encontrar la presencia de solo una secuencia de ADN por colonia; sin embargo, los resultados fueron diferentes. Esto puede deberse a que durante el proceso de almacenamiento de la biblioteca genómica se tomó más de una colonia al momento de guardarlas, esto explica porque se encuentra la amplificación de los tres pares de primers, pero además se tienen otras secuencias. Tal como se encuentran las colonias no pueden ser utilizadas en los procesos siguientes para realizar la expresión heteróloga del BGC ya que se necesita tener al plásmido con la secuencia aislado.

Por tal razón se aislaron colonias individuales directamente de la biblioteca genómica y se hizo PCR de colonia nuevamente para encontrar la que contenga el BGC. De las colonias aisladas, se encontró positiva la colonia 8 y para seleccionar únicamente a las colonias positivas la concentración de cloranfenicol se aumentó a 25 $\mu\text{g}/\text{mL}$. Tres de las cinco colonias aisladas dieron positivo (Figura 37, 38); sin embargo, en la amplificación de *stefInicio* se observó un patrón de bandas diferente para las colonias que dieron positivo. Se purificó este plásmido y para verificar que solo tuviera una secuencia, se debe volver a secuenciar; sin embargo, esto queda como perspectiva a realizar para futuros proyectos.

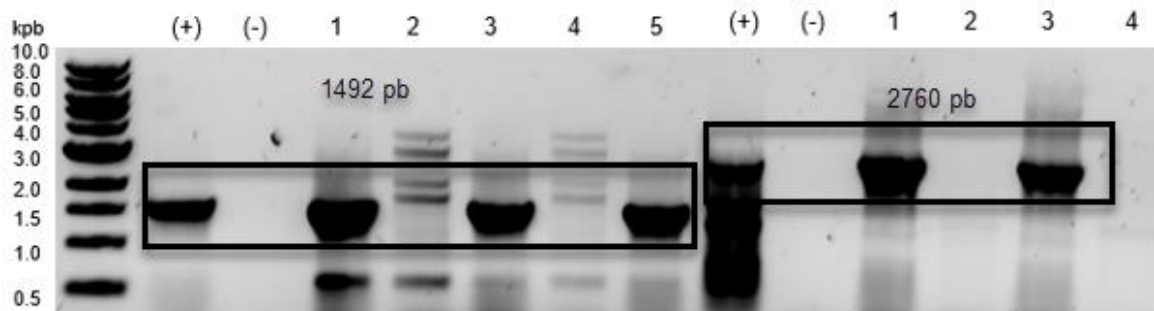


Figura 37. Identificación de *stefInicio* (1492 pb) para cinco colonias y *stefMedio* (2760 pb) para cuatro colonias (Rectángulo negro). Control (+) ADN genómico de *Embleya* sp. NF3, (-) colonia H11 del pool 21A4.

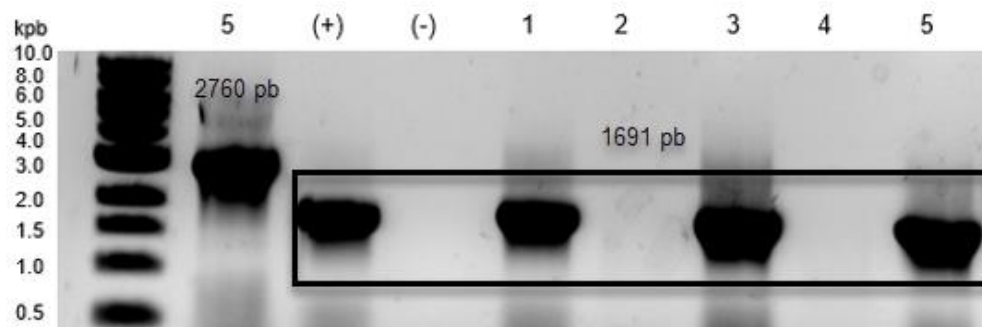


Figura 38. Identificación de *stefFinal* (1691 pb) para cinco colonias y *stefMedio* (2760 pb) de una colonia (Rectángulo negro). Control (+) ADN genómico de *Embleya* sp. NF3, (-) colonia H11 del pool 21A4.

8.5 Vector de expresión PIJ10257

El fósido donde se encuentra la biblioteca es únicamente de almacenamiento, por lo que una vez que se tenga el fósido con la secuencia buscada se debe de pasar esta secuencia a un vector de expresión. Con este fin se escogió al vector PIJ10257. Se eligió por tener la maquinaria para poder realizar conjugación desde una cepa de *E. coli* a una de *Streptomyces*, es un vector integrativo lo que permitirá que el clúster quede integrado en el genoma del microorganismo huésped para que no elimine el ADN externo y al momento de la producción del metabolito se pueda utilizar la maquinaria propia del organismo huésped (Negrete, 2018). Además, cuenta con el promotor constitutivo PermE* que ha sido utilizado en la expresión heteróloga de policétidos (Nivina et al., 2019; Wolpert et al., 2008).

Una vez que se hizo la transformación de acuerdo con el Anexo 4, se hizo la extracción del plásmido y se pudo comprobar su presencia en las colonias de *E. coli* DH5 α (Figura 39), ya que puede observarse el tamaño del vector (6670 pb), así como una de las isoformas que tiene el plásmido.

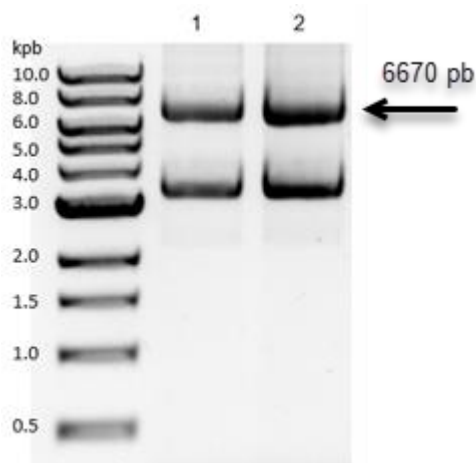


Figura 39. Resultado de la purificación del vector de expresión PIJ10257. Tamaño de 6670 pb en carril 1 y 2.

8.6 Producción y purificación del policétido

Debido a los resultados obtenidos para la expresión heteróloga del BGC 1.12 putativamente asociado a la estefimicina B y siguiendo con el objetivo de encontrar una alternativa para aumentar su producción, se decidió realizar una fermentación de *Embleya* sp. NF3 durante 21 días. Sabemos que *Embleya* sp. NF3 produce una diversidad de compuestos, entre ellos algunos que pertenecen a la familia de las estefimicinas con actividad dual, es decir antimicrobiana y anticancerígena. Con el objetivo de comprobar si con las condiciones de fermentación propuestas se

obtienen metabolitos secundarios con actividad antimicrobiana se decidió hacer ensayos de antibiosis con el medio de cultivo durante la fermentación, utilizando *M. luteus* por ser un microorganismo sensible. Se tomaron muestras del medio en los días 7 y 14, los resultados están en la Tabla 11.

Tabla 11. Resultados antibiosis del medio de cultivo.

Día de fermentación	Halo de inhibición (mm)
14	18.5
18	21

Podemos observar que aumenta la capacidad antibiótica del medio conforme pasan los días, esto es porque se sabe que los metabolitos secundarios son producidos en la idiofase subsecuente al crecimiento (Sanchez & Demain, 2019). Ya que se comprobó la actividad antimicrobiana de la fermentación indicando la presencia de compuestos de interés, se continuó con el proceso de purificación específicamente buscando a la estefimicina B.

Para este fin al término de la fermentación se agregó AmberLite™ XAD™ 7HP. Las resinas de amberlita son perlas de copolímeros polimerizados duros e insolubles, en este caso se trata de un polímero acrílico polar de éster. Cada perla se forma a partir de microperlas cementadas juntas durante la polimerización, dando una resina con estructura macroreticular, lo que resulta en una gran área de superficie en cada perla. La adsorción de los compuestos orgánicos en la superficie homogénea de las resinas es mediante un proceso de baja energía que involucra fuerzas de Van der Waals, lo que facilita la desorción de las moléculas adsorbidas y ha demostrado ser capaz de adsorber cantidades grandes de compuestos sin carga (Moore & Karasek, 1984). Por estas razones se eligió para iniciar con la purificación del compuesto. Después de que se dejó a la amberlita en contacto con el medio de cultivo, se empaquetaron las perlas en una columna y se hicieron las eluciones que se indican en la Tabla 3. De las fracciones obtenidas también se hicieron ensayos de antibiosis con *M. luteus* obteniendo los resultados que se muestran en la Tabla 12. Se obtuvieron dos muestras por fracción.

