

SCREENING DE GENES BIOSINTETICOS DE POLICÉTIDOS (PKS) EN UNA  
BIBLIOTECA METAGENOMICA PROVENIENTE DE SUELO ATLANTICO  
BRASILEIRO

ROGER DAVID CASTILLO ARTEAGA

UNIVERSIDAD DE NARIÑO  
FACULTAD DE CIENCIAS EXACTAS Y NATURALES  
PROGRAMA DE BIOLOGÍA  
SAN JUAN DE PASTO

2013

SCREENING DE GENES BIOSINTETICOS DE POLICÉTIDOS (PKS) EN UNA  
BIBLIOTECA METAGENOMICA PROVENIENTE DE SUELO ATLANTICO  
BRASILERO

ROGER DAVID CASTILLO ARTEAGA

Trabajo de grado presentado como requisito final de la asignatura trabajo de grado II

Asesor Externo:

GABRIEL PADILLA MALDONADO PhD.

Doctor en Genética de Microorganismos

*Instituto de Ciências Biomédicas*

*Universidade de São Paulo, Brasil*

Asesora interna:

EDITH MARIELA BURBANO ROSERO PhD.

Doctora en Ciencias (Microbiología)

Departamento de Biología

Universidad de Nariño, Colombia.

UNIVERSIDAD DE NARIÑO

FACULTAD DE CIENCIAS EXACTAS Y NATURALES

PROGRAMA DE BIOLOGÍA

SAN JUAN DE PASTO

2013

## NOTA DE RESPONSABILIDAD

“Las ideas y conclusiones aportadas en el trabajo de grado, son responsabilidades exclusivas del autor”

Artículo 1° del acuerdo No 324 de Octubre 11 de 1966, emanado del Honorable Consejo Directivo de la Universidad de Nariño.

Nota de aceptación:

---

---

---

---

---

---

---

---

Gabriel Padilla Maldonado PhD.  
Firma del director externo de trabajo de grado

---

Edith Mariela Burbano Rosero PhD.  
Firma del director interno del trabajo de grado

---

Carol Rosero Galindo PhD.  
Firma del jurado

---

Bianca Caetano de Almeida PhD.  
Firma del jurado

San Juan de Pasto, Octubre de 2013



**ACUERDO No.161**  
(25 de Noviembre)

Por la cual se asigna a un Trabajo de Grado la distinción de **LAUREADO**.

**EL CONSEJO DE LA FACULTAD DE CIENCIAS EXACTAS Y NATURALES DE LA UNIVERSIDAD DE NARIÑO**, en uso de sus atribuciones reglamentarias y estatutarias, y,

**CONSIDERANDO:**

Que el día martes 19 de Noviembre de 2013, a las 11:00 am en el Auditorio de la Facultad de Ciencias Agrícolas, el estudiante **ROGER DAVID CASTILLO ARTEAGA**, sustentó el proyecto de grado titulado **"SCREENING DE GENES BIOSINTETICOS DE POLICETIDOS (PKS) EN UNA BIBLIOTECA METAGENOMICA PROVENIENTE DEL SUELO ATLANTICO BRASILEIRO"**;

Que las Doctoras Carol Yovanna Rosero Galindo, Bianca Caetano de Almeida, jurados evaluadores de dicho trabajo, concedieron una calificación de 60 puntos en el trabajo escrito y 40 puntos en la sustentación oral;

Que los Doctores Edith Mariela Burbano y Gabriel Padilla Maldonado como asesores y jurados del trabajo de grado del estudiante conceden una calificación de 40 puntos en el trabajo escrito, 30 puntos en la sustentación oral y 30 puntos en la evaluación de competencias en investigación;

Que una vez promediadas las calificaciones del trabajo escrito, sustentación y competencias en investigación emitidas por el jurado incluyendo los asesores que desempeñaron como jurado, se obtiene un puntaje final de 100 puntos, lo cual corresponde cualitativamente a Trabajo de Grado con mención **LAUREADO**;

Que los Doctores Carol Yovanna Rosero Galindo, Bianca Caetano de Almeida, Edith Mariela Burbano y Gabriel Padilla Maldonado han emitido los conceptos mediante los cuales se argumenta la calificación de Trabajo de Grado con mención **LAUREADO** otorgado al estudiante **ROGER DAVID CASTILLO**;

Que el trabajo de grado del estudiante **ROGER DAVID CASTILLO** puede ser de mucha utilidad para la sociedad académica;

Que el trabajo desarrollado presenta aspectos de gran interés no solo para el público del área académica sino también para empresarios de la rama biotecnológica y farmacéutica;

Que el estudiante en mención investigó genes que pueden ser muy importantes en la producción de antimicrobianos, nuevos medicamentos y medicamentos que pueden ser utilizados en el combate de dolencias;

Que el estudiante hizo uso de técnicas en este estudio que fueron adecuadas a los objetivos del trabajo, toda vez que son interesantes en investigaciones que involucran muestras ambientales

Que se observó durante la exposición la capacidad de síntesis y argumentación soportada;

Que el estudiante respondió a todas las preguntas de forma correcta, sucinta y concreta;

Que se evidenció durante la sustentación que el estudiante se esmeró en su presentación, se notó pulcritud, organización coherente y secuencial y que el manejo del tiempo fue como lo establecido;

Que el estudiante asumió en forma ética, responsable e idónea las dificultades presentadas en el proceso de investigación;

Que argumentó de manera coherente y relevante los avances científicos que implicó el desarrollo de su trabajo;

Que en consecuencia de lo anterior el Comité Curricular y de Investigaciones del Departamento de Biología; mediante Proposición No.087 de 25 de noviembre del presente año, recomendó ante el Consejo de Facultad de Ciencias Exactas y Naturales, Otorgar la distinción de **LAUREADO** al trabajo de grado (Modalidad trabajo de Investigación) titulado "**SCREENING DE GENES BIOSINTETICOS DE POLICETIDOS (PKS) EN UNA BIBLIOTECA METAGENOMICA PROVENIENTE DEL SUELO ATLANTICO BRASILEIRO**"; presentado por el estudiante **ROGER DAVID CASTILLO ARTEAGA**;

Que este organismo consideró viable la petición.

**ACUERDA:**

ARTICULO 1º Otorgar la distinción de **LAUREADO** al trabajo de grado (Modalidad trabajo de Investigación) titulado "**SCREENING DE GENES BIOSINTETICOS DE POLICETIDOS (PKS) EN UNA BIBLIOTECA METAGENOMICA PROVENIENTE DEL SUELO ATLANTICO BRASILEIRO**"; presentado por el estudiante **ROGER DAVID CASTILLO ARTEAGA**; según la parte motiva de la presente Acuerdo;

COMUNÍQUESE Y CUMPLASE.

San Juan de Pasto, 25 de noviembre de 2013.

  
PABLO FERNÁNDEZ IZQUIERDO  
Presidente  
Proyectó: Duvi C.

  
DUVI CASTILLO MENESES  
Secretaria

## AGRADECIMIENTOS

A mis padres Adalberto Castillo y Aura Arteaga, a mi hermanita Anabel Castillo Arteaga por el esfuerzo tan grande a lo largo de su vida, por regalarme su ejemplo, valores, dedicación, a mirar hacia adelante, porque siempre han estado conmigo.

A la Profesora Edith Mariela Burbano, por su amistad y apoyo incondicional durante estos años, por todos sus consejos que permitieron mejorar mi formación académica y personal.

Al Profesor Dr. Gabriel Padilla Maldonado, por las innumerables orientaciones, recomendaciones, por su amistad y por brindarme un lugar en el laboratorio de bioproductos, por impulsarme a continuar en este proceso de formación.

A la Profesora. Dra. Carol Rosero y a la Dra. Bianca Caetano por sus observaciones y aportes fundamentales en este trabajo.

Al Profesor Dr. Pablo Fernández y amigos del grupo de investigación en Biotecnología Microbiana: Vanessa Pérez, Iván Otero, Maira Quiroz, Cristian Cabrera, Guido Villota, Duvi Castillo, donde comencé los primeros pasos en la investigación.

A Carlos, Raquel, Lizeth, Ricardo, Felipe, Camilo, y a todos mis amigos de biología, por su amistad durante estos años, por ser un grupo muy colaborador, unido, por sus consejos.

A mis compañeros y amigos del laboratorio de Bioproductos, por brindarme un acogimiento muy especial: Simone Ichiwaki, María Paula Parada, Fernanda Nogales, Carla Montanari, Ruth Ortiz, Felipe Monsalve, Camilo Contreras, Alejandra Ferreira, Renata Furlan, Leandro Garrido.

A mis amigos en São Paulo, Jennifer, Nadia, Catalina, Marcela, Lina, Juan C, Lizeth, Maira, María Paula y todas las personas con las que compartí muy buenos momentos en Brasil.

Al personal de la sección de Genética de microorganismos por la compañía en los experimentos, las charlas en los “lanches” y su amistad.

Al Instituto de *Ciências Biomédicas da Universidade de São Paulo*, por la asistencia durante el desarrollo del trabajo de conclusión de pregrado, las actividades realizadas bajo sus instalaciones y por brindarme un espacio para la investigación.

Al programa de Biología de la Universidad de Nariño y a cada una de las personas que hacen parte de él, por tener la oportunidad de crecer profesionalmente en todas las áreas de la biología.

## DEDICATORIA

A mis Padres, mi hermana y toda mi familia por ser una parte esencial en mi vida.

A todas las personas que contribuyeron positivamente, y compartieron conmigo buenos momentos a lo largo de estos años.

## TABLA DE CONTENIDO

	Pagina
1. INTRODUCCION	18
3 PLANTEAMIENTO DEL PROBLEMA	19
3. JUSTIFICACIÓN	21
4. OBJETIVOS	23
4.1 Objetivo general	23
4.2 Objetivos específicos	23
5. ANTECEDENTES	24
6. MARCO TEORICO	26
6.1 Comunidades bacterianas	26
6.1.1 Comunidades bacterianas en suelos	27
6.2 Metagenómica	28
6.3 Policétidos	42
6.3.1 Tipos de PKS	45
6.3.1.1 PKS tipo I	45
6.3.1.2 PKS II	46
7. MATERIALES Y MÉTODOS	48
7.1 Recuperación y análisis de viabilidad de la biblioteca metagenómica	48
7.2 Extracción de cromosoma artificial bacteriano pBAC OriV (~8kb)	48
7.3 <i>Screening</i> primario	49
7.4 Amplificación para detección de genes del sistema PKS I y II	50
7.5 Reacción de secuenciación	54
7.6 Análisis bioinformático	54
8. RESULTADOS Y DISCUSIÓN	56
8.1 Extracción de cromosoma artificial bacteriano pBAC OriV (~8kb)	56

8.2	<i>Screening</i> primario y amplificación para detección de genes del sistema PKS I y II	59
8.2.1	PKS I	59
8.2.2	PKS II	64
8.3	Modelamiento molecular y sitios de unión	72
	CONCLUSIONES	76
	RECOMENDACIONES	77
	REFERENCIAS BIBLIOGRAFICAS	78
	ANEXOS	85

## LISTA DE TABLAS

	Pagina.
Tabla 1. Principales fuentes y vectores para la inserción de fragmentos según su tamaño.	33
Tabla 2. Secuencias de los primers degenerados que se utilizaron para la detección de policétidos PKS de tipo I.	51
Tabla 3. Secuencias de los primers degenerados que fueron utilizados para la detección de policétidos PKS de tipo II.	51
Tabla 4. Protocolo inicial para estandarización de los genes PKS tipo I en DNA plasmidial.	52
Tabla 5. Protocolo inicial para la detección PKS tipo II en DNA total de suelo.	53
Tabla 6. Microorganismos usados en este estudio como controles positivos para amplificación de genes biosintéticos PKS I y II.	54
Tabla 7. Protocolo estandarizado para la detección de PKS tipo I en Bibliotecas de insertos grandes.	61
Tabla 8. Protocolo estandarizado para la detección de PKS tipo II en Bibliotecas de insertos grandes.	65

## LISTA DE FIGURAS

		Pagina
Figura 1.	Formación de microhábitats en suelos. Fuente: Adaptado de Chenu & Stotzky, 2002.	28
Figura 2.	Proceso de Fragmentación de DNA, inserción del fragmento en el vector apropiado (DNA ligasa). Fuente: Cultek, 2006.	29
Figura 3.	Ejemplo de un plásmido pUC18. Este plásmido posee un gran número de sitios de restricción en el sitio de clonación múltiple, localizado dentro del gen <i>lacZ</i> . Las bacterias que contienen pUC18 producen colonias de color azul si crecen en un medio que contenga X-gal. El DNA insertado en el sitio de clonación múltiple interrumpe el gen <i>lacZ</i> , resultando en colonias blancas, lo que permite la identificación directa de las colonias que tienen insertos de DNA clonados Fuente: (Cultek, 2006).	31
Figura 4.	Etapas para la clonación usando vector BAC. 1) Restricción del vector con enzimas; 2) Ligación del inserto en el vector usando DNA ligasa; 3) Trasformación del vector 4) Selección de colonias con BAC recombinante. Fuente: Adaptado de Cultek, 2006.	32
Figura 5.	Ensamblaje de secuencias en bibliotecas genómicas. El solapamiento de secuencias se denomina <i>contigs</i> y se organizan de tal manera que la información proporciona el origen y orden de la secuencia original Fuente: (Green, 2001).	35
Figura 6.	Flujograma mostrando las etapas para la construcción de bibliotecas metagenómicas a partir de muestras de suelo. Fuente: Adaptado de Daniel, 2005.	36
Figura 7.	Técnicas para el análisis de bibliotecas metagenómicas. Fuente: Adaptado de Schloss & Handelsman, 2003.	37
Figura 8.	Fotografía de nanoDESI con una colonia microbiana sobre la superficie de agar sólido en caja Petri Fuente: (Watrous <i>et al.</i> , 2012).	39
Figura 9.	<i>Screening</i> por hibridación de ácido nucleico. 1) La muestra con las colonias que componen la biblioteca se transfiere a una membrana 2) las bacterias son lisadas y el DNA desnaturizado, 3) la membrana se coloca en una bolsa con la sonda marcada, luego la membrana es retirada y lavada 4) los híbridos son detectados por auto radiografía de rayos X 5)	41

las colonias que contienen el gen de interés hibridan y son seleccionadas Fuente: Adaptado de (Allison, 2009).

- Figura 10. Ejemplos de policétidos como metabolitos secundarios indicando la actividad biológica. Fuente: (Hertweck, 2009). 43
- Figura 11. Mecanismo básico involucrado en la síntesis de a) ácidos grasos b) policétidos. (Hertweck, 2009). 44
- Figura 12. Clúster génico codificante de las sintasas que forman la eritromicina (PKS I) Fuente: (Staunton & Weissman, 2001). 46
- Figura 13. Ruta biosintética de la producción de la molécula actinorodina en *Streptomyces coelicolor*, indicando la catálisis de policétidos sintasas Fuente: (Keatinge-Clay *et al.*, 2004). 47
- Figura 14. Ejemplo de biosíntesis de doxorubicina (PKS II) Fuente: (Hertweck, 2009). 47
- Figura 15. Esquema de extracción de DNA plasmidial de cada una de las placas de polipropileno. Adaptado y modificado de Galindo, 2011. 50
- Figura 16. Extracción de DNA plasmidial con diversos kits comerciales. 1. Marcador I *Hind*III (*Promega*); 2-5. Clon 12A. Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciences*) teñido en solución de 5µg/mL de bromuro de etidio. 57
- Figura 17. Extracción clásica de DNA plasmidial (modificado). 1. Marcador 1Kb (*Thermo scientific*); 2-6. Clon 2A3F, clon 2A4E, clon 2A4F, Pool 2A, Pool 14B. Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciences*) teñido en solución de 5µg/mL de bromuro de etidio. 58
- Figura 18. Mapa de restricción vector pBAC oriV (~8kb) (Wild *et al.*, 2002) 59
- Figura 19. Restricción clones 1-4 2A1A 2A3F, 2A4E, 2A4F 5. Marcador I *Hind*III (*Promega*). Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciences*) teñido en solución de 5µg/mL de bromuro de etidio. 59
- Figura 20. Amplificación PKS I *Streptomyces olindensis*. 1. Marcador 1Kb plus (*Life technologies*); 2-7. Gradiente Temperatura 58° a 66°. Gel agarosa 1% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciences*) teñido 62

en solución de 5µg/mL de bromuro de etídio.