Tabla 12. Resultados antibiosis primera columna con amberlita.

Fracción	Halo de inhibición (mm)
70:30	6
50:50	8.5

30:70	13
Metanol	14.25
Acetona	9

Para la fermentación donde se colocó amberlita al final, se hizo el mismo procedimiento, los resultados de las antibiosis se muestran en la Tabla 13. En este caso se obtuvieron seis muestras por fracción.

Tabla 13. Resultados antibiosis segunda columna con amberlita.

Fracción	Halo de inhibición (mm)
70:30	6.83
50:50	14.83
30:70	21.00
Metanol	19.58
Acetona	20.00

Para conocer si la amberlita era capaz de adsorber todos los compuestos de interés del medio también se realizaron ensayos de antibiosis con el medio de cultivo después de que se retiraron las perlas y en ambos casos no se observó ninguna actividad antibiótica, por lo que puede suponerse que todos los compuestos con esta actividad fueron adsorbidos.

Una vez que se obtuvieron las fracciones se llevaron a sequedad y se hizo una extracción de ácidos grasos con hexano. Después se hizo una CCF para observar la separación de compuestos. La fase móvil utilizada, ya había sido reportada con anterioridad dando buenos resultados al separar a la estefimicina B del resto de sus derivados, por lo que se decidió utilizarla como referencia.

Los resultados preliminares mostraron presencia de un compuesto naranja-amarillo con un Rf de 0.80-0.85 (Figura 40), según lo reportado anteriormente el compuesto estefimicina B es el que tiene un Rf mayor respecto a otros derivados como la estefimicina debido a la naturaleza menos polar de la estefimicina B (Brodasky & Reusser, 1974).

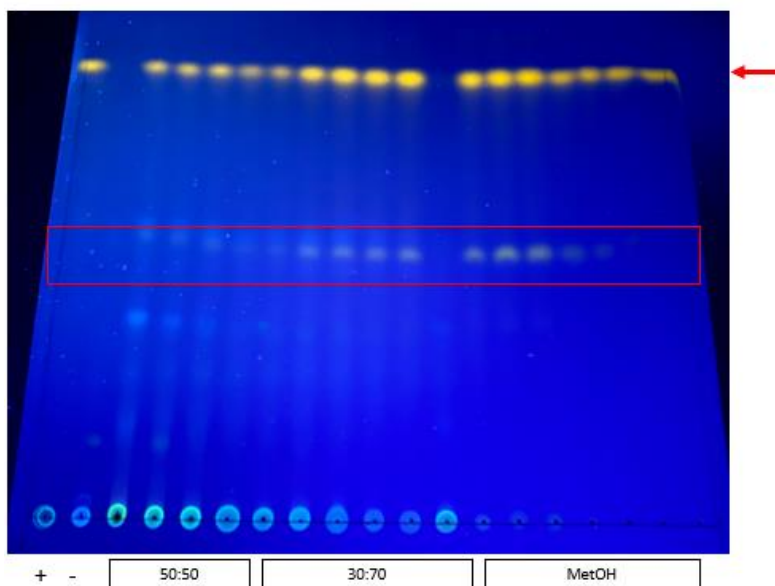


Figura 40. CCF de las fracciones 50:50, 30:70 y metanol. Control positivo (+) estefimicina B, control negativo (-) medio de cultivo. Elución con cloruro de metileno, acetona, hexano y metanol en proporción 30:10:3:2 y observado a λ 365 nm. Flecha roja: Estefimicina B. Rectángulo rojo: derivados de la estefimicina B.

Como se observa en la Figura 40, además del compuesto del interés que se encuentra en la parte superior (Flecha roja) hay presencia de compuestos que se separan muy bien (Rectángulo rojo), estos pueden ser estefimicina C, 8-demetoxi-10-deoxiestefimicinona, 7-deoxiestefimicinona, 8-demetoxiestefimicina, 8-demetoxi-10-deoxiestefimicina B, 7-deoxi-8-demetoxiestefimicinona y 7-deoxi-10-deoxiestefimicinona, compuestos ya descritos para la cepa NF3 y que gracias a su descubrimiento se ha logrado proponer una ruta alternativa para la producción de estefimicina B (Rodríguez-Peña et al., 2022). Por lo que es probable que estos compuestos sean lo que se observen a lo largo de la CCF. Para conocer la naturaleza de cada compuesto observado es necesario realizar más análisis.

El desarrollo de la CCF fue el mismo desde las últimas muestras de las fracciones 70:30 y continuó hasta las fracciones de metanol. Sin embargo, ninguna de las fracciones de la primera columna mostró el mismo desarrollo lo cual puede deberse a que la amberlita al estar todo el tiempo en contacto con el medio de producción absorbió a los compuestos precursores de la estefimicina B (Rodríguez-Peña, 2020) y esto no permitió la síntesis final del compuesto. La actividad antibiótica que pudo observarse probablemente es debida a la presencia lípidos metilados en el medio de cultivo que también ha sido reportada (Trenado-Uribe, 2016), ya que estos ensayos se realizaron antes de hacer la extracción de ácidos grasos con hexano.

También se hizo un ensayo de bioautografía con *M. luteus* en medio TSB, ya que sabemos que de todos los compuestos de la familia de las estefimicinas que es capaz de producir *Embleya* sp. NF3, la estefimicina B es la que mayor actividad antimicrobiana tiene. Para esto, la placa de la CCF con las muestras de la fracción de metanol se colocó en la parte inferior del medio. En la Figura 41 la flecha roja y el rectángulo indican la región donde no se observó crecimiento del microorganismo por lo que la actividad antimicrobiana es debida al compuesto mayoritario superior observado previamente. Los otros compuestos observados en la Figura 40 no presentaron actividad antibiótica.

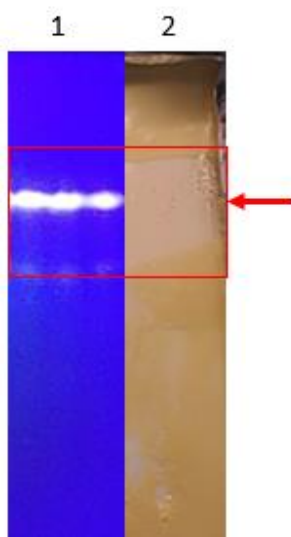


Figura 41. Bioautografía. Columna 1, CCF de las muestras de la fracción de metanol, columna 2 medio TSB inoculado con *M. luteus* colocado sobre la CCF. Rectángulo y flecha roja muestran a la estefimicina B en la CCF y el halo de inhibición en la bioautografía.

Ya que se obtuvo una buena separación con la fase móvil elegida se hizo una CCF preparativa con todas las muestras, y se recuperó el compuesto. Para comprobar que el compuesto estaba puro se volvió a correr una CCF de las muestras obtenidas y como puede observarse en la Figura 42, a diferencia de la Figura 40, únicamente se observa la presencia de un compuesto naranja que coincide con el control positivo (flecha roja) por lo que podemos suponer que la purificación fue exitosa ya que no se observa la presencia de algún otro compuesto en toda la cromatografía.



Figura 42. CCF del compuesto recuperado por CCF preparativa. Control positivo (+) estefimicina B, (-) medio de cultivo, carriles 1- 5 corresponden a las muestras obtenidas. Elución con cloruro de metileno, acetona, hexano y metanol en proporción 30:10:3:2 y observado a λ 365 nm. La flecha roja indica la presencia de la estefimicina B.

La cantidad de compuesto recuperada fue de 29.8 mg. Es decir, un rendimiento de 99.3 mg/mL superior al reportado por Rodríguez-Peña de 1.8 mg/mL y los 70 mg/L de Trenado-Urbe en el 2018.

Finalmente, para comprobar que el compuesto recuperado corresponde a la estefimicina B se analizó la muestra por LC-MS. Se obtuvieron los cromatogramas extraídos correspondientes a la masa 587.17, la cual corresponde a la masa de la estefimicina B en polaridad negativa, esto porque en espectrometría de masas cuando se registra en polaridad negativa los iones pierden un protón, por lo que vemos es la relación de la masa sobre la carga menos un protón es decir M-H.

Una vez que se obtuvieron los cromatogramas se comprobó que la masa 587.17 tiene un tiempo de retención de 1.20 minutos, se extrajeron las masas correspondientes al tiempo de retención 1.20 min tanto para el estándar como para la muestra (Figura 43).

También se obtuvo la composición elemental de la masa 587.17 (Figura 44), la cual nos muestra que el ión se encuentra compuesto por C₂₉, H₃₁, O₁₃, correspondiente a la composición de la estefimicina B ionizada negativamente y con un error en la masa de -.03 ppm, por lo que es altamente probable que el ión registrado correspondiente al analito sea la estefimicina B.

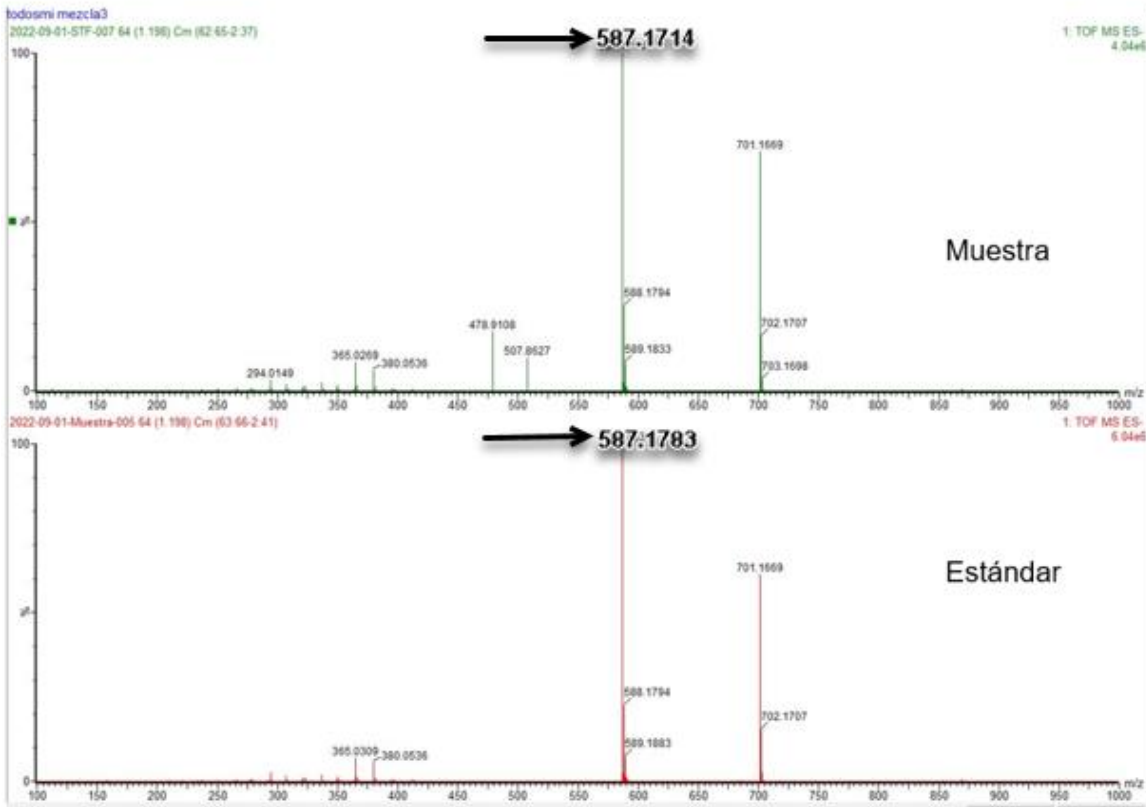


Figura 43. Masas presentes en el tiempo de retención 1.20 min. Estándar estefimicina B de Enzo Life Sciences ALX-380-220-M001 (Ann Arbor, MI, USA).

Mass	Calc. Mass	mDa	PPM	DBE	Formula	C	H	O
587.1763	587.1765	-0.2	-0.3	14.5	C₂₉ H₃₁ O₁₃	29	31	13

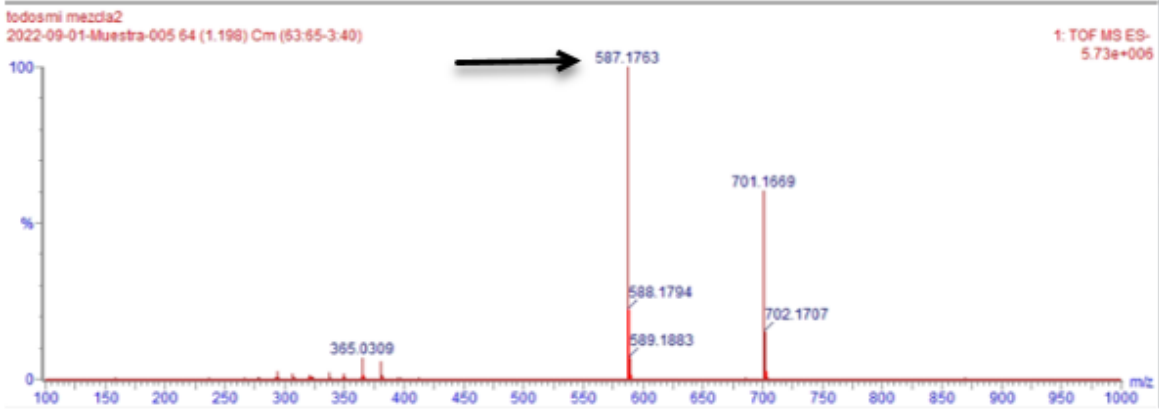


Figura 44. Composición elemental de la muestra obtenida mediante LC- MS.

8.7 Mejora en la solubilidad del policétido y evaluación de su actividad.

Ahora bien, aunque se ha probado la actividad biológica de este compuesto en ensayos *in vitro* (Brodasky & Reusser, 1974; Rodriguez-Peña, 2020), no se ha podido demostrar actividad *in vivo*. Esto debido posiblemente a la baja solubilidad del compuesto en medios acuosos. Se han hecho ensayos con soluciones de sodio y aunque el compuesto fue muy soluble la actividad disminuyó significativamente (Brodasky & Reusser, 1974). Para contrarrestar esta característica del compuesto, se hizo un complejo con la ciclodextrina β que es un oligosacárido cíclico formado por siete unidades de S glucopiranosas unidas por enlaces α (1-4) y se obtiene de la degradación enzimática del almidón. En su cavidad interior es capaz de contener moléculas más pequeñas ya sean orgánicas o inorgánicas (Hoskisson & van Wezel, 2019).

Por esta razón se hizo un complejo en proporción 1:3 con la ciclodextrina y se hicieron ensayos de antibiosis utilizando como control estefimicina B- DMSO. Los resultados se muestran en la Tabla 14.

Tabla 14. Resultados antibiosis complejos con estefimicina B.

Fracción	Halo de inhibición (mm)
Complejo Estefimicina B - β -CD proporción 1:3	16.13
Complejo Estefimicina B-DMSO	10

Como podemos observar, el halo de inhibición fue mayor en el caso del complejo con la ciclodextrina frente al control que es con DMSO. Esto se debe a que las ciclodextrinas presentan diferentes polaridades en las superficies interna y externa. En este tipo de moléculas los grupos hidroxilos primarios y secundarios están localizados en el borde más estrecho y ancho respectivamente (Figura 45).