- Figura 21. Amplificación PKS II *Streptomyces coelicolor*. 1. Marcador 1Kb (*Thermo scientific*); 2-7. Gradiente Temperatura 56° a 66°. Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciencies*) teñido en solución de 5µg/mL de bromuro de etídio. 66
- Figura 22. Amplificación PKS II . 1. Marcador 1Kb (*Thermo scientific*); 2. Pool 2A; 3. Pool 3C; 4. Pool 14A; 5. Pool 14B; 6. Control negativo *E. coli* 7. Control Positivo *S. olindensis*. Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciencies*) teñido en solución de 5µg/mL de bromuro de etídio. 67
- Figura 23. Amplificación PKS II clon 2A4F. 1. Marcador 1Kb (*Thermo scientific*); 2-7. Gradiente Temperatura 56° a 64°C. Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciencies*) teñido en solución de 5µg/mL de bromuro de etídio. 68
- Figura 24. Árbol de las secuencias genicas proveniente de los clones, control positivo (*S. olindensis*) comparacion con secuencias de las bases de datos. La escala indica el grado de similaridad. El árbol construido usando el metodo de Neighbor joining con el programa MEGA 5. 70
- Figura 25. Composición de aminoácidos de la enzima control *S. olindensis* usando la base de datos predict protein (Yadav *et al.*, 2003) 71
- Figura 26. Composición de aminoácidos de la enzima de origen metagenómico usando la base de datos predict protein (Yadav *et al.*, 2003) 71
- Figura 27. Sitios de unión en la proteína control cetosintasa usando la base de datos predict protein (Yadav *et al.*, 2003). 73
- Figura 28. Sitios de unión en la proteína de origen metagenómico usando la base de datos predict protein (Yadav *et al.*, 2003). 73
- Figura 29. Recombinación homologa de fragmentos de origen diferente 74  
Fuente: Madigan *et al.*, 2011.

## LISTA DE ANEXOS

	Pagina
Anexo A. Protocolos de extracción de DNA usados por Massini 2009, para la construcción de la librería.	85
Anexo B. Protocolo de miniprep de DNA plasmídico usando el método modificado de lisis alcalina.	88
Anexo C. Secuencias ( <i>Query</i> ) origen metagenómico y Cetosintase, usando la base de datos predict protein (Yadav <i>et al.</i> , 2003)	91
Anexo D. Código genético	95

## RESUMEN

Actualmente, la metagenómica ha permitido conocer y estudiar con más detalle el material genómico desconocido en todos los ecosistemas que componen la tierra. El suelo se ha convertido en una fuente potencial de estudio, puesto que es un sistema complejo que alberga una gran cantidad y diversidad de microorganismos. El objetivo del presente trabajo fue realizar un *screening* de genes relacionados con la biosíntesis de policétidos (PKS), con posible aplicación biotecnológica a partir de una biblioteca metagenómica del 2004 de suelo Atlántico Brasileiro, construida con 5000 clones utilizando como vector a *Escherichia coli* DH10B pBAC/oriV e insertos de 40 y 50 Kb. La caracterización de los sistemas implicados en la biosíntesis de estas moléculas, sería una alternativa para dilucidar más ampliamente el potencial biotecnológico inmerso en las comunidades microbianas, así como entender más profundamente su papel ecológico. La primera etapa de este estudio fue la recuperación de los clones en medio LBAC. Posteriormente, se experimentaron varios protocolos para la extracción del vector: GeneJET Plasmid Miniprep (Thermo Scientific), AxyPrep Plasmid Miniprep (Axygen Biosciences), PureYield™ Plasmid Miniprep System (Promega) y QIAprep Spin Miniprep (Qiagen). No obstante, no se logró obtener buenos resultados, por ello fue necesario realizar un protocolo clásico modificado, con el cual se obtuvo buena calidad y cantidad de DNA. La amplificación de los genes asociados a los sistemas PKS I y PKS II fue realizada utilizando primers degenerados. Los clones positivos para los sistemas PKS fueron aislados y sus secuencias nucleotídicas y proteicas fueron analizadas por análisis bioinformático. El *screening* permitió la obtención de tres clones positivos para genes relacionados con el sistema PKS II. Se observaron variaciones en la secuencia de origen metagenómico y la secuencia del control positivo cuando fueron comparadas de acuerdo a su conformación molecular proteica, esto permitió dilucidar la posible relación existente en la conformación de rutas de biosíntesis de otros metabolitos secundarios.

## ABSTRACT

Actually, the metagenomic has allowed to know and to study in detail the genomic material present in all environments on earth. The soil has become a potential source of study, since it is a complex system that has a large number and variety of microorganisms. The aim in these study was perform a *screening* of the biosynthesis PKS I and PKS II related genes, from a metagenomic library constructed in 2004 from Brazilian Atlantic soil with 5000 clones, using *Escherichia coli* DH10B pBAC/oriV as a vector and inserts between 40 and 50 Kb. The characterization of the biosynthesis related systems of these molecules could be a great choice to elucidate the biotechnological potential in microbial communities and their ecological roles. The first step of this study was to recover the clones in LBAC medium. Later, different kits were performed for the vector extraction: GeneJET Plasmid Miniprep (*Thermo Scientific*), AxyPrep Plasmid Miniprep (*Axygen Biosciences*), PureYield™ Plasmid Miniprep System (*Promega*) and QIAprep Spin Miniprep (*Qiagen*). However, the results were not satisfactory, for that it was necessary to perform and modify the classic extraction, in order to obtain good quality of DNA. The amplification of the related genes whit PKS I and PKS II systems was performed using degenerated primers. The positive clones for PKS systems were isolated, and their nucleotide sequences were analyzed through bioinformatics tools. The *screening* allowed obtaining three positive clones for genes related with PKS II. Furthermore, variations were observed in the sequences of metagenomic origin when they were compared with the positive control, according to protein molecular conformation. This results allowed elucidate the possible relationship in the formation of other biosynthetic pathways secondary metabolites.

## GLOSARIO

ACP [Acyl carrier protein] - proteína transportadora del grupo acil

AT [Acyltransferase] -acil transferasa

BAC [Bacterial artificial chromosome] -cromosoma artificial bacteriano

DH [Dehydratase] -deshidratasa

ER [Enoyl reductase] -enoil reductasa

FAS [Fatty acid synthase] -sintasa de ácidos grasos

MT [Methyltransferase] -metiltransferasa

NCBI *National Center of Biotechnology Information*

KS [Ketosynthase]- cetosintasa

PKS [Polyketide synthase] -policetido sintasa

TE [Thioesterase] –tioesterasa

NRPS [Nonribosomal peptide synthetases]- Sintetasas de péptidos no ribosomales.

## 1. INTRODUCCION

El aislamiento de metabolitos secundarios producidos por microorganismos, históricamente ha sido efectivo en la identificación de nuevos compuestos con propiedades bioactivas. Los suelos son considerados ambientes que contienen una diversidad considerable de microorganismos, de los cuales solamente el 1% es cultivable en laboratorio (Hugenholtz & Goebel, 1998; Curtis & Sloan, 2005). Una pequeña parte de los microorganismos que son cultivados en laboratorio, suministran una importante cantidad de moléculas que en la actualidad son de uso comercial, entre los que se destacan: antibióticos, antifúngicos, inmunosupresores, y antitumorales. Sin embargo una de las fuentes promisorias para la búsqueda de este tipo de compuestos, se encuentra en los microorganismos que no son cultivables por las técnicas tradicionales de cultivo (Gillespie *et al*, 2002). Una de las alternativas para el estudio de este tipo de microorganismos, es la metagenómica, que ofrece el análisis de la totalidad de genomas contenidos en una muestra ambiental.

En este estudio se realizó un *screening* molecular de genes pertenecientes a los policétidos tipo I y II, que son moléculas que dependiendo de su conformación estructural, las propiedades bioactivas pueden variar (Hopwood, 2004; Hertweck, 2004); a partir de una biblioteca metagenómica, construida con vector pBAC oriV, con insertos de DNA total próximos a los 50kb provenientes de suelo atlántico Brasileño.

## 2. PLANTEAMIENTO DEL PROBLEMA

Desde hace mucho tiempo, los microorganismos han sido explorados con base en cultivos y aislamientos bacterianos en laboratorio, sin embargo una pequeña parte de estos es conocida; una de las principales causas es la dificultad de crecimiento con los métodos tradicionales, lo que ha generado un vacío sobre la diversidad microbiana real (Torsvik *et al.*, 2002).

Uno de los hábitats con mayor diversidad microbiana es el suelo (Ludwig, & Schleifer, 1995; Schloss & Handelsman, 2006), a pesar de su poco conocimiento, actualmente existe una nueva alternativa viable para el estudio de estas comunidades, denominada metagenómica. Esta rama de la genómica actual, le ha dado un nuevo sentido a la problemática de analizar y conocer la totalidad de genomas contenidos en una muestra ambiental, además de ser una herramienta útil para rastrear genes de interés, y resolver incógnitas fisiológicas o ecológicas de las comunidades microbianas (Bonilla *et al.*, 2008).

La metagenómica permite un estudio extenso de diversos genes, entre ellos los de interés biotecnológico. Estos genes están ampliamente relacionados con el metabolismo, y codifican para una amplia gama enzimática que representa un gran potencial en el campo biotecnológico. La mayoría de metabolitos con potencial bioactivo, son: antibióticos, antifúngicos, inmunosupresores, antitumorales, inhibidores de enzimas, toxinas, pigmentos, entre otros (Demain, 1992; Massini, 2009).

Entre los metabolitos secundarios con mayor interés en el ámbito biotecnológico se encuentran los que presentan actividad farmacológica; estos metabolitos son una prioridad en el mundo actual, debido al incremento de microorganismos patógenos multiresistentes, causada por el uso irracional de medicamentos (antibacterianos y antifúngicos). El objetivo principal de la biomedicina actual es la búsqueda de nuevos medicamentos como moléculas bioactivas para el tratamiento de enfermedades complejas como el cáncer y enfermedades cardiovasculares. Esto ha impulsado el desarrollo de nuevas alternativas para la producción de nuevos fármacos (Waugh & Long, 2002; Strobel & Daisy, 2003).

Una de las alternativas para la producción de nuevos fármacos se incluye la búsqueda de un grupo de moléculas denominadas “policétidos”, que se caracterizan por tener propiedad bioactiva y pertenecen a una familia de metabolitos estructuralmente muy diversa con un amplio espectro de actividad; además comparten un mismo patrón biosintético por medio de enzimas denominadas policétido sintetasas (Hopwood, 2004; Massini, 2009).

La búsqueda de este tipo de moléculas usando el abordaje de la genómica actual, es una herramienta útil que puede ser aplicada al *screening* de una biblioteca construida en 2004 con DNA total de suelo proveniente del bosque atlántico brasilero, con insertos de aproximadamente 50 Kb, la pertinencia del análisis se basa en que hasta ahora no se han rastreado genes codificantes de este tipo de moléculas en la mencionada construcción.

En atención a los anteriores argumentos, el presente trabajo, plantea realizar el *screening* molecular de una biblioteca metagenómica para conocer genes codificantes de policetidos, que representan un gran potencial de información que puede ser utilizada con fines académicos, ecológicos, biotecnológicos y económicos.

Con este trabajo se plantea resolver la siguiente pregunta de investigación:

¿Qué genes relacionados con la biosíntesis de policétidos (PKS) están presentes en una biblioteca metagenómica de suelo Atlántico Brasileiro?

### 3.JUSTIFICACIÓN

En Brasil, los bosques del Atlántico (*Mata Atlântica*) son considerados refugios de la biosfera, y presentan una diversidad bastante alta, por lo cual esta zona fue declarada por la UNESCO como un patrimonio de la Humanidad (<http://www.pmma.etc.br/>). En ecosistemas tan importantes como estos, los cuerpos terrestres juegan un papel importante en la formación de microhábitats, estos están influenciados por factores ambientales que ejercen presión sobre los microorganismos permitiendo su diversificación (Chenu & Stotzky, 2002).

Los microhábitats terrestres pueden ser comprendidos mejor con las actuales metodologías independientes de cultivo. Una de las principales razones para estudiar los microhábitats es que se ha comprobado que la diversidad de microorganismos en el suelo es de aproximadamente  $10^9$  células por gramo de muestra, de acuerdo con estudios de observación directa y recuento de células al microscopio (FISH-Microscopia óptica) (Amann *et al.*, 1995; Schloss & Handelsman, 2006), representando entre 4.600 a 10.000 genomas distintos (Torsvik *et al.*, 1990), muchas de las cuales, >90%, corresponden a bacterias viables pero no cultivables (Schloss & Handelsman, 2006).

La mayoría de fármacos de importancia como los antibióticos del grupo de las tetraciclinas, eritromicinas, vancomicinas, cefalosporinas, rifamicinas y diferentes  $\beta$  lactámicos, provienen de microorganismos de suelo; motivo por el cual es un ambiente promisorio para estudios de bioprospección (Osborne *et al.*, 2002; Lorenz *et al.*, 2002). Entre las moléculas bioactivas que presentan mayor variabilidad estructural y química se encuentran los policétidos, que dependiendo de los dominios enzimáticos, las propiedades del compuesto son diferentes (Khosla & Zawada, 1996; Yadav, Gokhale, & Mohanty, 2003).

Partiendo que el conjunto de genes de una comunidad microbiana se puede estudiar por medio de la metagenómica. La construcción de una biblioteca involucra una serie de procedimientos complejos, que requieren optimización en la extracción de DNA total, clonación, manejo y almacenamiento de un número bastante alto de clones.

Por ello se plantea realizar el *screening* de una construcción metagenómica del año 2004, que aún no ha sido analizada para este tipo de genes, la aplicación de herramientas de “*screening*” molecular en la biblioteca construida, puede generar una mejor alternativa en la búsqueda de nuevos genes involucrados en la biosíntesis de compuestos con posible uso farmacéutico.

## 4.OBJETIVOS

### 3.1 Objetivo general

Realizar un *screening* de genes relacionados con la biosíntesis de policétidos (PKS), con posible aplicación biotecnológica, en una biblioteca metagenómica de suelo Atlántico Brasileiro.

### 3.2 Objetivos específicos

1. Recuperar y evaluar una biblioteca metagenómica de suelo Brasileiro construida con *Escherichia coli* DH10B, el vector pBAC/oriV e insertos entre 40 y 50 Kb.
2. Realizar *screening* por PCR de genes PKS I y II en una biblioteca metagenómica.
3. Seleccionar y secuenciar los clones positivos producto del *screening* de genes PKS I y II.
4. Analizar las secuencias génicas obtenidas con genes reportados en bases de datos.

## 5. ANTECEDENTES

La producción y prospección de nuevos fármacos se ha estudiado ampliamente en microorganismos filamentosos como los actinomicetos. Los estudios basados en caracterización química de metabolitos secundarios, es un proceso clásico que aún es ampliamente utilizado en la actualidad. Entre las especies más estudiadas involucradas en la producción de metabolitos secundarios, se encuentra la producción de estreptomina, espectinomicina y tetraciclina por: *Streptomyces griseus*, y *S. aureofaciens*. La modificación genética en *Streptomyces* también se ha estudiado con el fin de obtener variantes que produzcan nuevos metabolitos. Entre las nuevas moléculas que se han obtenido con el uso de cepas recombinantes se encuentran: polienos, nistatina, amfotericina y pimaricina (Mariano, 2009).

En la década de los 80 una gran parte de herramientas moleculares y tecnologías basadas en secuenciación, fueron desarrolladas con el fin de superar las limitaciones de los cultivos en laboratorio (Borneman & Triplett, 1997; Torsvik *et al.*, 1998). El uso de estas herramientas les permitieron a los investigadores clonar fragmentos grandes de DNA con la finalidad de rastrear genes en bibliotecas metagenómicas (Borneman & Triplett, 1997; Gillespie *et al.*, 2002; Lefevre *et al.*, 2008).