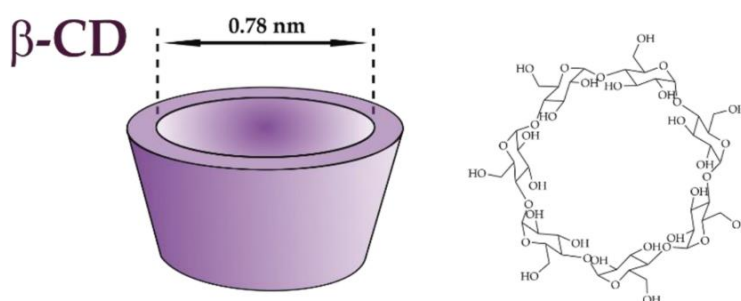


Figura 45. Dimensiones y estructura de la β -ciclodextrina (Modificado de Cid-Samamed et al., 2022).

El complejo de inclusión con compuestos hidrofóbicos como la estefimicina B ocurre principalmente por interacciones hidrofóbicas entre la cavidad en las paredes de las ciclodextrinas y las moléculas huésped. Otras fuerzas como las interacciones dipolo-dipolo y de Van der Waals, pueden estar implicadas también en esta unión (Cid-Samamed et al., 2022). Gracias a estas propiedades es que las ciclodextrinas han sido empleadas en la industria farmacéutica y son consideradas generalmente seguras por la FDA.

Para continuar evaluando la actividad antimicrobiana que tiene el compuesto aislado se hicieron ensayos para determinar la concentración mínima inhibitoria (MIC) en *M. luteus* y *B. subtilis*, los resultados se muestran en la Tabla 15.

Tabla 15. Resultados CMI complejos con estefimicina B.

Muestra	<i>M. luteus</i>	<i>B. subtilis</i>
SteffB:ciclodextrina β	0.08 $\mu\text{g/ml}$	20 $\mu\text{g/ml}$
SteffB:DMSO	0.63 $\mu\text{g/ml}$	30 $\mu\text{g/ml}$

A pesar de obtener mejores resultados con el uso de la ciclodextrina frente al DMSO, el compuesto no fue soluble en su totalidad y esto pudo observarse en las muestras. Por lo que deberán seguirse buscando otras alternativas o concentraciones del complejo para lograr este objetivo.

9. CONCLUSIÓN

Se realizó minería genómica de tres actinobacterias endófitas seleccionando a *Embleya* sp. NF3 por contar con los BGC de policétidos más prometedores con posible actividad anticancerígena. Se detectaron dos regiones importantes, la región 1.4 corresponde a una PKS del tipo I y con los análisis bioinformáticos se pudo determinar su posible funcionalidad al ser expresada de manera heteróloga, así como su probable actividad anticancerígena al ser evaluado en CLC-Pred.

La región 1.12 corresponde a una PKS del tipo II, los análisis bioinformáticos y antecedentes sugieren que es la región asociada probablemente a la producción de la estefimicina B. Por la actividad anticancerígena previamente reportada se exploraron dos alternativas para aumentar su producción.

La alternativa de realizar la expresión heteróloga de la región se llevó a cabo utilizando la biblioteca genómica de *Embleya* sp. NF3 y después de evaluarla se identificó una colonia que contiene las tres regiones del BGC, precedente clave para continuar con la estrategia para su expresión heteróloga.

También se optimizó el método de purificación de la estefimicina B utilizando AmberLite™ XAD™ 7HP como auxiliar para la recuperación del compuesto en medio de fermentación y se obtuvo un rendimiento de 99.3 mg/mL, superior a los rendimientos reportados previamente. La identidad del compuesto se verificó mediante LC-MS.

Por último, debido a los problemas de solubilidad la estefimicina B, se comparó la actividad antibiótica utilizando ciclodextrina β y DMSO como disolventes. Los resultados muestran una mayor actividad antibiótica gracias al uso de la ciclodextrina β , lo que permitiría seguir con los ensayos en líneas celulares para evaluar su actividad anticancerígena.

10. PERSPECTIVAS

1. Enviar a secuenciar el fósido de la colonia 8 positiva para el BGC 1.12 y continuar con la estrategia de expresión heteróloga.
2. Realizar más ensayos con la ciclodextrina β para poder lograr la completa solubilidad de la estefimicina B.
3. Diseñar la estrategia de expresión del BGC 1.4 ya que parece ser un candidato prometedor para obtener un policétido novedoso con actividad anticancerígena.

11. REFERENCIAS

- Baerson, S. R., & Rimando, A. M. (2007). A plethora of polyketides: Structures, biological activities, and enzymes. *ACS Symposium Series*, 955, 2–14. <https://doi.org/10.1021/bk-2007-0955.ch001>
- Bekiesch, P., Basitta, P., & Apel, A. K. (2016). Challenges in the heterologous production of antibiotics in *Streptomyces*. *Archiv Der Pharmazie*, 594–601. <https://doi.org/10.1002/ardp.201600058>
- Belknap, K. C., Park, C. J., Barth, B. M., & Andam, C. P. (2020). Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria. *Scientific Reports*, 10(1), 1–9. <https://doi.org/10.1038/s41598-020-58904-9>
- Bergkessel, M., & Guthrie, C. (2013). Colony PCR. *Methods in Enzymology*, 529, 299–309. <https://doi.org/10.1016/B978-0-12-418687-3.00025-2>
- Bhattarai, K., Bastola, R., & Baral, B. (2020). Antibiotic drug discovery: Challenges and perspectives in the light of emerging antibiotic resistance. In *Advances in Genetics* (1st ed., Vol. 105). Elsevier Inc. <https://doi.org/10.1016/bs.adgen.2019.12.002>
- Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., van Wezel, G. P., Medema, M. H., & Weber, T. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Research*, 49(May), 29–35. <https://doi.org/10.1093/nar/gkab335>
- Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., Medema, M. H., & Weber, T. (2019). AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research*, 47(W1), W81–W87. <https://doi.org/10.1093/nar/gkz310>
- Brodasky, T. F., & Reusser, F. (1974). Steffimycin B, a new member of steffimycin family: Isolation and characterization. *The Journal of Antibiotics*, 1(1), 810–813.
- Buchholz, T. J., Kittendorf, J. D., & Sherman, D. H. (2012). Natural products in chemical biology. In N. Civjan (Ed.), *Natural Products in Chemical Biology* (Vol.

66, pp. 163–187).

- Centeno-Leija, S., Vinuesa, P., Karol Rodríguez-Peña, Trenado-Uribe, M., Yair Cárdenas-Conejo, Serrano-Posada, H., Romina Rodríguez-Sanoja, & Sánchez, S. (2016). Draft genome sequence of an endophytic *actinoplanes* species, encoding uncommon trans-acyltransferase polyketide synthases. *Genome Announcements*, 4(2), 4–5. <https://doi.org/10.1128/genomeA.00164-16>
- Chen, A., Re, R. N., & Burkart, M. D. (2018). Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions. *Natural Product Reports*, 35(10), 1029–1045. <https://doi.org/10.1039/c8np00040a>
- Cid-Samamed, A., Rakmai, J., Mejuto, J. C., Simal-Gandara, J., & Astray, G. (2022). Cyclodextrins inclusion complex: Preparation methods, analytical techniques and food industry applications. *Food Chemistry*, 384(February), 132467. <https://doi.org/10.1016/j.foodchem.2022.132467>
- Cragg, G. M., Grothaus, P. G., & Newman, D. J. (2009). Impact of natural products on developing new anti-cancer agents †. *Chemical Reviews*, 109(7), 3012–3043. <https://doi.org/10.1021/cr900019j>
- Cuevas, X. (2005). A revision of the genus *Amphipterygium* (Julianiaceae). *Ibugana*, 13(1), 27–48.
- Dechert-Schmitt, A. M. R., Schmitt, D. C., Gao, X., Itoh, T., & Krische, M. J. (2014). Polyketide construction via hydrohydroxyalkylation and related alcohol C-H functionalizations: Reinventing the chemistry of carbonyl addition. *Natural Product Reports*, 31(4), 504–513. <https://doi.org/10.1039/c3np70076c>
- Demain, A. L., & Vaishnav, P. (2011). Natural products for cancer chemotherapy. *Microbial Biotechnology*, 4(6), 687–699. <https://doi.org/10.1111/j.1751-7915.2010.00221.x>
- El-Naggar, M. Y., & Yousry, M. G. (2001). Induction of sporulation and antibacterial activity in the aerial mycelium negative mutants of *Streptomyces nasri*. *Egyptian Journal of Biology*, 3, 23–35.
- Filimonov, D. A., Lagunin, A. A., Glorizova, T. A., Rudik, A. V., Druzhilovskii, D. S., Pogodin, P. V., & Poroikov, V. V. (2014). Prediction of the biological activity spectra of organic compounds using the pass online web resource. *Chemistry of Heterocyclic Compounds*, 50(3), 444–457. <https://doi.org/10.1007/s10593-014-1496-1>
- Gao, Y., Zhao, Y., He, X., Deng, Z., & Jiang, M. (2021). Challenges of functional expression of complex polyketide biosynthetic gene clusters. *Current Opinion in Biotechnology*, 69, 103–111. <https://doi.org/10.1016/j.copbio.2020.12.007>