Con respecto a los genes que hacen parte de sistemas PKS, los primeros estudios comenzaron con la construcción de primers homólogos a regiones consenso presentes en los genes KS $\alpha$  y KS $\beta$  (codificantes de unidades involucradas en el factor de elongación de la cadena), lo que amplió la posibilidad de analizar variantes de estos genes en nuevas muestras (Hutchinson & Fujii, 1995). Los estudios de amplificación de genes involucrados en las vías del grupo de los políctidos comenzaron en 1999, cuando Metsä-Ketela y colaboradores construyeron primers degenerados homólogos al gen KS $\alpha$  para analizar microorganismos productores de políctidos (Metsä-Ketelä *et al.*, 1999).

Esto conllevó a realizar análisis en muestras ambientales como las procesadas en el año 2003 por Courtois y colaboradores, usando primers degenerados homólogos a PKS de tipo I como sondas moleculares para analizar 5000 clones de una biblioteca metagenómica, de las cuales se obtuvieron 11 nuevas secuencias codificantes para sintasas de policétidos de tipo I (Courtois *et al.*, 2003).

En biomas Brasileños como la Mata Atlántica se han realizado estudios de metagenómica; Andrielli (2006) amplificó genes codificantes de sintasas de tipo II, en una muestra de DNA total en suelo, sugiriendo la existencia de microorganismos productores de esta clase de productos representados, principalmente por agentes antitumorales en esa comunidad bacteriana.

Massini (2004), construyó una biblioteca metagenómica proveniente de suelo de la *Mata Atlántica*. Las técnicas de extracción de DNA para este tipo de suelo fueron estandarizadas y optimizadas, obteniendo un DNA total de buena calidad, que fue digerido con la enzima *Hind* III, obteniendo fragmentos de aproximadamente 50 Kb, estos fragmentos posteriormente fueron clonados en un vector pBAC/oriV (Wild *et al.*, 2002).

El presente trabajo tiene como objetivo el rastreo de clones positivos que estén asociados con el sistema PKS I y II, o con rutas biosintéticas de metabolitos secundarios como policétidos.

## 6.MARCO TEORICO

### 6.1 Comunidades bacterianas

A pesar de la importancia de los microorganismos procariontes (arqueas y bacterias), ésta solo fue reconocida a partir de la década de los sesenta, cuando Kluver y van Niel, (1956) estimaron que 95% del CO<sub>2</sub> disponible en la atmósfera era producido por la mineralización bacteriana. Así mismo, las bacterias son responsables del cambio atmosférico más importante (hasta ahora) en el planeta Tierra: la producción de oxígeno diatómico y con ello la conversión de una atmósfera reductora a una oxidante. Hoy en día, solo los organismos procariontes son capaces de realizar la fijación de nitrógeno atmosférico, mineralizar nitrógeno hasta nitrato, respirar sulfatos o hierro, por mencionar algunos de los múltiples procesos que involucran los ciclos biogeoquímicos (Bonilla *et al.*, 2008).

A lo largo de la historia, la microbiología clásica ha tratado de comprender mejor los procesos fisiológicos y bioquímicos que son llevados a cabo por los microorganismos en los diferentes ambientes. Actualmente, los estudios de biología molecular han ampliado el conocimiento sobre metabolismo y diversidad en estos ambientes, lo que permitió reconocer que el dominio con mayor biomasa, mayor diversidad taxonómica y funcionalidad es del dominio Bacteria (Bonilla *et al.*, 2008).

El estudio de los microorganismos, dado su pequeño tamaño y sus tasas generacionales muy variables, (generalmente muy rápidas), implica una problemática de estudio. Actualmente, se acepta que tan solo se conoce el 1% de la diversidad total de bacterias en el planeta y que en un gramo de suelo existen, al menos, miles de especies (Torsvik *et al.*, 2002; Schloss & Handelsman, 2006). Dado que las escalas espaciales y temporales de los organismos procariontes son una dificultad de estudio, también representan una oportunidad única de entender los fenómenos ecológicos y metabólicos.

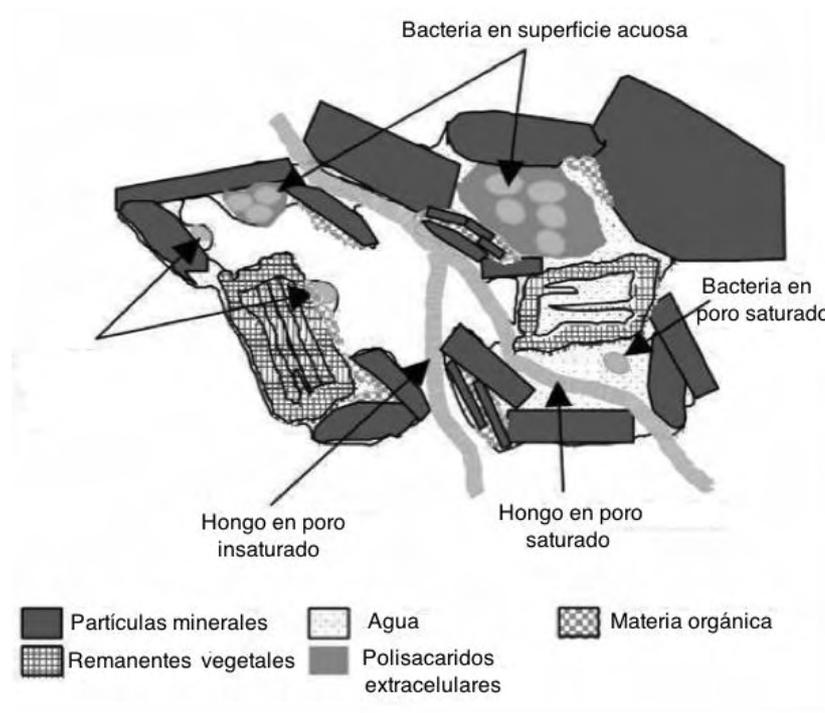
La revolución de la genómica actual, debido al desarrollo de tecnologías de secuenciación, le ha dado un nuevo sentido al análisis de comunidades bacterianas, porque ahora es posible conocer el total de los genes contenidos en una muestra ambiental para resolver incógnitas ecológicas y/o metabólicas de los organismos (Bonilla *et al.*, 2008).

#### 6.1.1 Comunidades bacterianas en suelos

El suelo está estructurado por la combinación de agregados, que se forman dependiendo de factores bióticos como: liberación de compuestos de origen vegetal, actividad microbiana y animal, factores abióticos que incluyen reacciones fisicoquímicas (formación de compuestos, interacciones iónicas, precipitación, aireación, movimiento, retención de agua, erosión, reciclado de nutrientes, y penetración por raíces ( Tisdall & Oades, 1982; Perfect & Grove, 2002; Bronick & Lal, 2005).

Los agregados de partículas sólidas resultan en un complejo discontinuo, formando espacios de varios tamaños y formas, muchos de ellos llenos de agua o aire, dando como resultado diferentes hábitats para los microorganismos (Figura 1) (Chenu & Stotzky 2002).

Figura 1. Formación de microhábitats en suelos.



Fuente: Adaptado de Chenu & Stotzky, 2002.

La proporción de procariontes en la tierra fue estimada en  $4,6 \times 10^{30}$  células y entre los ambientes considerados como los mayores reservorios de diversidad microbiana en el planeta, se encuentran los suelos (Robe *et al.*, 2003).

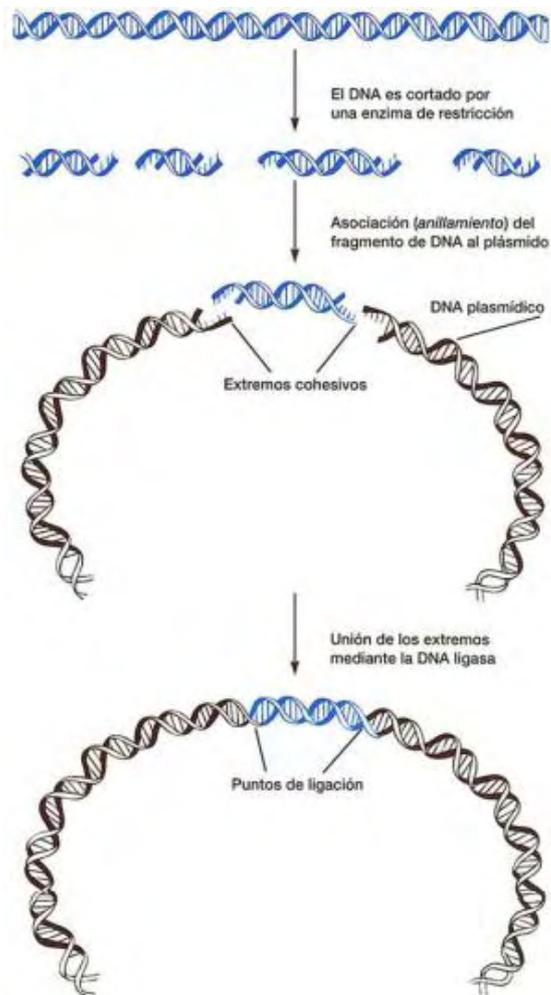
## 6.2 Metagenómica

La metagenómica representa una aproximación totalmente nueva al estudio de las comunidades microbianas, y se define como la aplicación de técnicas genómicas modernas para el estudio directo de comunidades de microorganismos en su entorno natural, evitando la necesidad de aislar y cultivar cada una de las especies que componen la comunidad (Handelsman *et al.*, 1998; Chen & Pachter, 2005). Las bibliotecas metagenómicas se construyen de la misma forma que una biblioteca genómica y se dividen en tres etapas:

1. Fragmentación de DNA total por medio de enzimas de restricción o de forma mecánica:

La metagenómica permite eliminar la etapa de cultivo, porque comienza con la extracción directa de DNA de una muestra que incluye la fragmentación del DNA total para la inserción en un vector apropiado (Figura 2) (Plásmidos, Cósmidos, Fosmidos, BAC, entre otros) (Handelsman *et al*, 1998).

Figura 2. Proceso de Fragmentación de DNA, inserción del fragmento en el vector apropiado (DNA ligasa).



Fuente: Cultek, 2006.

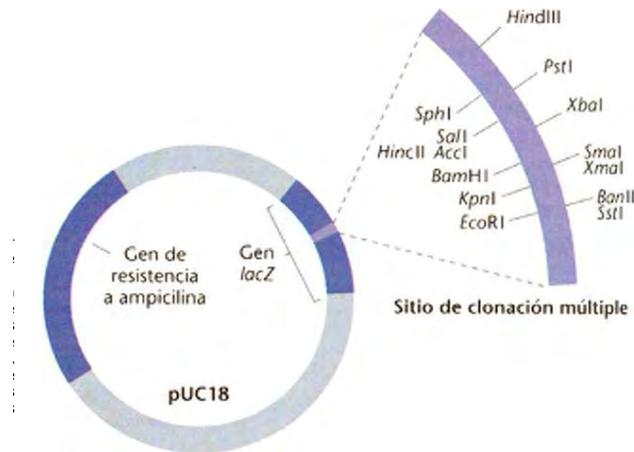
## 2. Inserción de los fragmentos de DNA en un vector adecuado

Un factor importante al momento de la construcción de una biblioteca metagenómica es el tamaño del inserto que va a ser clonado. Bibliotecas menores a 10 Kb pueden ser construidas con plásmidos y generalmente se usan para la identificación de unos cuantos genes (Handelsman, 2004).

Los plásmidos están compuestos por moléculas de DNA de doble cadena extracromosómicas que tienen un origen de replicación (*ori+*) y que se replican autónomamente en las células bacterianas. Para poder utilizarlos en diferentes estudios, se han modificado o diseñado muchos plásmidos de manera que contengan un número limitado de sitios de restricción y genes de resistencia a antibióticos específicos (Cultek, 2006).

Los plásmidos actuales contienen un fragmento de un gen bacteriano denominado *lacZ* (Figura 3), el gen posee un sitio de clonación múltiple (*polylinker*). El gen *lacZ* normal, codifica la  $\beta$ -galactosidasa, enzima que corta moléculas de azúcar. Cuando se introduce en una célula hospedera que tiene el gen *lacZ* mutante, se produce  $\beta$ -galactosidasa no funcional. La  $\beta$ -galactosidasa puede ser separada en dos péptidos: LacZ $\alpha$  and LacZ $\Omega$ . Estos péptidos no son activos individualmente pero si están en presencia del otro se activan espontáneamente. Esta propiedad es usada porque si se insertan fragmentos de DNA en el vector, el gen se ve interrumpido y las células no van a mostrar actividad de  $\beta$  -galactosidasa. Esta actividad puede ser detectada por X-gal que tiñe de azul cuando es cortado por la  $\beta$  -galactosidasa, por lo tanto si hubo una inserción de DNA al vector, la tinción se ve blanca. Las colonias blancas, resistentes a ampicilina, contienen plásmidos con el DNA de interés (Cultek, 2006).

Figura 3. Ejemplo de un plásmido pUC18. Este plásmido posee un gran número de sitios de restricción en el sitio de clonación múltiple, localizado dentro del gen *lacZ*. Las bacterias que contienen pUC18 producen colonias de color azul si crecen en un medio que contenga X-gal. El DNA insertado en el sitio de clonación múltiple interrumpe el gen *lacZ*, resultando en colonias blancas, lo que permite la identificación directa de las colonias que tienen insertos de DNA clonados.



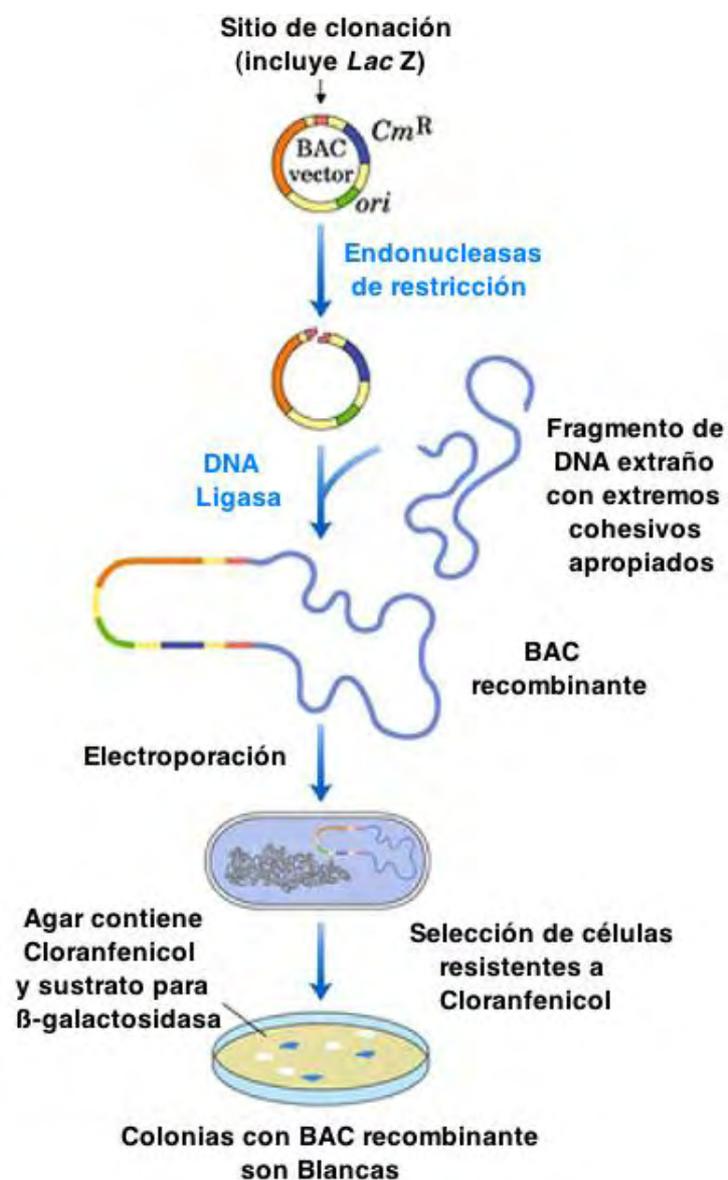
Fuente: (Cultek, 2006).