- Gómez, M. P. (2021). *Caracterización in silico y química de sideróforos novedosos identificados por minería genómica aplicada a Embleya sp. NF3 y construcción de su biblioteca genómica [Tesis de maestría]*. Universidad Nacional Autónoma de México.
- Guimarães, T. C., Gomes, T. S., Fernandes, C. D., Barros, F. D., Oliveira, K. V., Bilal, M., Bharagava, R. N., Ferreira, L. F. R., & Hollanda, L. M. (2020). Antitumor microbial products by actinomycetes isolated from different environments. In P. Kumar (Ed.), *Microbial Technology for Health and Environment* (pp. 113–160). Springer India. https://doi.org/10.1007/978-981-15-2679-4_5
- Gullón, S., Olano, C., Abdelfattah, M. S., Braña, A. F., Rohr, J., Méndez, C., & Salas, J. A. (2006). Isolation, characterization, and heterologous expression of the biosynthesis gene cluster for the antitumor anthracycline steffimycin. *Applied and Environmental Microbiology*, 72(6), 4172–4183. <https://doi.org/10.1128/AEM.00734-06>
- Guzmán-Trampe, S. M., Rodríguez-Peña, K., Espinosa-Gómez, A., Sánchez-Fernández, R., Macías-Rubalcava, M. L., Flores-Cotera, L. B., & Sánchez, S. (2015). Endophytes as a potential source of new antibiotics. In S. Sanchez & A. L. Demain (Eds.), *Antibiotics: Current innovations and future trends* (First, pp. 175–204).
- Hoskisson, P. A., & van Wezel, G. P. (2019). *Streptomyces coelicolor*. *Trends in Microbiology*, 27(5), 468–469. <https://doi.org/10.1016/j.tim.2018.12.008>
- Hu, Y., Nie, Q.-Y., Pan, H.-X., & Tang, G.-L. (2020). Bacterial type II polyketide synthases. In *Comprehensive Natural Products III*. <https://doi.org/10.1016/b978-0-12-409547-2.14626-8>
- Jenke-Kodama, H., & Dittmann, E. (2009). Bioinformatic perspectives on NRPS/PKS megasynthases: Advances and challenges. *Natural Product Reports*, 26(7), 874. <https://doi.org/10.1039/b810283j>
- Karpiński, T. M., & Adamczak, A. (2018). Anticancer activity of bacterial proteins and peptides. *Pharmaceutics*, 10(2). <https://doi.org/10.3390/pharmaceutics10020054>
- Katsuyama, Y., & Ohnishi, Y. (2012). Type III polyketide synthases in microorganisms. In *Methods in Enzymology* (1st ed., Vol. 515). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-394290-6.00017-3>
- Keatinge-Clay, A. T. (2012). The structures of type I polyketide synthases. *Natural Product Reports*, 29(10), 1050–1073. <https://doi.org/10.1039/c2np20019h>
- Keatinge-Clay, A. T., Shelat, A. A., Savage, D. F., Tsai, S. C., Miercke, L. J. W., O'Connell, J. D., Khosla, C., & Stroud, R. M. (2003). Catalysis, specificity, and

- ACP docking site of *Streptomyces coelicolor* malonyl-CoA:ACP Transacylase. *Structure*, 11(2), 147–154. [https://doi.org/10.1016/S0969-2126\(03\)00004-2](https://doi.org/10.1016/S0969-2126(03)00004-2)
- Koss-Mikołajczyk, I., Todorovic, V., Sobajic, S., Mahajna, J., Gerić, M., Tur, J. A., & Bartoszek, A. (2021). Natural products counteracting cardiotoxicity during cancer chemotherapy: The special case of doxorubicin, a comprehensive review. *International Journal of Molecular Sciences*, 22(18). <https://doi.org/10.3390/ijms221810037>
- Kwan, D. H., & Leadlay, P. F. (2010). Mutagenesis of a modular polyketide synthase enoylreductase domain reveals insights into catalysis and stereospecificity. *ACS Chemical Biology*, 5(9), 829–838. <https://doi.org/10.1021/cb100175a>
- Lagunin, A. A., Dubovskaja, V. I., Rudik, A. V., Pogodin, P. V., Druzhilovskiy, D. S., Glorizova, T. A., Filimonov, D. A., Sastry, N. G., & Poroikov, V. V. (2018). CLC-Pred: A freely available web-service for in silico prediction of human cell line cytotoxicity for drug-like compounds. *PLoS ONE*, 13(1), 1–13. <https://doi.org/10.1371/journal.pone.0191838>
- Lancini, G., & Demain, A. L. (2013). Bacterial pharmaceutical products. In E. Rosenberg, E. F.- DeLong, S. Lory, E. Stackebrandt, & F. Thomson (Eds.), *The prokaryotes* (4th ed., pp. 257–280). Springer Heidelberg.
- Li, Z., Zhu, D., & Shen, Y. (2018). Discovery of novel bioactive natural products driven by genome mining. *Drug Discoveries & Therapeutics*, 12(6), 318–328. <https://doi.org/10.5582/ddt.2018.01066>
- Liu, C. Y., Li, Y. L., Lu, J. H., Qian, L. L., Xu, K., Wang, N. N., Chang, W. Q., & Lou, H. X. (2021). Steffimycin F, a new steffimycin-type derivative from the lichen-derived actinomycetes *streptomyces* sp. *Journal of Molecular Structure*, 1227. <https://doi.org/10.1016/j.molstruc.2020.129352>
- Ludwig-Müller, J. (2015). Plants and endophytes: equal partners in secondary metabolite production? *Biotechnology Letters*, 37(7), 1325–1334. <https://doi.org/10.1007/s10529-015-1814-4>
- Ma, R., & Kwok, H. F. (2020). New opportunities and challenges of venom-based and bacteria-derived molecules for anticancer targeted therapy. *Seminars in Cancer Biology, March*. <https://doi.org/10.1016/j.semcan.2020.08.010>
- Martinez-Klimova, E., Rodríguez-Peña, K., & Sánchez, S. (2017). Endophytes as sources of antibiotics. *Biochemical Pharmacology*, 134, 1–17. <https://doi.org/10.1016/j.bcp.2016.10.010>
- Martins-Teixeira, M. B., & Carvalho, I. (2020). Antitumour anthracyclines: Progress and perspectives. *ChemMedChem*, 15(11), 933–948. <https://doi.org/10.1002/cmdc.202000131>