Las bibliotecas que poseen insertos superiores a 10Kb son ideales para la identificación de vías metabólicas compuestas por varios genes (Ginolhac *et al.*, 2004) y las bibliotecas que contienen insertos grandes de DNA superiores a 20 Kb son clonadas en cósmidos, fosmidos, cromosomas artificiales de bacterias (BAC) o levaduras (YAC) (Tabla 1) y son ideales para la búsqueda de clúster de genes que incluyan vías metabólicas completas (Handelsman, 2004).

BAC (cromosoma artificial bacteriano): estos vectores están basados en el factor F de bacterias. El factor F es un plásmido que se replica independientemente y está implicado en la transferencia de información genética durante la conjugación bacteriana. Puesto que los factores F pueden transportar fragmentos del cromosoma bacteriano de hasta 1Mb, han sido diseñados para que funcionen como vectores de DNA de tamaños grandes. Los vectores BAC tienen los genes de replicación y de

número de copias del factor F, incorporan un marcador de resistencia a un antibiótico y sitios de restricción para insertar el DNA exógeno que se desea clonar (Figura 4). Además, el sitio de clonación está flanqueado por regiones promotoras que pueden utilizarse para generar moléculas de RNA y expresar genes clonados, pueden ser utilizadas como sondas de mapeo cromosómico, o para secuenciar el inserto de DNA (Allison, 2009).

Figura 4. Etapas para la clonación usando vector BAC. 1) Restricción del vector con enzimas; 2) Ligación del inserto en el vector usando DNA ligasa; 3) Trasformación del vector 4) Selección de colonias con BAC recombinante.



Fuente: Adaptado de Cultek, 2006.

Tabla 1. Principales fuentes y vectores para la inserción de fragmentos según su tamaño

VECTOR	FUENTE	TAMAÑO
PLÁSMIDO	Plásmidos replicación múltiple (origen natural y modificados)	≤ 10 Kb
FAGO	Bacteriofago λ	5-20 Kb
COSMIDO	Plásmido + sitio Cos del bacteriofago λ	30-45 Kb
BAC (CROMOSOMA ARTIFICIAL BACTERIANO)	<i>Escherichia coli</i> (modificado)	50-300 Kb
YAC (CROMOSOMA ARTIFICIAL DE LEVADURA)	<i>Saccharomyces cerevisiae</i> +centromero+ telomero+secuencia de replicación autónoma	100-1000 Kb

Fuente: (Allison, 2009).

### 3. Transformación del vector en un hospedero adecuado

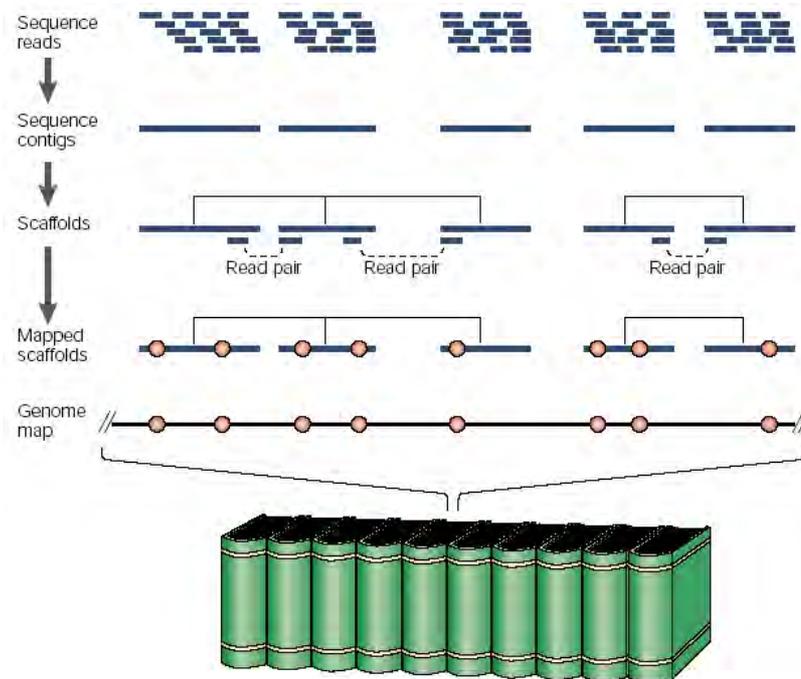
El DNA recombinante que incluye la fusión del inserto con el vector es introducido en un hospedero cultivable (generalmente *E. coli*) para la construcción de la biblioteca (Handelsman, 2004).

Cuando el DNA recombinante es aceptado exitosamente, la biblioteca entra en la última fase de construcción. Los clones detectados como “positivos” se seleccionan y almacenan generalmente en placas de Elisa de 96 pozos, para la posterior selección (*screening*) (Bonilla *et al.*, 2008).

Las tecnologías actuales de secuenciación permiten el análisis de una mayor cantidad de información en menor tiempo, además la disminución de costos, brinda la posibilidad de estudiar un conjunto bastante grande de genes, generalmente denominado “*cluster*”. Esta agrupación es típica de rutas biosintéticas involucradas en la producción de metabolitos secundarios (Bonilla *et al.*, 2008).

Después del proceso de secuenciación, se debe realizar el análisis y ensamblaje de las secuencias. El ensamblaje de secuencias se basa en la sobreposición de diferentes fragmentos “*contigs*” de DNA y así reconstruir secuencias lineales completas conocidas como “*scaffolds*” y la secuencia lineal completa del genoma (Figura 5). Muchas veces los fragmentos secuenciados no cubren la totalidad del genoma (existen *contigs* que no pueden ensamblarse con ningún otro fragmento porque las regiones que los unían no fueron secuenciadas). Obviamente, si los fragmentos producidos por la secuenciación son más grandes es más fácil ensamblar que si los fragmentos son más pequeños. Sin embargo, para saber con precisión si una secuencia es en realidad la secuencia del genoma y no un artefacto de la secuenciación, es necesario que cada base en cada sitio a lo largo de la secuencia lineal se encuentre repetida varias veces, es decir, que cada base en cada fragmento haya sido secuenciada varias veces para que al alinear todos los diferentes fragmentos secuenciados de una misma región sean coherentes entre sí (Green, 2001). A esto se le conoce como profundidad de la cobertura de secuenciación, un genoma completo se considera confiable si tiene más de 6x. Esto quiere decir, que la totalidad de bases secuenciadas es 6 veces el tamaño del genoma inferido, y por lo tanto, se esperaría que cada base del genoma completo estuviera representada al menos en seis fragmentos independientes. Cabe señalar que el esfuerzo computacional que requiere este análisis es considerable y, aunque hemos observado enormes avances, el ensamblaje automático dista mucho de ser perfecto y frecuentemente requiere la asistencia humana (Green, 2001).

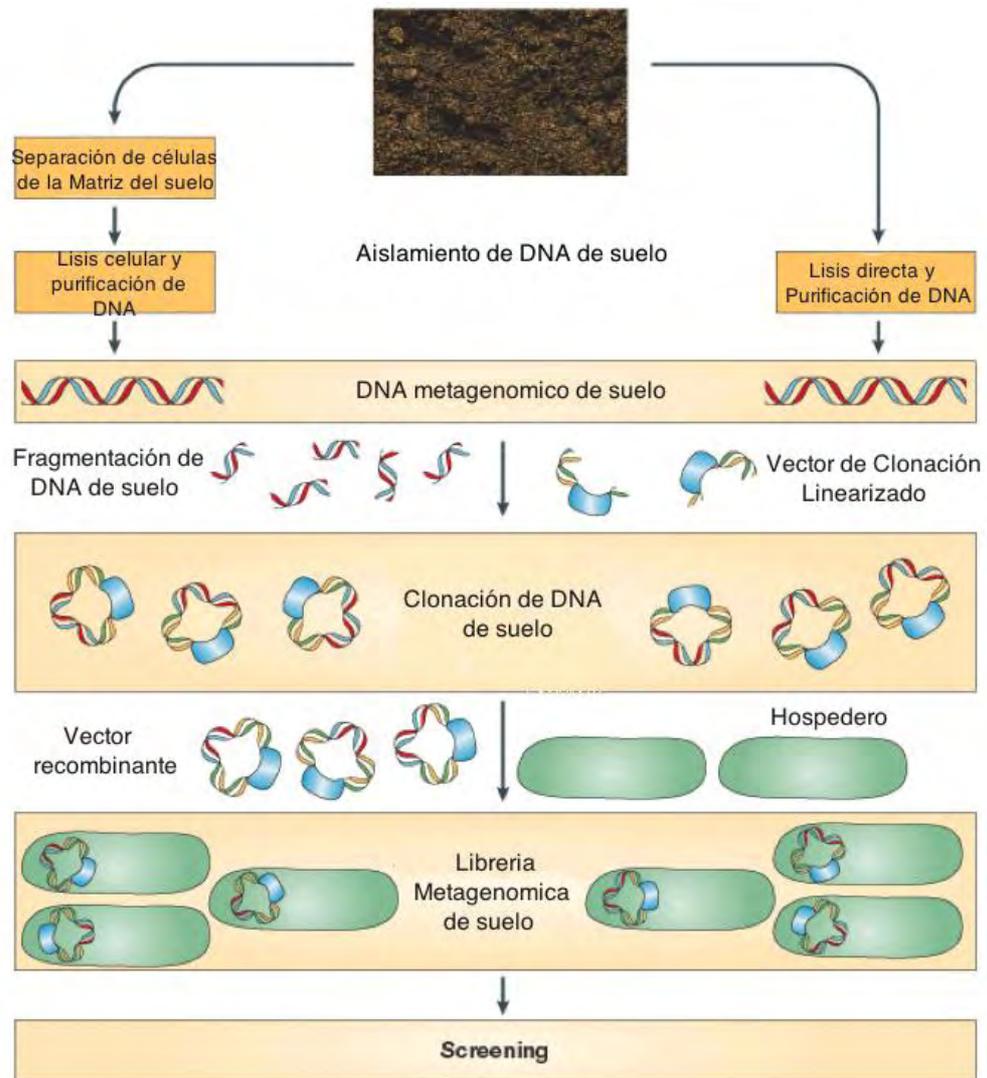
Figura 5. Ensamblaje de secuencias en bibliotecas genómicas. El solapamiento de secuencias se denomina *contigs* y se organizan de tal manera que la información proporciona el origen y orden de la secuencia original



Fuente: Green, 2001.

Con el estudio de la metagenómica se puede acceder a una mayor cantidad de genes, porque además de ser secuenciados, estos se pueden aislar y transformar en vectores de expresión eficientes para formar los nuevos compuestos con mayor facilidad (Lorenz *et al.*, 2002; Rees *et al.*, 2003; Daniel, 2005).

Figura 6. Flujo de trabajo mostrando las etapas para la construcción de bibliotecas metagenómicas a partir de muestras de suelo.



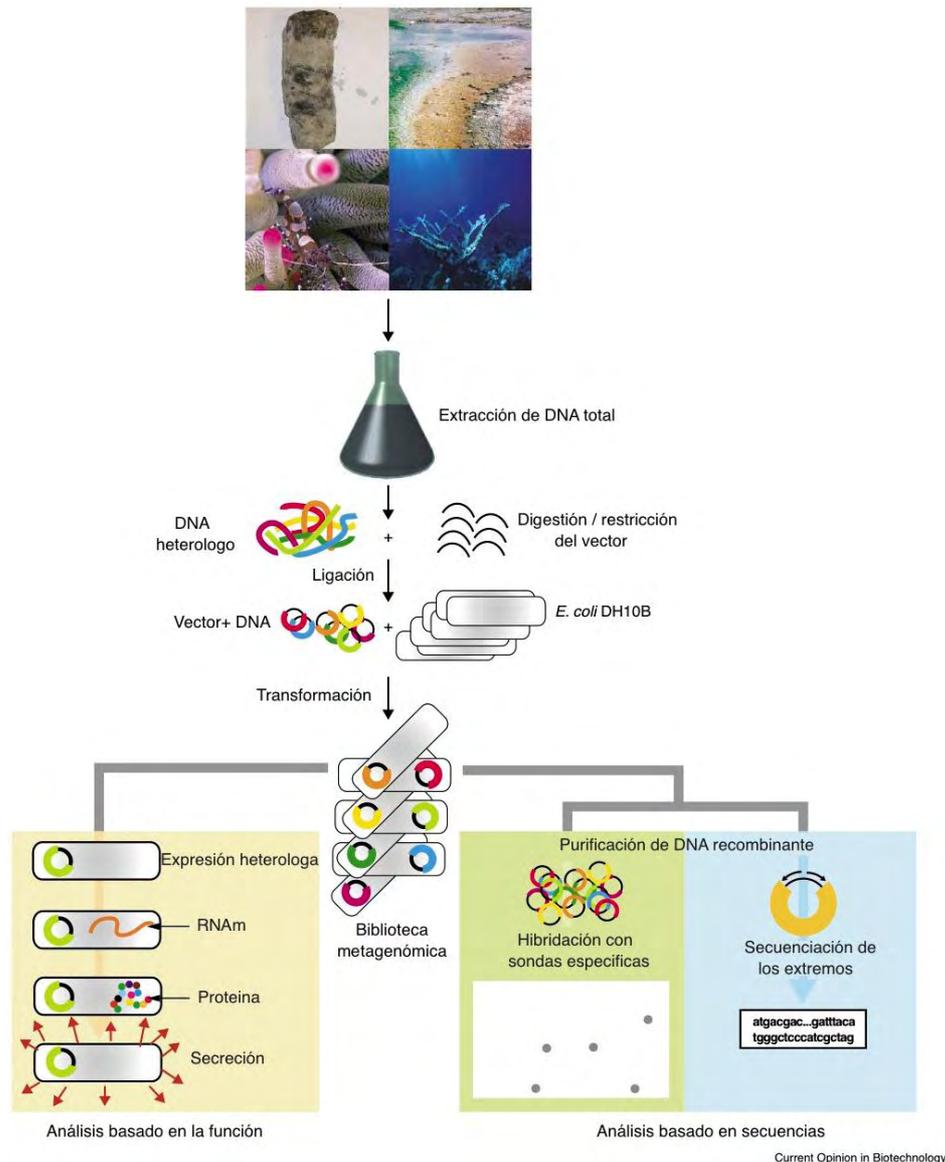
Nature Reviews | Microbiology

Fuente: Adaptado de Daniel, 2005.

El análisis de las bibliotecas metagenómicas se puede realizar a través de dos métodos el primero se basa en la función y el segundo se basa en la secuencia (Schloss & Handelsman, 2003) (Figura 7). El método basado en la función identifica clones que expresan una enzima o una vía metabólica deseada (Jeon *et al.*, 2009). La identificación de expresión se realiza en un medio de cultivo sólido con un sustrato selectivo para la función de interés, generalmente los clones expresan un halo en la

colonia hidrolizando el medio. El otro método está basado en la amplificación de secuencias conservadas de interés por lo cual se usan sondas para la hibridación de primers que amplifican la región de interés. La secuenciación de estos clones permite analizar las secuencias que posiblemente tienen genes nuevos de interés (Schloss & Handelsman, 2003).

Figura 7. Técnicas para el análisis de bibliotecas metagenómicas.



Fuente: Adaptado de Schloss & Handelsman, 2003.

### 6.2.1 *Screening* de bibliotecas metagenómicas

Una biblioteca puede contener varios cientos de miles de clones. El problema ahora es identificar y seleccionar sólo el clon o clones que contienen el gen o genes que nos interesa, y determinar si un clon determinado contiene todo el fragmento o sólo una parte de él. Hay varias maneras de hacerlo, y la elección depende de las circunstancias y de la información disponible. Los métodos que se describen a continuación utilizan diversos enfoques para encontrar una secuencia específica de DNA en una biblioteca.