- Medema, M. H., Breitling, R., & Takano, E. (2011). Synthetic biology in *streptomyces* bacteria. In C. Voigt (Ed.), *Synthetic Biology, Part A* (1st ed., Vol. 497, pp. 485–502). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-385075-1.00021-4>
- Méndez, C., & Salas, J. A. (2001). The role of ABC transporters in antibiotic-producing organisms: Drug secretion and resistance mechanisms. *Research in Microbiology*, 152(3–4), 341–350. [https://doi.org/10.1016/S0923-2508\(01\)01205-0](https://doi.org/10.1016/S0923-2508(01)01205-0)
- Meyer, K. J., & Nodwell, J. R. (2021). Biology and applications of co-produced, synergistic antimicrobials from environmental bacteria. *Nature Microbiology*, 6(9), 1118–1128. <https://doi.org/10.1038/s41564-021-00952-6>
- Miller, K. I., Qing, C., Sze, D. M. Y., & Neilan, B. A. (2012). Investigation of the biosynthetic potential of endophytes in traditional Chinese anticancer herbs. *PLoS ONE*, 7(5), 1–12. <https://doi.org/10.1371/journal.pone.0035953>
- Mohar-Betancourt, A., Reynoso-Noverón, N., Armas-Texta, D., Gutiérrez-Delgado, C., & Torres-Domínguez, J. A. (2017). Cancer trends in Mexico: Essential data for the creation and follow-up of public policies. *Journal of Global Oncology*, 3(6), 740–748. <https://doi.org/10.1200/JGO.2016.007476>
- Moore, R. A., & Karasek, F. W. (1984). Extraction of organic compounds from aqueous media by amberlite XAD resins. *International Journal of Environmental Analytical Chemistry*, 17(3–4), 187–202. <https://doi.org/10.1080/03067318408076972>
- Nah, H. J., Pyeon, H. R., Kang, S. H., Choi, S. S., & Kim, E. S. (2017). Cloning and heterologous expression of a large-sized natural product biosynthetic gene cluster in *Streptomyces* species. *Frontiers in Microbiology*, 8(MAR), 394. <https://doi.org/10.3389/fmicb.2017.00394>
- Newman, D. J., & Cragg, G. M. (2020). Natural Products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *Journal of Natural Products*, 83(3), 770–803. <https://doi.org/10.1021/acs.jnatprod.9b01285>
- Nivina, A., Yuet, K. P., Hsu, J., & Khosla, C. (2019). Evolution and diversity of assembly-line polyketide synthases [Review-article]. *Chemical Reviews*, 119(24), 12524–12547. <https://doi.org/10.1021/acs.chemrev.9b00525>
- Olano, C., Abdelfattah, M. S., Gullón, S., Braña, A. F., Rohr, J., Méndez, C., & Salas, J. A. (2008). Glycosylated derivatives of steffimycin: Insights into the role of the sugar moieties for the biological activity. *ChemBioChem*, 9(4), 624–633. <https://doi.org/10.1002/cbic.200700610>
- Olano, C., Méndez, C., & Salas, J. A. (2009). Antitumor compounds from marine actinomycetes. *Marine Drugs*, 7(2), 210–248. <https://doi.org/10.3390/md7020210>

- Pashaei, Y., Mehrabi, M., & Shekarchi, M. (2020). A review on various analytical methods for determination of anthracyclines and their metabolites as anti-cancer chemotherapy drugs in different matrices over the last four decades. *TrAC - Trends in Analytical Chemistry*, 130, 115991. <https://doi.org/10.1016/j.trac.2020.115991>
- Praveen, V., Bihari, V., Srivastava, S. C., Division, F. T., & Manzil, C. (2008). Production of Actinomycin-D by the mutant of a new isolate of *Streptomyces sindenensis*. *Brazilian Journal of Microbiology*, 689–692.
- Quiñonez-Bastidas, G. N., & Navarrete, A. (2021). Mexican plants and derivatives compounds as alternative for inflammatory and neuropathic pain treatment—a review. *Plants*, 10(5). <https://doi.org/10.3390/plants10050865>
- Rath, C. M., Scaglione, J. B., Kittendorf, J. D., & Sherman, D. H. (2010). NRPS/PKS hybrid enzymes and their natural products. *Comprehensive Natural Products II: Chemistry and Biology*, 1, 453–492. <https://doi.org/10.1016/b978-008045382-8.00725-5>
- Ridley, C. P., & Khosla, C. (2009). Polyketides. In M. Schaechter (Ed.), *Encyclopedia of Microbiology* (pp. 472–481). Elsevier. <https://doi.org/10.1016/B978-012373944-5.00158-9>
- Rodríguez-Peña, K. (2020). *Búsqueda de metabolitos secundarios de interés farmacológico en microorganismos endófitos de Amphipterygium adstringens [Tesis de doctorado]*. Universidad Nacional Autónoma de México.
- Rodríguez-Peña, K., Gómez-Román, M. P., Macías-Rubalcava, M. L., Rocha-Zavaleta, L., Rodríguez-Sanoja, R., & Sánchez, S. (2022). Bioinformatic comparison of three *Embleya* species and description of steffimycins production by *Embleya* sp. NF3. *Applied Microbiology and Biotechnology*, 106(8). <https://doi.org/10.1007/s00253-022-11915-0>
- Rodríguez-Peña, K., Macías-Rubalcava, M. L., Rocha-Zavaleta, L., Trenado-Uribe, M., Rodríguez-Sanoja, R., & Sanchez, S. (2018). *Streptomyces scabrisporus*, an endophyte isolated from *Amphipterygium adstringens* as producer of an anthracycline active against different cancer cell lines. *Global Drugs and Therapeutics*, 3(5), 1–10. <https://doi.org/10.15761/gdt.1000158>
- Rohr, J., & Hertweck, C. (2010). Type II PKS. *Comprehensive Natural Products II: Chemistry and Biology*, 1, 227–303. <https://doi.org/10.1016/b978-008045382-8.00703-6>
- Romero-Rodríguez, A., Robledo-Casados, I., & Sánchez, S. (2015). An overview on transcriptional regulators in *Streptomyces*. In *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* (Vol. 1849, Issue 8, pp. 1017–1039). Elsevier B.V. <https://doi.org/10.1016/j.bbagr.2015.06.007>

- Rückert, C., Albersmeier, A., Busche, T., Jaenicke, S., Winkler, A., Fridjónsson, Ó. H., Hreggvidsson, G. Ó., Lambert, C., Badcock, D., Bernaerts, K., Anne, J., Economou, A., & Kalinowski, J. (2015). Complete genome sequence of *Streptomyces lividans* TK24. *Journal of Biotechnology*, 199, 21–22. <https://doi.org/10.1016/j.jbiotec.2015.02.004>
- Sanchez, S., & Demain, A. L. (2019). Secondary metabolites. In M. Moo-Young (Ed.), *Comprehensive Biotechnology* (2nd ed., Vol. 1, Issue 1). Elsevier. <https://doi.org/10.1016/B978-0-444-64046-8.00012-4>
- Selim, M. S. M., Abdelhamid, S. A., & Mohamed, S. S. (2021). Secondary metabolites and biodiversity of actinomycetes. *Journal of Genetic Engineering and Biotechnology*, 19(1). <https://doi.org/10.1186/s43141-021-00156-9>
- Severi, E., & Thomas, G. H. (2019). Antibiotic export: Transporters involved in the final step of natural product production. *Microbiology (United Kingdom)*, 165(8), 805–818. <https://doi.org/10.1099/mic.0.000794>
- Sharifi-Rad, J., Ozleyen, A., Tumer, T. B., Adetunji, C. O., Omari, N. El, Balahbib, A., Taheri, Y., Bouyahya, A., Martorell, M., Martins, N., & Cho, W. C. (2019). Natural products and synthetic analogs as a source of antitumor drugs. In *Biomolecules* (Vol. 9, Issue 11). <https://doi.org/10.3390/biom9110679>
- Shimizu, Y., Ogata, H., & Goto, S. (2016). Type III polyketide synthases: Functional classification and phylogenomics. *ChemBioChem*, 18(1), 50–65. <https://doi.org/https://doi.org/10.1002/cbic.201600522>
- Singh, A., Chaudhary, S., Shankar, A., & Prasad, V. (2018). Polyketide and its derivatives. In *New and Future Developments in Microbial Biotechnology and Bioengineering: Microbial Genes Biochemistry and Applications*. Elsevier B.V. <https://doi.org/10.1016/B978-0-444-63503-7.00013-9>
- Smith, S., & Tsai, S. C. (2007). The type I fatty acid and polyketide synthases: A tale of two megasynthases. *Natural Product Reports*, 24(5), 1041–1072. <https://doi.org/10.1039/b603600g>
- Sousa, M. F. V. Q., Lopes, C. E., & Pereira, N. (2002). Development of a bioprocess for the production of actinomycin-D. *Brazilian Journal of Chemical Engineering*, 19(3), 277–285. <https://doi.org/10.1590/S0104-66322002000300002>
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 71(3), 209–249. <https://doi.org/10.3322/caac.21660>
- Trenado-Uribe, M. (2016). *Identificación de metabolitos secundarios antibacterianos producidos por actinomicetos endófitos de la planta Amphipterygium adstringens [Tesis de maestría]*. Universidad Nacional Autónoma de México.