En metagenómica las bibliotecas de clones se rastrean por medio de expresión o técnicas moleculares:

Las formas tradicionales para selección de clones por análisis de expresión son por medio de observación fenotípica que envuelven resistencia a metales pesados o antibióticos (Olsen *et al.*, 1994); estos análisis incluyen cambios visibles en las colonias en medios diferenciales y generalmente incluyen color, formación de claros (halo), e hidrólisis ( Rondon *et al.*, 2000; Fu *et al.*, 2011; Park, Shin, & Kim, 2011). La ventaja de usar este método es que no requiere equipamientos especiales para la detección, solo para el posterior análisis del producto por cromatografía de gases (HPLC), o cromatografía de gases acoplada a masas (GS-MS). La desventaja de realizar *screening* por observación es que algunos clones tienen expresión tan baja que no hay un cambio aparente en la colonia, además los análisis pueden tardar mucho tiempo por la gran cantidad de clones que compone una biblioteca (Olsen *et al.*, 1994).

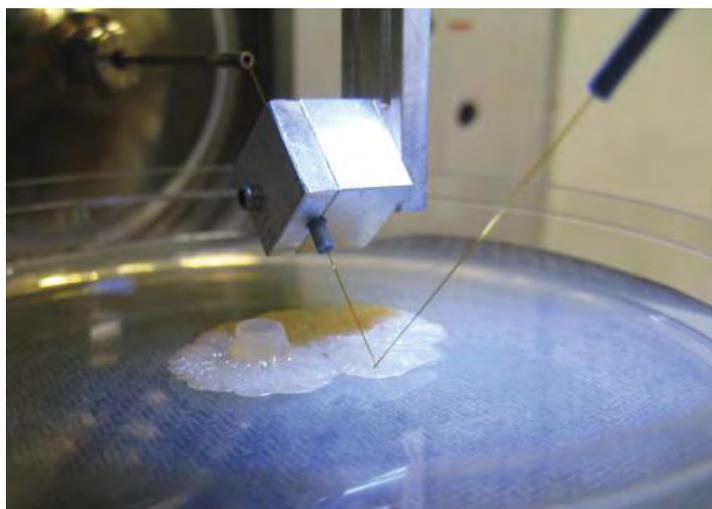
Actualmente, se comenzaron a usar metodologías que incluyen robots, estos pueden analizar placas de 96 o 384 clones, posibilitando y viabilizando el trabajo que manualmente puede tardar mucho tiempo (Streit & Daniel, 2010). La aplicación de automatización robótica para el *screening* de expresión no sólo reduce los costos de mano de obra, también aumenta la reproducibilidad y comparabilidad entre las muestras.

Una de las tecnologías disponibles desde el año 2012 es el análisis de

metabolitos directamente de colonia, “Nanospray Desorption ElectroSpray Ionization” Técnica que usa el equipo NanoDESI, este es un nanoespectrómetro de masas para el análisis detallado del contenido molecular de la muestra biológica (Watrous *et al.*, 2012).

El equipo incluye dos pequeños tubos de vidrio o capilares en serie, una corriente continua de disolvente se pone en contacto con la superficie de la muestra (Figura 8). Este disolvente puede ser cualquier muestra biológica o colonia. Cuando el disolvente toca la muestra, el material sobre el sustrato se disuelve. El segundo capilar, que actúa como una bomba, convierte los aerosoles disueltos en iones, se analizan en un espectrómetro de masas LTQ-Orbitrap alta resolución (Watrous *et al.*, 2012).

Figura 8. Fotografía de nanoDESI con una colonia microbiana sobre la superficie de agar solido en caja Petri.



Fuente: (Watrous *et al.*, 2012).

El *screening* de bibliotecas por medio de técnicas moleculares incluyen PCR como una de las técnicas más empleadas que se basa en secuencias específicas de oligonucleótidos para buscar en estas bibliotecas genes de interés a partir de DNA total (Courtois *et al.*, 2003; Ginolhac *et al.*, 2004; Parachin & Gorwa-Grauslund, 2011). Sin embargo, la PCR puede tener dificultad porque depende de la información ya existente para el diseño de iniciadores en la amplificación; lo que tornaría limitante la

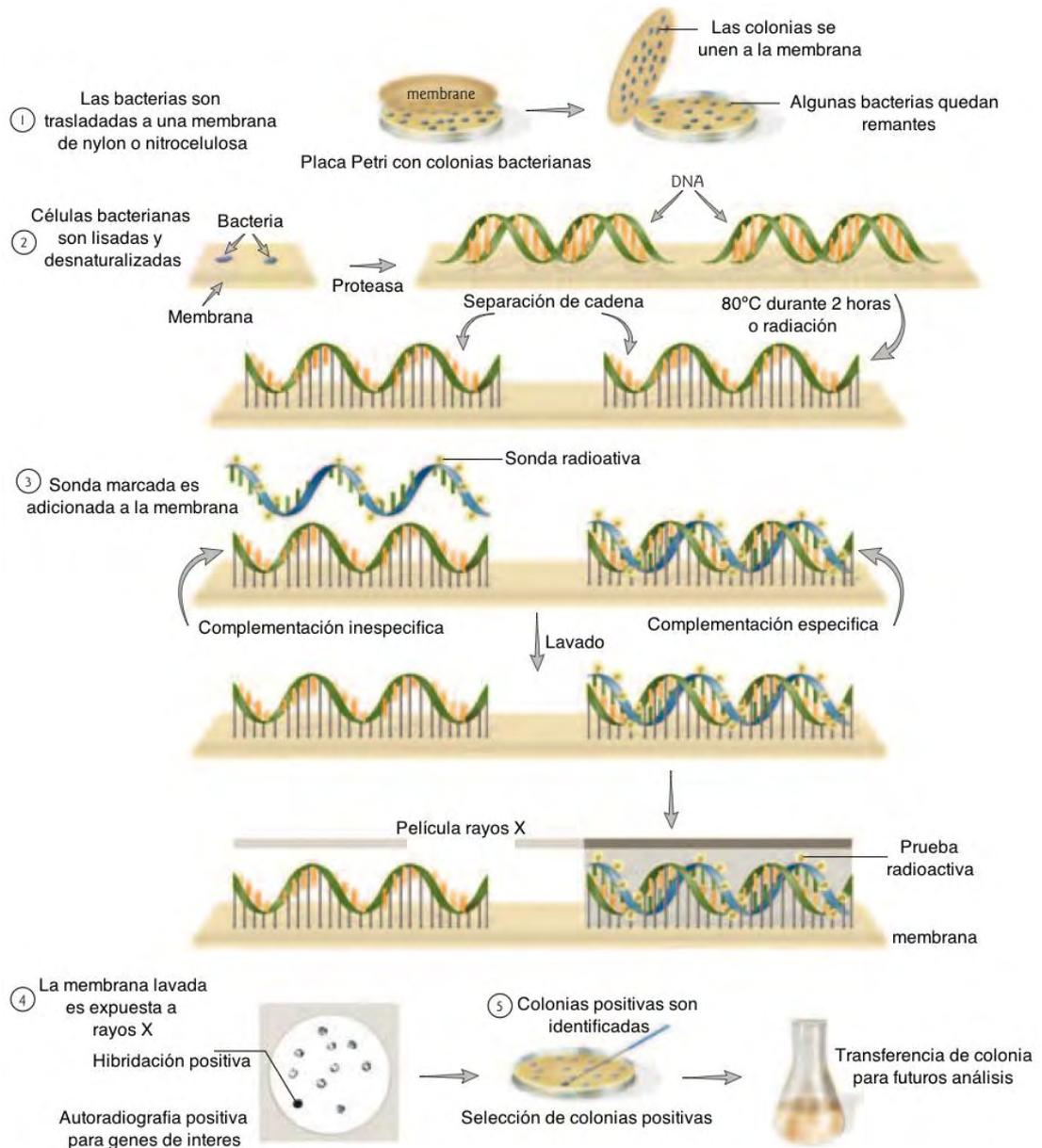
detección de nuevos genes; alternativamente se puede abordar la PCR con iniciadores degenerados o de secuencias conservadas que permiten tener acceso a variantes de genes conocidos (Cowan *et al.*, 2005).

La hibridación de colonia es también una técnica usada para *screening* de bibliotecas, el método se lleva a cabo por medio de sondas marcadas específicamente, estas usan un isótopo radioactivo, usualmente P<sup>32</sup> es la alternativa de radioisótopo más utilizada ( Ginolhac *et al.*, 2004; Angelov *et al.*, 2009).

El método comienza con la replicación de las colonias a un filtro de nitrocelulosa, y las colonias se rastrean mezclando el filtro con las sondas de ácido nucleico. Por este método se debe conocer parte de la secuencia del gen o del DNA que se está rastreando. Si se utiliza una sonda radioactiva para rastrear la biblioteca, la sonda de DNA se desnaturaliza para formar moléculas de cadena sencilla y se añade a la solución en la que el filtro está sumergido. Si la secuencia de la sonda es complementaria a cualquiera de las secuencias clonadas de las colonias, se formará un híbrido DNA-DNA entre la sonda y el inserto. Después se lava el filtro, para eliminar las sondas que no hibridaron, y se cubre el filtro con un trozo de película fotográfica. La sonda marcada cuando es complementaria al DNA, forma un complejo radioactivo que puede ser identificado como una impresión sobre la película. Las colonias positivas pueden recuperarse de la placa inicial, y el DNA clonado que lleva, puede utilizarse para nuevos experimentos (Figura 9). Las sondas no radioactivas utilizan métodos quimioluminiscentes o colorimétricos para detectar el clon (o clones) de interés (Cultek, 2006).

Tanto la técnica de PCR como la de hibridación requieren de iniciadores y sondas apropiadas que son derivadas de regiones conservadas de genes conocidos y de productos génicos (Daniel, 2005).

Figura 9. *Screening* por hibridación de ácido nucleico. 1) La muestra con las colonias que componen la biblioteca se transfiere a una membrana 2) las bacterias son lisadas y el DNA desnaturalizado, 3) la membrana se coloca en una bolsa con la sonda marcada, luego la membrana es retirada y lavada 4) los híbridos son detectados por autorradiografía de rayos X 5) las colonias que contienen el gen de interés hibridan y son seleccionadas.



Fuente: Adaptado de (Allison, 2009).

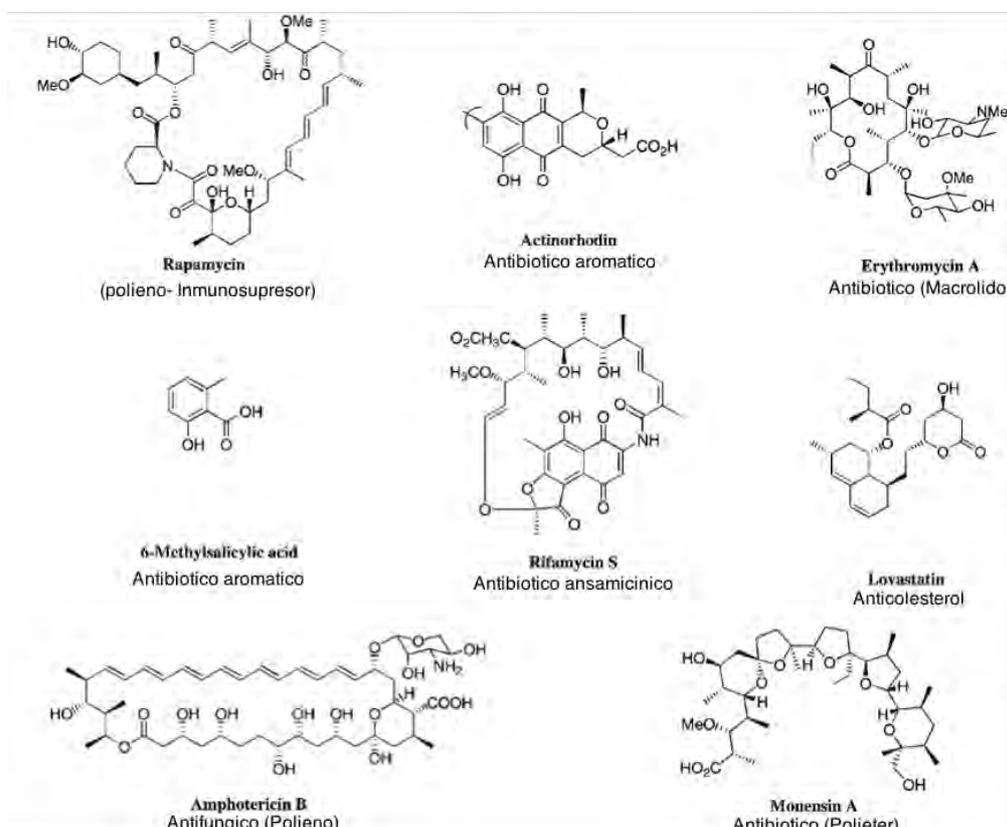
### 6.3 Policétidos

El crecimiento celular está determinado por las actividades metabólicas primarias, esta actividad produce enzimas, ácidos orgánicos, etanol, entre otros. Sin embargo, los metabolitos secundarios son sintetizados como acompañantes no esenciales en algunas fases del crecimiento celular (Kleinkauf, 1996).

El metabolismo secundario se torna interesante debido a la producción de una gran variedad de compuestos de interés (antibióticos, antifúngicos, inmunosupresores, antitumorales, inhibidores de enzimas, toxinas, pigmentos, entre otros), los cuales son moléculas orgánicas complejas producidas por una cadena de reacciones enzimáticas (Demain, 1992; Massini, 2009).

Entre estos metabolitos secundarios se encuentran los policétidos, que son una familia de compuestos estructuralmente muy diversa con un amplio espectro de actividad (Figura 10); más de un tercio de los productos que se producen por organismos vivos aprobados como fármacos son moléculas policetónicas, a pesar de su diversidad todos comparten un mismo patrón biosintético, donde las macromoléculas clave para la biosíntesis son enzimas denominadas policétido sintetasas (Hopwood, 2004).

Figura 10. Ejemplos de policétidos como metabolitos secundarios indicando la actividad biológica.

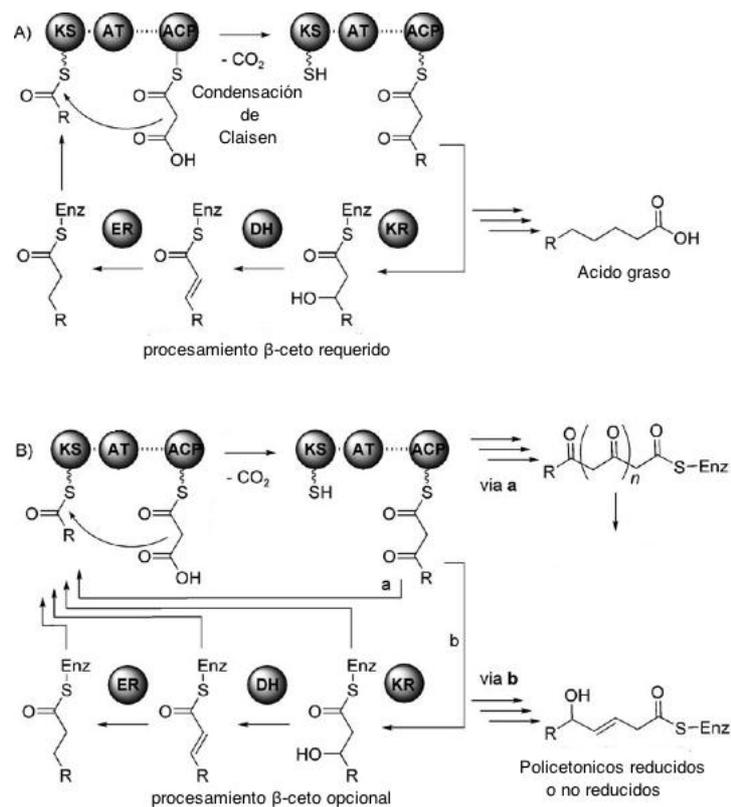


Fuente: (Hertweck, 2009).

En las rutas de síntesis de policétidos existe alta similitud con los mecanismos de síntesis de ácidos grasos, las enzimas sintetas (Hutchinson & Fujii, 1995). El sistema enzimático que cataliza la síntesis de ácidos grasos es llamado por ácido grasos sintetas (FAS) y está involucrado en la síntesis de los policétidos (PKS), ambos los FAS y los PKS necesitan que los grupos acetil y malonil, que participan de la condensación, de las cadenas finales. La esencia de la síntesis de ácidos grasos y de policétidos está en la condensación descarboxilativa en la cual los ácidos carboxílicos simples se unen para producir una cadena de 6 a 50 átomos de carbono, siendo que cada unidad contribuye con 2 carbonos de la estructura principal en la cadena. (Hopwood, 2004) .