- Trenado-Urbe, M., Silva-Miranda, M., Rivero-Cruz, J. F., Rodríguez-Peña, K., Espitia-Pinzón, C. I., Rodríguez-Sanoja, R., & Sánchez, S. (2018). Antimycobacterial activity of an anthracycline produced by an endophyte isolated from *Amphipterygium adstringens*. *Molecular Biology Reports*, *45*(6), 2563–2570. <https://doi.org/10.1007/s11033-018-4424-0>
- Van der Zanden, S. Y., Qiao, X., & Neefjes, J. (2021). New insights into the activities and toxicities of the old anticancer drug doxorubicin. *FEBS Journal*, *288*(21), 6095–6111. <https://doi.org/10.1111/febs.15583>
- Vázquez-Martínez, M. (2017). *Minería genómica de una cepa de Streptomyces sp. aislada del cuachalalate: Alternativa para la búsqueda de nuevos compuestos con actividad farmacéutica. [Tesis de maestría]*. Universidad Nacional Autónoma de México.
- Wang, J., Zhang, R., Chen, X., Sun, X., Yan, Y., Shen, X., & Yuan, Q. (2020). Biosynthesis of aromatic polyketides in microorganisms using type II polyketide synthases. *Microbial Cell Factories*, *19*(1), 1–11. <https://doi.org/10.1186/s12934-020-01367-4>
- Wang, X., Tian, X., Wu, Y., Shen, X., Yang, S., & Chen, S. (2018). Enhanced doxorubicin production by *Streptomyces peucetius* using a combination of classical strain mutation and medium optimization. *Preparative Biochemistry and Biotechnology*, *48*(6), 514–521. <https://doi.org/10.1080/10826068.2018.1466156>
- Waterhouse, A. M., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., De Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. <https://doi.org/10.1093/nar/gky427>
- Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>
- Weber, T., Blin, K., & Shaw, S. (2019). *antiSMASH 5 USER MANUAL antiSMASH 5 antiSMASH database MANUAL*.
- Wiegand, I., Hilpert, K., & Hancock, R. E. W. (2008). Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nature Protocols*, *3*(2), 163–175. <https://doi.org/10.1038/nprot.2007.521>
- Wolpert, M., Heide, L., Kammerer, B., & Gust, B. (2008). Assembly and heterologous expression of the coumermycin A1 gene cluster and production of new derivatives by genetic engineering. *ChemBioChem*, *9*(4), 603–612.

<https://doi.org/10.1002/cbic.200700483>

- Xu, W., Qiao, K., & Tang, Y. (2013). Structural analysis of protein-protein interactions in type I polyketide synthases. *Critical Reviews in Biochemistry and Molecular Biology*, 48(2), 98–122. <https://doi.org/10.3109/10409238.2012.745476>
- Yan, X., Ge, H., Huang, T., Hindra, Yang, D., Teng, Q., Crnovčić, I., Li, X., Rudolf, J. D., Lohman, J. R., Gansemans, Y., Zhu, X., Huang, Y., Zhao, L. X., Jiang, Y., van Nieuwerburgh, F., Rader, C., Duan, Y., & Shen, B. (2016). Strain prioritization and genome mining for enediyne natural products. *MBio*, 7(6), 1–12. <https://doi.org/10.1128/mBio.02104-16>
- Zierep, P. F., Ceci, A. T., Dobrusin, I., Rockwell-Kollmann, S. C., & Günther, S. (2020). Sempi 2.0—a web server for pks and nrps predictions combined with metabolite screening in natural product databases. *Metabolites*, 11(1), 1–27. <https://doi.org/10.3390/metabo11010013>
- Zierep, P. F., Padilla, N., Yonchev, D. G., Telukunta, K. K., Klementz, D., & Günther, S. (2017). SeMPI: A genome-based secondary metabolite prediction and identification web server. *Nucleic Acids Research*, 45(W1), W64–W71. <https://doi.org/10.1093/nar/gkx289>

12. ANEXOS

12.1 Anexo 1. Medios de cultivo y soluciones

Tabla 16. Medio Terrific-Broth.

Solución	Reactivo	Cantidad
A) 900 mL	Triptona	12 g
	Extracto de levadura	24 g
	Glicerol	4 mL
B) 100 mL	KH ₂ PO ₄	23.1 g/L
	K ₂ HPO ₄	125.4 g/L

Tabla 17. Medio LB.

Reactivo	Cantidad
Extracto de levadura	5.0 g/L
NaCl	10.0 g/L
Triptona de caseína	10.0 g/L

Tabla 18. Medio YMG líquido suplementado con CaCO₃.

Reactivo	Cantidad
Glucosa	4 g/L
Extracto de levadura	4 g/L
Extracto de malta	10 g/L
CaCO ₃	2 g/L

Tabla 19. Medio YMG líquido.

Reactivo	Cantidad
Glucosa	4 g/L
Extracto de levadura	4 g/L
Extracto de malta	10 g/L

Tabla 20. Solución I.

Reactivo	Concentración
Glucosa	50 mM
Tris-HCl	25 mM
EDTA	10 Mm
Lisozima	4 mg/mL

Tabla 21. Solución II.

Reactivo	Concentración
NaOH	0.2 N
SDS	1%

Tabla 22. Solución III.

Reactivo	Cantidad
Acetato de potasio 5M	60 mL
Ácido acético glacial	11.5 mL
Agua	28.5 mL

Tabla 23. Buffer TE (Tris-EDTA)

Reactivo	Concentración
Tris-HCl	10 mM
EDTA	1 mM
	pH: 8.0

Tabla 24. Solución 0.5 de McFarland.

Para 100 mL:

Reactivo	Volumen
Ácido sulfúrico 0.18 M	99.5 mL
Cloruro de Bario 0.048 M	0.5 mL

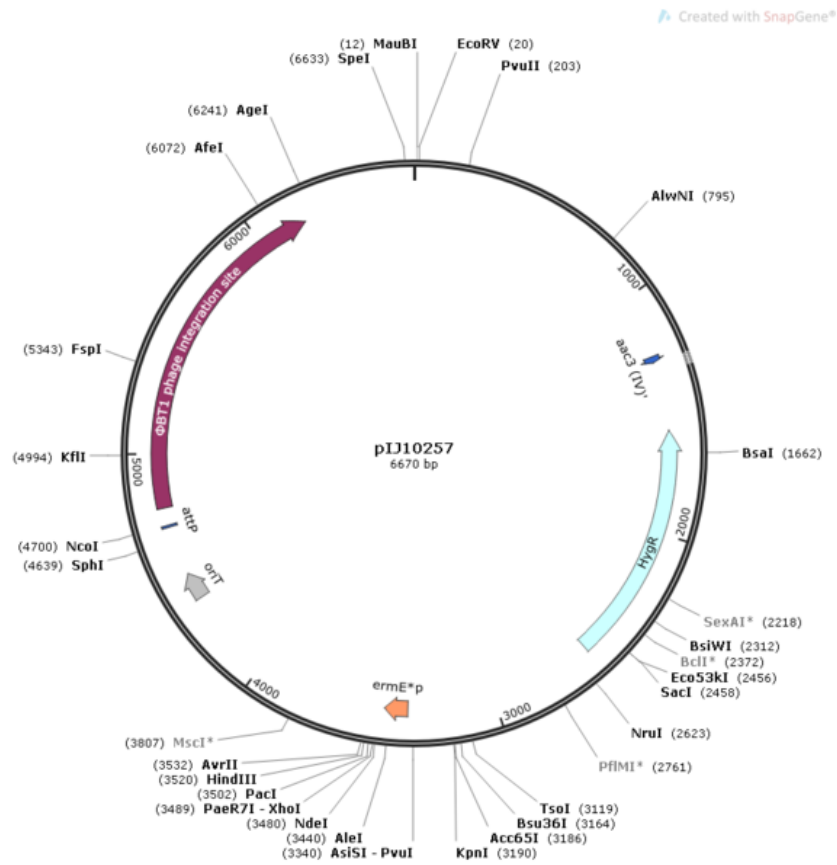
12.2 Anexo 2. Secuencia de primers.