La síntesis de moléculas policetónicas comienza con la unión de unidades de Acetil CoA o Malonil CoA, por medio de la unidad KS, conocida como iniciadora, luego la unidad AT transfiere moléculas acilas a la proteína de transferencia ACP, la unidad KS continua con las condensaciones descarboxilativas entre las moléculas acilas para generar  $\beta$ -ceto-acil-ACP. Por medio de este proceso repetitivamente se originan cadenas con múltiples grupos ceto dentro de su estructura (policétidos) (Figura 11) (Staunton & Weissman, 2001; Hopwood, 2004), la síntesis de moléculas policetónicas, esta acompañada de otros grupos que participan en el procesamiento final de la molécula, como : Cetoreductasas (KR) que cataliza la reducción del grupo  $\beta$ -ceto, Dehidratasa (DH) esta unidad transforma alcohol (-OH) en C=C y la unidad reductasa (ER) reduce el C=C a C-C (Reeves, 2003; Hertweck, 2009).

Figura 11. Mecanismo básico involucrado en la síntesis de a) ácidos grasos b) policétidos



Fuente: Hertweck, 2009.

### 6.3.1 Tipos de PKS

Los policétidos se dividen en 3 grandes grupos:

Tipo I: Son moléculas complejas que pueden ser grandes o pequeñas, que son sintetizadas por sintetasas I como estainas y eritromicinas (Austin & Noel, 2003).

Tipo II: Son moléculas aromáticas sintetizadas por enzimas sintetasas II multifuncionales, como las antraciclinas (antitumorales) (Austin & Noel, 2003).

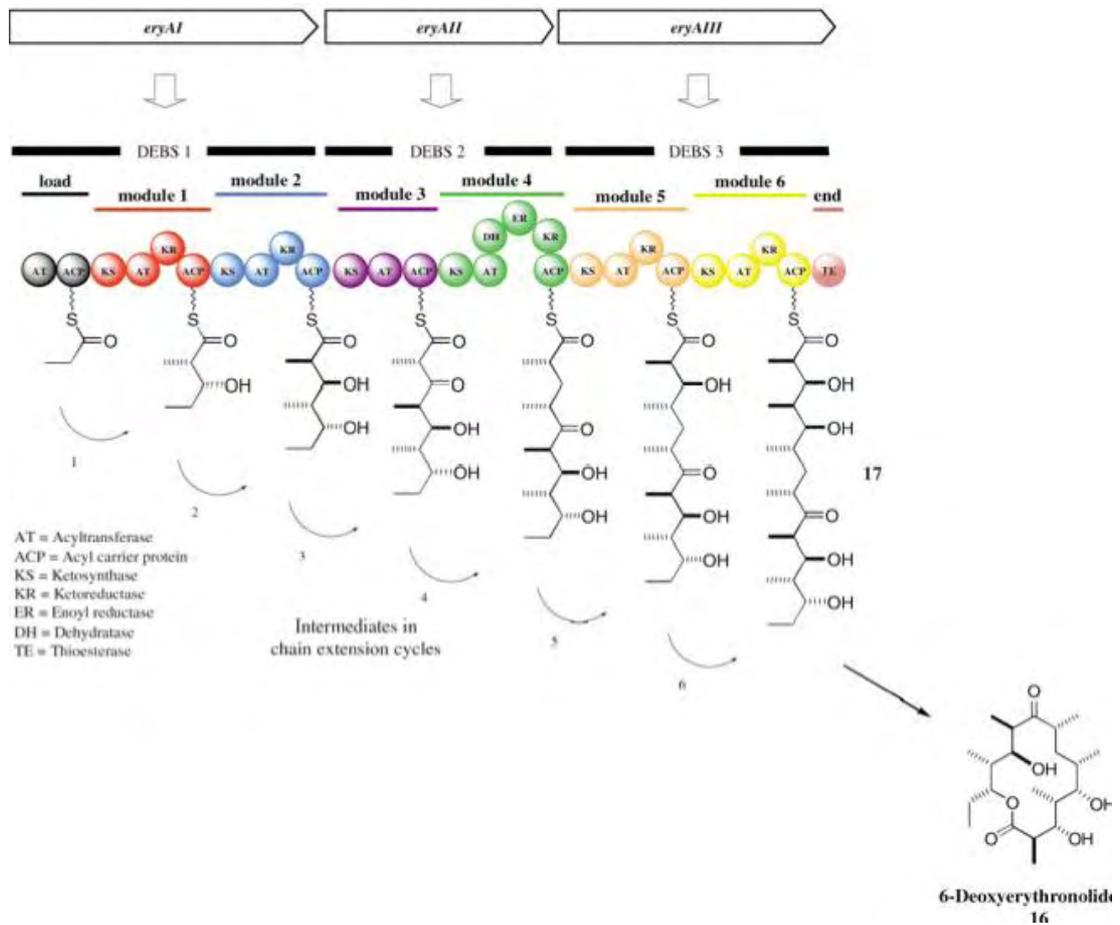
Tipo III: Son moléculas producidas principalmente por plantas superiores, por sintetasas III; como por ejemplo los flavonoides (Austin & Noel, 2003).

#### 6.3.1.1 PKS tipo I

Los complejos multienzimáticos que hacen parte del sistema PKS I contienen múltiples sitios activos, denominados dominios catalíticos, los cuales actúan como módulos de elongación para formar la cadena policetídica. Los dominios catalíticos necesarios para que se lleve a cabo la biosíntesis de la cadena son KS, AT y ACP, mientras que la presencia de los dominios catalíticos KR, DH, y ER, va a depender del grado de reducción de la cadena en formación (Hopwood, 2004).

La vía ocurre en tres pasos para catalizar la extensión de la cadena en niveles diferentes de la reducción de la cadena en cada fase. La condensación en primera instancia se realiza sobre el grupo malonato, seguido de una segunda condensación reducción y deshidratación de un grupo ceto recién formado, consiguiente a este proceso la cadena sufre ciclación (Staunton & Weissman, 2001). A medida que se incorporan unidades de elongación mediante la acción de los polipéptidos que constituyen las PKS tipo I, se va formando la cadena policetídica que se libera del complejo enzimático. Además, la cadena ciclada por acción de las citocromo P450 da lugar a la estructura final por medio de transferasas (Kim *et al.*, 2005) como ejemplo de este tipo de moléculas encontramos la eritromicina (Figura 12).

Figura 12. Clúster génico codificante de las sintasas que forman la eritromicina (PKS I)



Fuente: Staunton & Weissman, 2001.

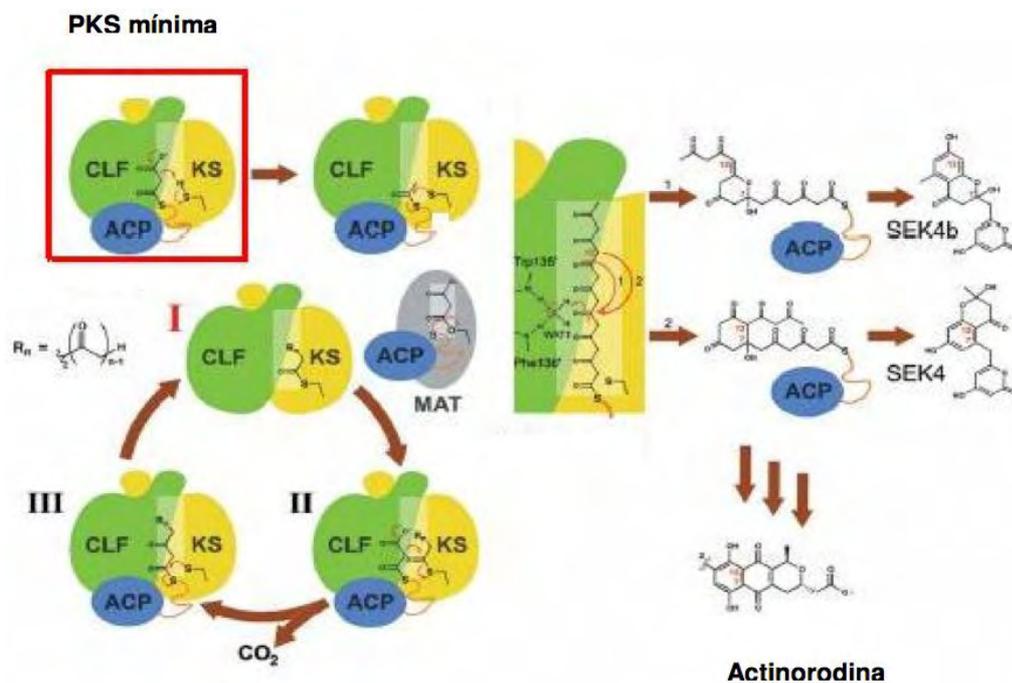
### 6.3.1.2 PKS II

Los mecanismos de las enzimas PKS II son completamente diferentes a las tipo I, porque estas enzimas son un conjunto de proteínas monofuncionales que catalizan la formación de compuestos aromáticos cíclicos y no requieren amplios ciclos de reducción o deshidratación (Figura 13) (Hutchinson & Fujii, 1995; Hertweck *et al.*, 2007).

En las PKS de tipo II, el complejo enzimático se denomina PKS mínima (Figura 14), con dos sub unidades  $KS\alpha$ ,  $KS\beta$  y ACP. La proteína  $KS\alpha$  está implicada en la condensación, la proteína  $KS\beta$  determina la longitud de la cadena policetídica y la ACP

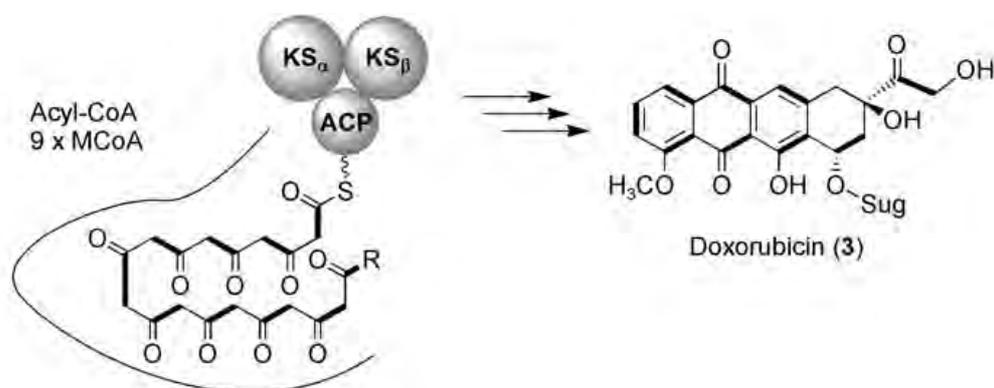
se encarga del transporte de grupos acilo. Sin embargo, la acción de las enzimas codificadas por las PKS mínima no determinan la formación del compuesto final, por lo que necesitan enzimas como ciclasas, aromatasas, y reductasas para la formación final de policétidos tipo II (Metsä-Ketelä *et al.*, 1999; Keatinge-Clay *et al.*, 2004).

Figura 13. Ruta biosintética de la producción de la molécula actinorodina en *Streptomyces coelicolor*, indicando la catálisis de policétidos sintasas.



Fuente: (Keatinge-Clay *et al.*, 2004).

Figura 14. Ejemplo de biosíntesis de doxorubicina (PKS II)



Fuente: Hertweck, 2009.

## 7. MATERIALES Y MÉTODOS

### 7.1 Recuperación y análisis de viabilidad de la biblioteca metagenómica

La biblioteca metagenómica, está compuesta por clones de *Escherichia coli* DH10B con fragmentos de próximos a 50 Kb de DNA total de una muestra proveniente de suelo Atlántico Brasileiro. La construcción de la Biblioteca fue realizada por Massini en el año 2004, usando múltiples protocolos de extracción (Anexo1), como resultado obtuvo DNA de buena calidad, la relación de absorbancia 260/280 osciló entre 1,8 a 2.0. El ácido nucleico fue digerido con la enzima *Hind* III, generando fragmentos cercanos de 50 Kb. El DNA fue ligado con un vector pBAC OriV (~8kb), el producto fue transformado en *E.coli* DH10 en medio LBAC (1% de triptona; 1% NaCl; 0,5% de extracto de levedura); 12,5µg/ml de cloranfenicol; 2g/L de arabinosa e incubado a 37°C. Las células transformantes fueron pasadas en medio de cultivo líquido en placas de polipropileno de 96 pozos y preservadas con glicerol al 60% a -80°C para posteriores estudios.

Los clones de la biblioteca fueron removidos del freezer a -80°C y descongelados en baño de hielo durante 5 minutos, los clones fueron recuperados en placas de polipropileno de 96 pozos con 150 µL medio (LBAC), arabinosa 2 g/L y cloranfenicol 12,5 µg/µL. La transferencia de los clones fue realizada usando un replicador de 96 dientes metálicos estériles, flameado aproximadamente 4 veces con alcohol al 70%. Luego de este periodo, fue registrado el crecimiento de los clones en placas por triplicado, los clones que presentaron crecimiento fueron preservados con 50 µl de glicerol al 60% a una temperatura de -80°C.

## 7.2 Extracción de cromosoma artificial bacteriano pBAC OriV (~8kb)

Debido al tamaño del inserto en la construcción de la biblioteca fueron evaluados cuatro Kits de extracción de vectores:

- 1) GeneJET Plasmid Miniprep Kit (Thermo Scientific)
- 2) AxyPrep Plasmid Miniprep Kit (Axygen Biosciences)
- 3) PureYield™ Plasmid Miniprep System (Promega)
- 4) QIAprep Spin Miniprep Kit (Qiagen)

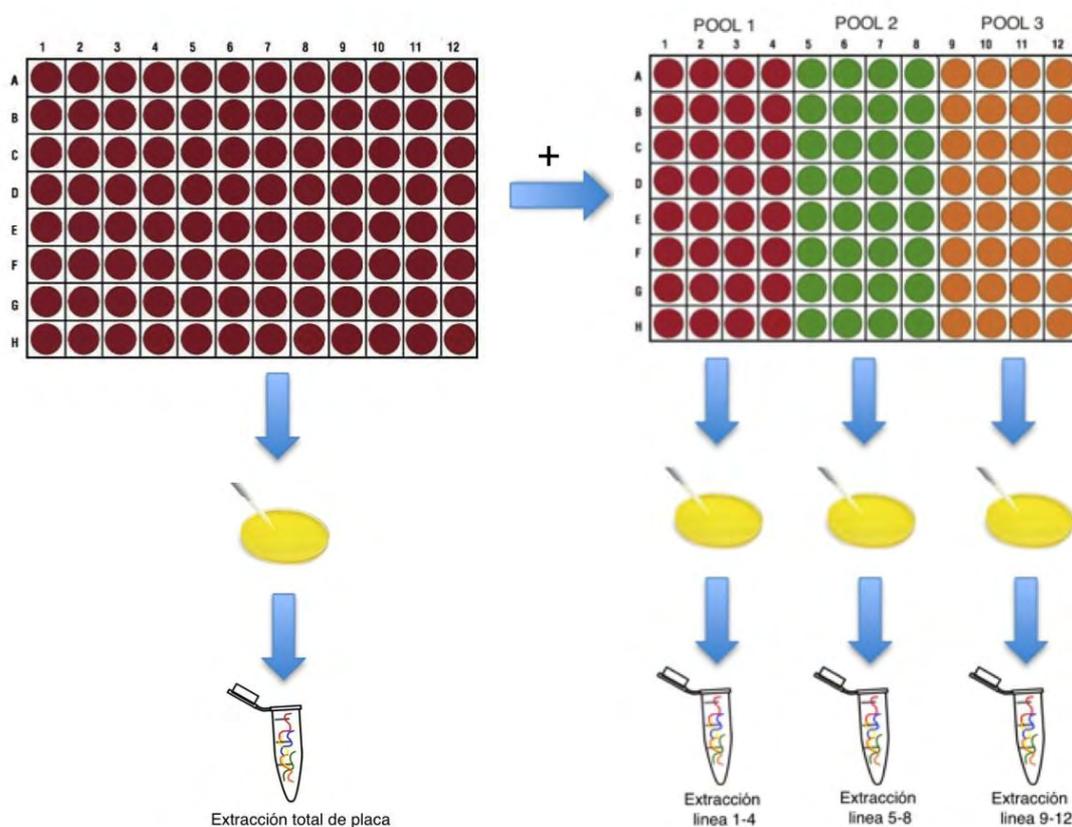
Los ensayos se realizaron en clones identificados como transformantes de la biblioteca, para determinar cuál es el mejor método de extracción de DNA del vector pBAC con el inserto. Los clones seleccionados se cultivaron en medio líquido con 4 mL de medio líquido LBAC a 37°C durante 18 horas, las extracciones se realizaron de acuerdo a las instrucciones de los fabricantes, adicionalmente, se realizó una nueva extracción teniendo en cuenta la sugerencia de los fabricantes en incrementar la temperatura de los buffers de elución.

La pureza y concentración del DNA (BAC) fue evaluada por espectrofotometría (260nm/280nm) (Sambrook *et al.*, 1989), se usó el equipo (Nanodrop ND 2000 *Thermo Scientific*). El DNA (BAC) fue sometido a electroforesis en gel de agarosa a 0,8%, corrido en una cámara de electroforesis (*Amersham biosciences*) a 100 V durante 50 minutos, el gel se tiñó con bromuro de etidio y se visualizó en fotodocumentador *Kodak gel logic 1000 Imaging System* acoplado a luz UVP.

## 7.3 *Screening* primario

Una vez estandarizado el proceso de extracción de cromosoma artificial bacteriano, se realizó la extracción y amplificación para genes PKS I y PKS II, inicialmente en grupos de 96 clones, los grupos positivos se dividieron en 3 grupos menores de 32 clones denominados (*pools*) (Figura 15) Galindo (2011); una vez identificados los *pools* positivos, se realizó la extracción y amplificación de cada uno de los clones que componen el *pool*.

Figura 15. Esquema de extracción de DNA plasmidial de cada una de las placas de polipropileno.



Fuente: Adaptado y modificado de Galindo, 2011.

#### 7.4 Amplificación para detección de genes del sistema PKS I y II

Para la amplificación de los genes biosintéticos de moléculas policetónicas, se utilizaron primers degenerados que se complementan a regiones conservadas de genes que codifican a enzimas esenciales en la biosíntesis de moléculas policetónicas. Para los genes del sistema PKS I se usaron primers degenerados que son complementarios a regiones altamente conservadas del sitio activo de la enzima asociado al sistema (Tabla 2) (Courtois *et al.*, 2003). De igual manera, los primers de genes biosintéticos del sistema PKS II se detectaron por medio de la amplificación de la región conservada KSα. (Tabla 3) (Metsä-Ketelä *et al.*, 1999).

Tabla 2. Secuencias de los primers degenerados que se utilizaron para la detección de policétidos PKS de tipo I.

SECUENCIA	REFERENCIA
5'- CCSCAGSAGCGCSTSTTSCTSGA-3'	Curtois, <i>et al.</i> , 2003
5'- GTSCCSGTSCCGTGSGTSTCS-3'	
5'- CCSCAGSAGCGCSTSCTSCTSGA-3'	
5'- GTSCCSGTSCCGTGSGCCTCSA-3'	

Símbolo de los *primers* degenerados: S= C o G; Y= C o T; B= C, G o T y N= A, T, C o G.

Tabla 3. Secuencias de los primers degenerados que fueron utilizados para la detección de policétidos PKS de tipo II.

SECUENCIA	REFERENCIA
5'- TSGCSTGCTTCGAYGCSATC-3'	Metsa-Ketelä <i>et al.</i> , 1999
3'- TCGCCBAAGCCGCCNAAGGT-5'	

Símbolo de los *primers* degenerados: S= C o G; Y= C o T; B= C, G o T y N= A, T, C o G.

Se estandarizaron en su totalidad las condiciones y concentraciones de reactivos para la amplificación por PCR de genes del sistema PKS I y (Curtois, *et al.*, 2003 y Metsa-Ketelä *et al.*, 1999 (Tablas 4 y 5), tanto para la amplificación de los primers PKS I y II. Las condiciones de amplificación se describen a continuación.

Tabla 4. Protocolo inicial para estandarización de los genes PKS tipo I en DNA plasmidial.

PROGRAMA DE TERMOCICLADO	
DESNATURALIZACIÓN INICIAL	95°C por 15 minutos
DESNATURALIZACIÓN CÍCLICA	95°C por 1 minuto
ANILLAMIENTO INICIAL	65°C 1 minuto
EXTENSIÓN INICIAL	72°C 1 minuto
ANILLAMIENTO CÍCLICO	62°C 1 minuto
EXTENSIÓN CÍCLICA	1 minuto a 72°C
EXTENSIÓN FINAL	1 ciclo de extensión final de 10 minutos.
CONCENTRACIÓN DE LOS REACTIVOS	
Concentración inicial ( <i>Stock</i> )	Concentración final de la reacción
DNTP's ( <i>Thermo Scientific</i> 25mM)	200 µM de dNTP ( <i>Thermo scientific</i> )
Primers ( <i>Exxtend</i> -100 mM)	0,4 µM de cada <i>primer</i>
DNA Polimerasa ( <i>Thermo Scientific</i> 5 U/µl)	2,5 U de <i>Taq</i> polimerasa ( <i>Thermo scientific</i> )
Buffer Para PCR ( <i>Thermo Scientific</i> )	10 X (20 mM Tris-HCL pH 8,4; 50 mM KCL)
MgCl <sub>2</sub> ( <i>Thermo Scientific</i> )	2,5 mM de MgCl
Volumen Final	50 µL
Ciclos De Repetición	30

Tabla 5. Protocolo inicial para la detección PKS tipo II en DNA total de suelo.

PROGRAMACIÓN TERMOCICLADO	
DESNATURALIZACIÓN INICIAL	96°C por 2 minutos
DESNATURALIZACIÓN CÍCLICA	96°C por 1 minuto
ANILLAMIENTO CÍCLICO DE LOS PRIMERS	anillamiento de 1 minuto con temperaturas entre 56° C a 64 °C
EXTENSIÓN CÍCLICA	1 minuto a 72°C
EXTENSIÓN FINAL	1 ciclo de extensión final de 10 minutos. 30 Ciclos
CONCENTRACIÓN DE LOS REACTIVOS	
Concentración inicial ( <i>Stock</i> )	Concentración final de la reacción
DNTP's ( <i>Thermo Scientific 25mm</i> )	1 mM de dNTP' ( <i>Thermo scientific</i> )
Primers ( <i>Exxtend-100 Mm</i> )	100 pM de cada <i>primer</i>
DNA Polimerasa ( <i>Thermo Scientific 5 U/μl</i> )	2 U de <i>Taq</i> polimerasa ( <i>Thermo scientific</i> )
Buffer Para PCR ( <i>Thermo Scientific</i> )	10 X (20 mM Tris-HCL pH 8,4; 50 mM KCL)
MgCl <sub>2</sub> ( <i>Thermo Scientific</i> )	2 mM de MgCl <sub>2</sub>
Volumen Final	50 μL

Tabla 6. Microorganismos usados en este estudio como controles positivos para amplificación de genes biosintéticos PKS I y II

Controles positivos que se usaron en la amplificación.

MICROORGANISMOS	CARACTERÍSTICAS DEL CONTROL
<i>Streptomyces olindensis</i> DAUFPE 5622 Universidade Federal de Pernambuco	Productor de PKS tipo I y II
<i>Streptomyces coelicolor</i> A3-2 Lab De Bioproductos ICB-USP	Productor de PKS tipo I y II

#### 7.5 Reacción de secuenciación.

Los productos amplificados y purificados incluyendo uno de los controles positivos (*S. olindensis*), fueron secuenciados, en el centro de estudios de Genoma Humano de la Universidad de São Paulo con un secuenciador ABI 3730 DNA Analyser (*Applied Biosystems* 4849USA). Las reacciones de secuenciación se realizaron utilizando BigDye® Terminator v3.1 Cycle Sequencing Kit.

#### 7.6 Análisis bioinformático

La edición de las secuencias se realizó con el programa 4 peaks (1.7.1) (Griekspoor & Groothuis, 2006). El alineamiento y la obtención de la secuencia consenso se verificó en el programa Chromas Pro 1.7.5; para el alineamiento múltiple se usó la herramienta Clustal W con el programa MEGA (5.01) (Tamura *et al.*, 2011).

Para el análisis de similaridad, las secuencias de nucleótidos fueron comparadas con la base de datos del GenBank, utilizando respectivamente los programas Blastn y Blastx del centro Nacional de Información y Biotecnología (NCBI).

El árbol filogenético se construyó con la versión (5.01) del programa MEGA (Tamura *et al.*, 2011), se empleó como estrategia de agrupamiento, el algoritmo de Neighbor joining.

Para la comparación de la estructura proteica se realizo un análisis bioinformático usando la base de datos *predict protein* (Yadav *et al.*, 2003), de las secuencias amplificadas asociadas al sistema PKS II, tanto para la secuencia de origen metagenómico como para el control positivo (*S. olindensis*) que corresponde a una cetosintasa. Los sitios activos fueron analizados con el método propuesto por (Ofran y Rost, 2007) que hace parte de la base de datos *predict protein*, y el modelamiento molecular fue realizado con el programa Cn3D del NCBI basado en el blasx y PSI Blast (Anexo 4) de las secuencias analizadas.

## 8. RESULTADOS Y DISCUSIÓN

El proceso de recuperación de la biblioteca, se realizó por medio de la visualización directa de sedimentación en cada uno de los clones en placas de polipropileno con medio fresco LBAC. No existió pérdida de clones con respecto a la biblioteca replicada en el año 2009, sin embargo el número de clones es inferior a la biblioteca original de 2004 con un total de 3456 de un total de 5000. Probablemente, el tiempo de preservación mayor a 5 años sin verificación continua, fue un factor determinante en la disminución de la tasa de viabilidad, y está directamente relacionada con las condiciones de crioprotección y de los linajes bacterianos usados en este estudio (Silva *et al.*, 1992).

### 8.1 Extracción de cromosoma artificial bacteriano pBAC OriV (~8kb)

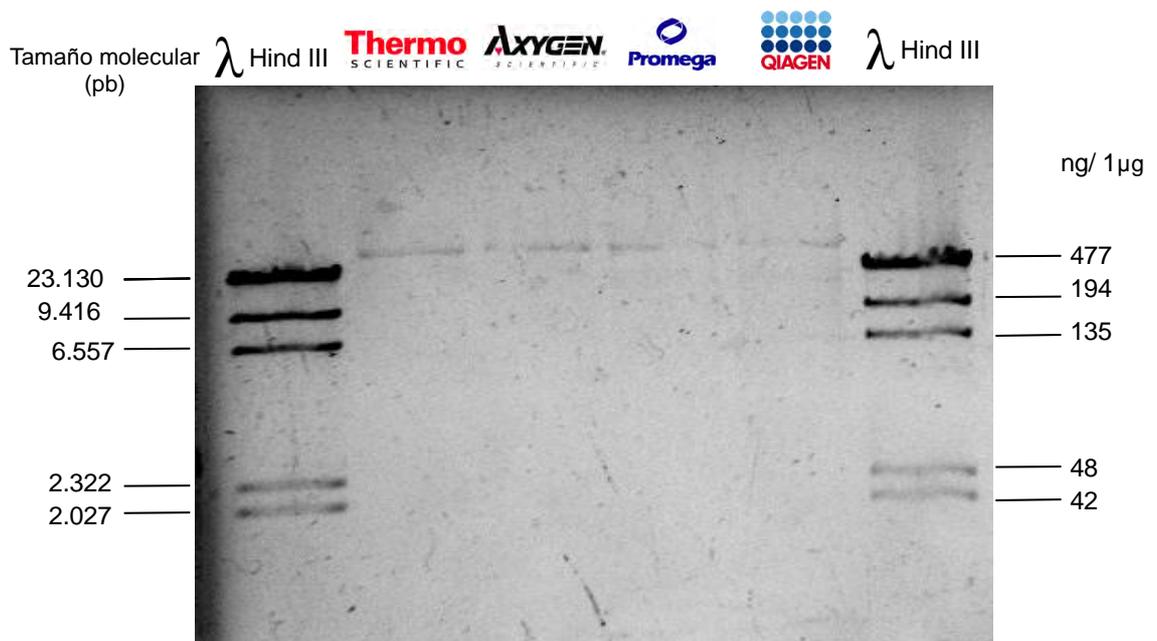
El proceso de extracción es una parte fundamental para el análisis genético y molecular de una biblioteca, el principal inconveniente para ello, fue establecer el protocolo de extracción del vector con insertos cercanos a 50Kb.

La extracción de DNA fue verificada por medio de espectrofotometría y electroforesis en gel de agarosa (Figura 16). Los valores de relación de absorbancia a 260/280 fueron considerados óptimos, todas las muestras presentaron una relación cercana a 2 (Sambrook, *et al.*, 1989).

Con relación a la concentración de DNA, se obtuvieron valores de 13 ng/uL para el kit de *Promega*, 26,4 ng/uL para el Kit de *Qiagen*, 33ng/uL para el Kit de *Axygen* y de 34,6 ng/uL para el Kit de *Thermo*; de forma general se consideraron concentraciones bajas para proseguir con el *screening*.

Se realizaron nuevas extracciones con modificaciones en la temperatura y en el tampón de elución, sin embargo, no se obtuvo un cambio significativo en la concentración. La causa de este resultado es atribuida a la posible persistencia del vector en las columnas de los kits, en principio podría estar relacionada con el gran tamaño del inserto.

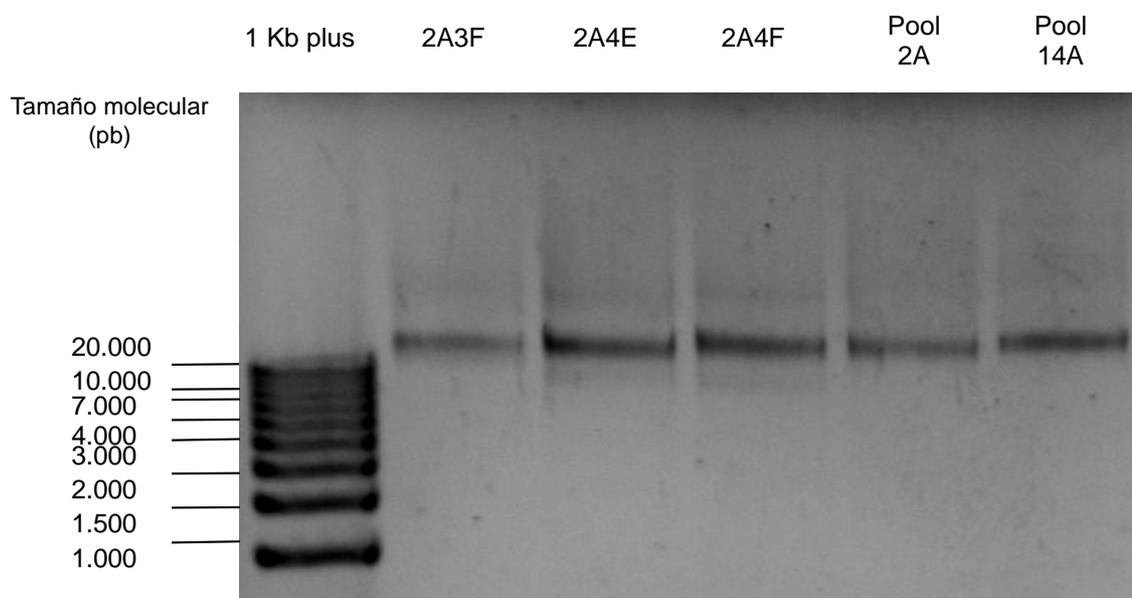
Figura 16. Extracción de DNA plasmidial con diversos Kits comerciales. 1. Marcador  $\lambda$  HindIII (Promega); 2-5. Clon 12A. Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (Amersham biosciencias) teñido en solución de 5µg/mL de bromuro de etidio.