Tabla 25. Secuencias de primers diseñados para el escaneo de la biblioteca genómica.

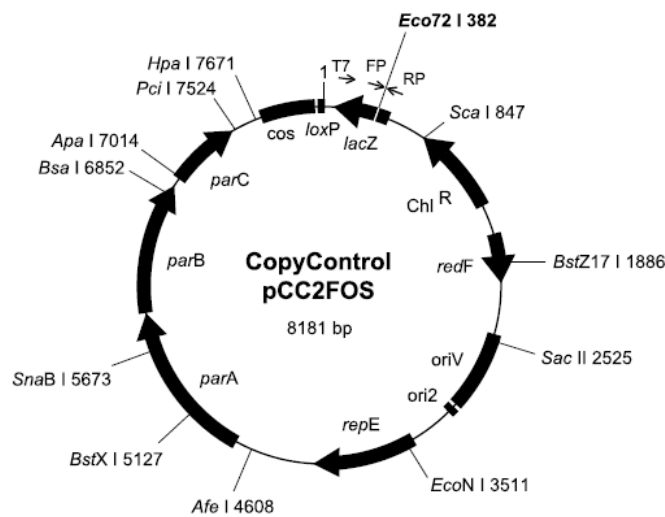
Nombre	Secuencia	Tamaño primer	Tamaño amplicón
StefInicioF	5' GGTGCTCAGCGCGAACAGCATGTCTCGAACCTGC 3'	32 pb	1492 pb
StefInicioR	5' GTCACCGGAGGCACGATCGAAGACCACATCGC 3'	32 pb	
StefMedioF	5'GACTGAGCGAGCTGATCACCGCGGCCGACC 3'	30 pb	2760 pb
StefMedioR	5' GGATCACCGCCTTCGACCCGCAACGCTATTCC 3'	32 pb	
StefFinalF	5'GCATCGCCTACACGCCTGCCCCGAACGACATC 3'	32 pb	1691 pb
StefFinalR	5'GGTGAACAGGTCCGAAGCGAGGCTGTCTGTGAC 3'	32 pb	

12.3 Anexo 3. Vectores utilizados

1. Vector de expresión pIJ10257



2. Fósmido pCC2FOS de la biblioteca genómica



12.4 Anexo 4. Protocolo para transformación de células electrocompetentes.

1. Las células se descongelan en hielo y se agitan cuidadosamente
2. A una alícuota de células competentes de 50 μL se añaden 2 μL de ADN.
3. Transferir la suspensión de células y ADN a una cubeta de electroporación fría.
4. Electroporar 5 ms a 1250 V.
5. Remover la cubeta de la cámara y añadir inmediatamente 950 μL de medio LB frío y glucosa 100 mM.
6. Resuspender las células y transferirlas a un tubo de 1.5 mL a 37°C por 1 hora.
7. Plaquear las células en medio LB con el antibiótico. Distribuir bien en la placa e incubar a 37°C.

12.5 Anexo 5. Aislamiento de plásmido de bajo número de copias.

1. Haga un cultivo de 250 mL con medio LB y el antibiótico adecuado. Incube toda la noche a 37°C y 200 rpm
2. Centrifugue el cultivo 15 minutos a 6,000 rpm. Deseche el sobrenadante cuidando que no queden restos del medio de cultivo.
3. Adicione por cada volumen de 250 mL, 30 mL de solución I (5 mM de sacarosa, 10 mM de EDTA, 25 mM de Tris pH 8) y deje incubar 5 minutos a temperatura ambiente.
4. Resuspenda las células, si es necesario utilice vortex. Añada 125 μL de Rnasa I.
5. En hielo, adicione 60 mL de solución II (0.2 N de NaOH, 1% de SDS) seguido de 50 mL de solución III (Acetato de sodio 3M pH 4.8). Incube por 15 minutos.
6. Centrifugue a 10,000 rpm por 10 minutos y precipite el sobrenadante con 1 volumen de isopropanol x 30 minutos.
7. Resuspenda cada botón en 5 mL de buffer TE.

12.6 Anexo 6. Minipreparación de plásmido de *E. coli* de la biblioteca genómica de *Embleya* sp. NF3.

1. Inocular 15 mL de medio Terrific-Broth que contenga cloranfenicol a una concentración de 12.5 mg/mL con una colonia de la biblioteca genómica de *Embleya* sp. NF3.
2. Cultivar toda la noche a 37°C con 200 rpm.
3. Tomar 12 mL en un tubo de Eppendorf y centrifugar 10 minutos a 6,000 rpm a 4°C.
4. Descartar el sobrenadante cuidando que no queden restos del medio de cultivo. Adicionar 100 mL de solución I por cada mililitro de medio extraído, agitar suavemente. Incubar 5 minutos a temperatura ambiente.
5. Añadir 100 mL/ mL de medio de solución II y mezclar por inversión el tubo. Incubar 5 minutos en hielo.
6. Adicionar 150 mL/mL de medio de solución III preenfriada y mezclar. Incubar 10 minutos en hielo.
7. Centrifugar 15 minutos a 12,000 rpm, traspasar el sobrenadante y añadirle 1 volumen de isopropanol. Incubar 40 minutos a temperatura ambiente.
8. Centrifugar 10 min a 10,000 rpm. Tirar el sobrenadante y agregar 2-2.5 volumen de etanol al 70% y volver a centrifugar a 12,000 rpm por 12 minutos.
9. Tirar el sobrenadante y dejar secar el botón a temperatura ambiente.

12.7 Anexo 7. Protocolo para limpieza de ARN

1. Adicionar Rnasa a una concentración final de 20 µg/mL considerando el volumen de medio extraído. Dejar reposar por una hora a 37° C.
2. Adicionar un volumen de fenol: cloroformo y centrifugar a 10,000-12,000 rpm por 5 minutos.
3. Recuperar la fase acuosa (superior) y adicionar un volumen de cloroformo. Centrifugar a 12,000 rpm por 5 minutos.
4. Repetir el paso 3.
5. Adicionar NaCl a concentración de 100 mM y 2.1 volúmenes de etanol absoluto, mezclar por inversión.
6. Dejar a -20°C por al menos 40 minutos.
7. Centrifugar a 15,000 rpm por 15 minutos y tirar el sobrenadante.
8. Adicionar 1 mL de etanol al 70% y centrifugar 10 minutos a 12,000 rpm.
9. Repetir el paso 8.
10. Dejar evaporar el etanol a temperatura ambiente.
11. Resuspender en agua miliQ o buffer TE.

12.8 Anexo 8. Protocolo PCR de colonia

1. Se seleccionan las colonias aisladas a evaluar.
2. En tubos de 1.5 mL se adicionan 50 mL de agua miliQ esterilizada. Con palillos estériles se pica la colonia y se enjuaga la punta del palillo.
3. Los tubos se hierven por 5 minutos.
4. Para la mezcla de reacción se toma como base una reacción de 50 mL de PCR.
5. Una vez que se tiene la mezcla, se distribuyen 9 mL de dicha mezcla en tubos de PCR y se adiciona a cada tubo 1 mL de ADN de colonia.
6. Finalmente se corre la PCR con las condiciones previamente establecidas.

Tabla 26. Lista de reactivos y concentración para la PCR de colonia.

Reactivo	Volumen	Concentración final
ADN de colonia	10 (1 mL c/u)	~50-100 ng
dNTP's 10 mM	2 mL	0.2 mM
Primer forward [10 pM/ μL]	2 mL	0.4 mM
Primer reverse [10 pM/ μL]	2 mL	0.4 mM
DMSO	5 mL	5%
MgCl ₂	3 mL	1.5 mM
Buffer de PCR 10X, - Mg	10 mL	5
Platinum™ Taq ADN Polymerase	1 mL	0.05 U
Agua miliQ	65 mL	A volumen final