Teniendo en cuenta que la concentración inicial fue muy baja, se optó por realizar extracción clásica usando el protocolo de extracción de vectores de baja escala (miniprep) propuesto por (Sambrook, *et al.*, 1989), método que cuando ensayado sin modificaciones no obtuvo buenos resultados, no se visualizó DNA en la corrida electroforética e igualmente no fue detectable en el NANODROP 2000 (Thermo Scientific); por lo cual se realizaron modificaciones en el tiempo y velocidad de centrifugación, en etapas claves donde se observaron residuos en la extracción inicial, además todos los procesos se realizaron en baño de hielo, y en la etapa de

sedimentación se usó una centrifuga refrigerada a 4°C (*Eppendorf 5415R*). La precipitación de DNA se incrementó por medio de la adición e incubación de etanol absoluto durante 1 hora, esto incrementó la estabilidad y facilitó la obtención del DNA (Anexo 2). Como resultado, se obtuvo un DNA puro, con concentraciones que oscilaron entre 2000 y 3400 ng/uL, igualmente se verificó los productos en gel de agarosa al 0,8% (Figura 17).

Figura 17. Extracción clásica de DNA plasmidial (modificado). 1. Marcador 1Kb plus (*Thermo scientific*); 2-6. Clon 2A3F, clon 2A4E, clon 2A4F, Pool 2A, Pool 14B Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciences*) teñido en solución de 5µg/mL de bromuro de etídio.).



Adicionalmente, se verificó el tamaño aproximado de algunos clones, retirando el inserto por medio de la enzima *Not I*, según el mapa de restricción (Figura 18) del vector pBAC oriV (~8kb) (Wild *et al.*, 2002).

Figura 18. Mapa de restricción vector pBAC oriV (~8kb) (Wild *et al.*, 2002).

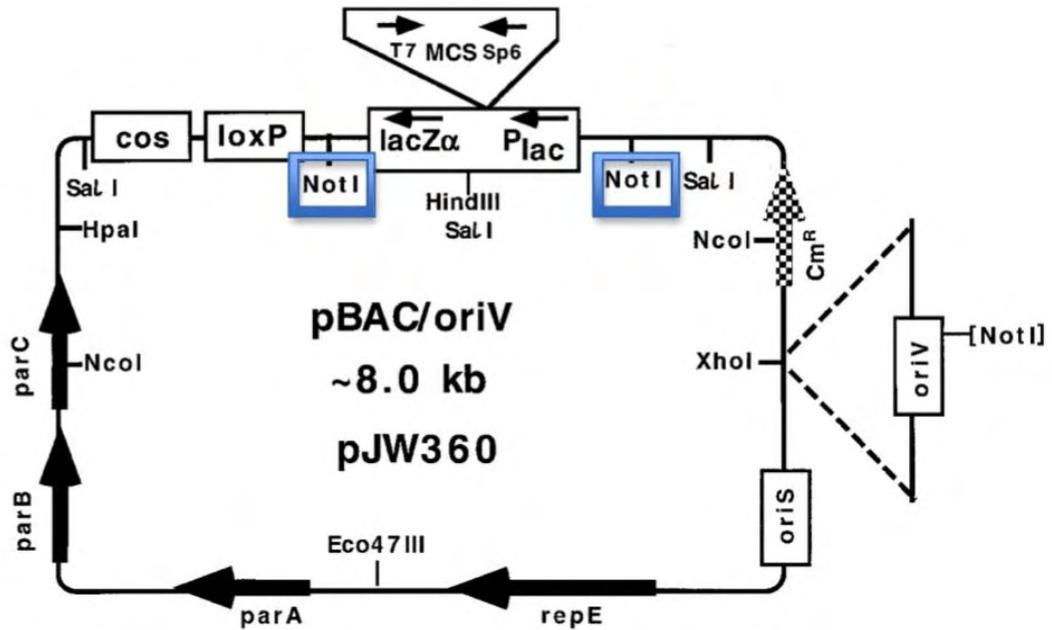
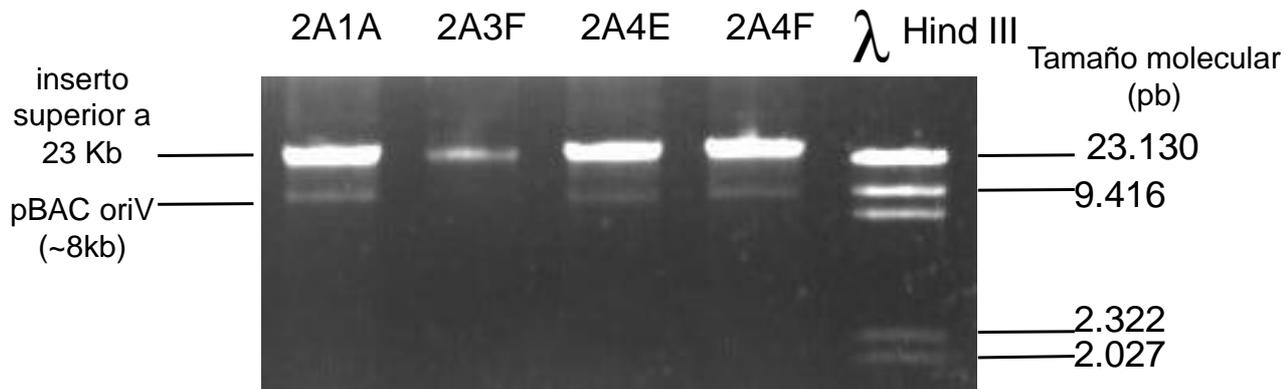


Figura 19. Restricción clones 1-4 2A1A 2A3F, 2A4E, 2A4F 5. Marcador  $\lambda$  *Hind*III (Promega) Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (Amersham biosciencias) teñido en solución de 5µg/mL de bromuro de etidio.



## 8.2 *Screening* primario y amplificación para detección de genes del sistema PKS I y II

Se estandarizaron y optimizaron condiciones de PCR para los sistemas PKS I y II primeramente en controles positivos, luego se utilizaron estas condiciones para los clones que componen la biblioteca metagenómica.

### 8.2.1 PKS I

Los primers utilizados para la amplificación de genes PKS I fueron diseñados por Courtois y colaboradores (2003) que utilizaron regiones consenso para un grupo amplio de firmicutes, estos primers fueron utilizados para realizar un *screening* en una biblioteca metagenómica de origen ambiental.

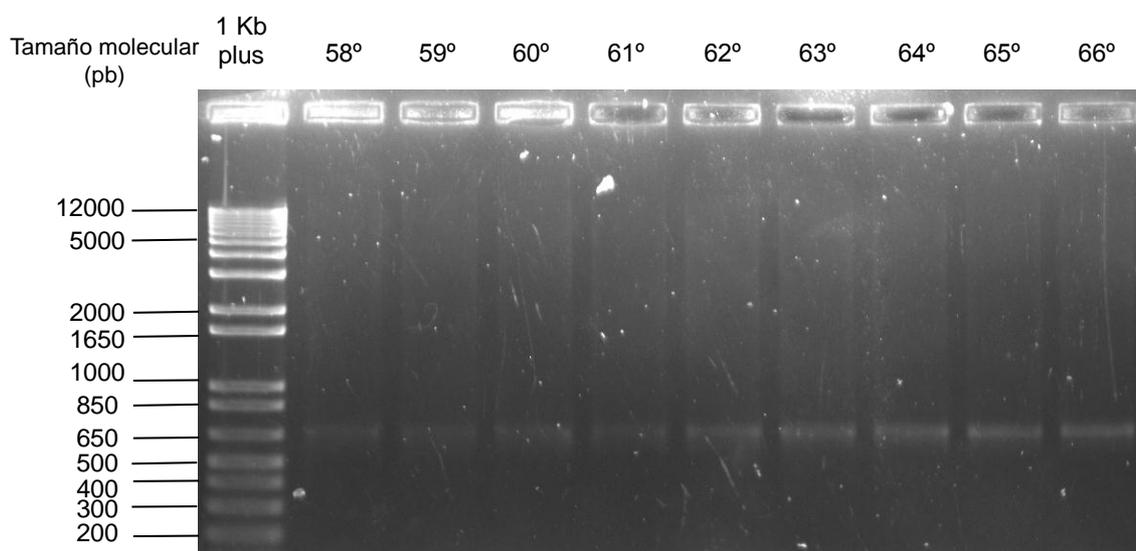
Para este estudio se usaron inicialmente las concentraciones de reactivos y condiciones de PCR, reportadas por Andrielli en 2006 y Massini en 2009, para otras bibliotecas construidas con amplificadas de DNA total, cercanos a 700 pb; sin embargo, para la biblioteca con insertos cercanos a los 50 Kb de este estudio, no se logró amplificación, por lo cual se realizó la optimización de las concentraciones y condiciones de PCR, inicialmente con controles positivos para este sistema en bacterias del genero *Streptomyces*.

Se realizó la amplificación por PCR, usando gradiente de temperatura con 6 temperaturas de anillamiento, en las bacterias *S. olindensis* y *S. avermitilis*, adicionalmente a las condiciones del termociclador (Tabla 8), el mix de reacción incluyo 7% de DMSO (Figura 20).

Tabla 7. Protocolo estandarizado para la detección de PKS tipo I en bibliotecas de insertos cercanos a 50 kb.

PROGRAMACIÓN TERMOCICLADO	
DESNATURALIZACIÓN INICIAL	95°C por 15 minutos
DESNATURALIZACIÓN CÍCLICA	95°C por 1 minuto
ANILLAMIENTO INICIAL	65°C 1 minuto
EXTENSIÓN INICIAL	72°C 1 minuto
ANILLAMIENTO CÍCLICO	62°C 1 minuto
EXTENSIÓN CÍCLICA	1 minuto a 72°C
EXTENSIÓN FINAL	1 ciclo de extensión final de 10 minutos.
CONCENTRACIÓN DE LOS REACTIVOS	
Concentración inicial ( <i>Stock</i> )	Concentración final de la reacción
DNTP's ( <i>Thermo Scientific 25mM</i> )	1 µM de dNTP' ( <i>Thermo scientific</i> )
Primers ( <i>Exxtend-100 mM</i> )	100 pM de cada <i>primer</i>
DNA Polimerasa ( <i>Thermo Scientific 5 U/µl</i> )	2,5 U de <i>Taq</i> polimerasa ( <i>Thermo scientific</i> )
Buffer Para PCR ( <i>Thermo Scientific</i> )	10 X (20 mM Tris-HCL pH 8,4; 50 mM KCL)
MgCl <sub>2</sub> ( <i>Thermo Scientific</i> )	2,5 mM de MgCl
Volumen Final	50 µL
Ciclos De Repetición	30

Figura 20. Amplificación PKS I *Streptomyces olindensis*. 1. Marcador 1Kb plus (*Life technologies*); 2-7. Gradiente Temperatura 58° a 66°. Gel agarosa 1% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciences*) teñido en solución de 5µg/mL de bromuro de etídio.



Según la figura 20, las temperaturas ideales de amplificación se encuentran entre 63 y 66°C; estas temperaturas fueron la base para amplificar los grupos de clones en la biblioteca.

Se realizó extracción de DNA y amplificación por PCR en los grupos de 32 clones *pools*, con las condiciones optimizadas para los controles positivos. La amplificación con las temperaturas base, únicamente fue positiva para los controles *S. olindensis* y *S. avermitilis*; este resultado conllevó a que se realicen nuevas variaciones en la reacción de PCR para el DNA *templete* proveniente de cada uno de los grupos de clones. Se comenzó con un gradiente con temperaturas de 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, y 68°C, además se incluyeron reacciones conteniendo PCR Master Mix (*Promega*) y Platinum® PCR SuperMix (*invitrogen*) (Tabla7). En todos los casos, únicamente amplificó el control positivo, por ello, se consideró que la biblioteca no presentaba clones positivos para genes del sistema PKS I.

Las posibles hipótesis que pueden explicar el porque no se detectaron clones positivos en el *Screening* de PKS son:

La pérdida de clones de la biblioteca construida en 2004 con un número inicial de clones de 5000, a la biblioteca actual que cuenta con 3456 clones, este número representa una pérdida considerable de viabilidad del 30,88%. Los factores determinantes como conservación, manipulación o estabilidad de los linajes hospederos son importantes en el mantenimiento de la viabilidad de este tipo de colecciones biológicas.

Como toda disciplina, la metagenómica tiene algunas dificultades; una de las más importantes es la ausencia de estandarización en las metodologías de protocolos de extracción de DNA que varían en sesgos y eficiencia, se usan diferentes volúmenes de muestra y varía mucho en el número de células por muestra, calidad y cantidad de DNA total utilizado en el ensamblaje de bibliotecas, por lo que la comparación confiable entre metagenomas actualmente es difícil (Raes *et al*, 2007), además la aplicación de metodologías moleculares como: extracción, purificación, amplificación por PCR, clonaje y secuenciación de DNA en estudios que son independientes de cultivo de microorganismos, presentan algunas limitaciones, y la mayoría está asociada a la obtención directa de DNA ambiental (Theron & Cloete, 2000; Kirk *et al.*, 2004;). La lisis desigual de bacterias Gram-positivas y Gram negativas durante los procesos de extracción de DNA y la pérdida de material genético durante los procesos de purificación de DNA ambiental, co-extracción de compuestos considerados como “inhibidores enzimáticos” como ácidos orgánicos, son una de las principales dificultades inherentes del abordaje de las técnicas independientes de cultivo (Lorenz *et al.*, 2002; Tringe & Rubin, 2005).

Courtois y colaboradores en 2003 construyeron una biblioteca metagenómica usando como vector Cósmidos, el trabajo da como resultado 11 clones positivos que presentan sistemas génicos PKS tipo I, de un total de 5000 clones.

Ginolhac y colaboradores en 2004, construyeron una biblioteca metagenómica usando como vector fosmidos, con un total de 100.000 clones, de los cuales se detectaron 40 clones con secuencias ya publicadas en las bases de datos y 23 clones con secuencias nuevas que fueron incluidas en el análisis de ese estudio.

Van Elsas y colaboradores en 2008, construyeron una biblioteca usando DNA total de suelo, como vector usaron fosmidos, la biblioteca estaba compuesta por 15.000 clones, el objetivo del trabajo fue realizar un *Screening* de policétidos tipo I, con la finalidad de rastrear compuestos bioactivos contra *Rhizoctonia solani*. Como resultado, los autores encontraron 3 clones positivos de este tipo, lo cual corrobora que existe un número considerablemente bajo con respecto al número total de clones.

Parsley y colaboradores en 2010, buscando genes relacionados con la síntesis de policétidos, a partir de una biblioteca de 10.000 clones, lograron obtener 34 clones positivos por la técnica de Southern Blot, estos clones no fueron detectados por la técnica tradicional de PCR, esto indicaría que es importante establecer más de una metodología de *screening* para bibliotecas, Parsley incluye rastreo con macroarrays obteniendo resultados similares a los de Southern Blot.

La ausencia de fragmentos amplificados de sintasas de policetónicos de tipo I, pudo haber ocurrido por la baja representatividad de este gen en el DNA total de suelo. O con respecto a otros trabajos puede significar que este tipo de genes biosintéticos está en muy bajas proporciones en los ambientes terrestres debido al número muy bajo de clones positivos de un total de miles de clones como se demuestra en algunos trabajos previos.

### 8.2.1 PKS II

Los primers utilizados para la amplificación de genes PKS II fueron diseñados por Metsa- Ketela y colaboradores (1999) que utilizaron regiones consenso de este sistema para Actinobacterias. Inicialmente fueron usadas las concentraciones y condiciones de PCR reportadas por Andrielli en 2006, y Massini en 2009, para bibliotecas construidas con amplificados de este sistema, que constaban de insertos no mayores a 700 pb; sin embargo para la biblioteca de este estudio, con insertos cercanos a los 50 Kb no se logró amplificación, por lo que fue necesario modificar las concentraciones y condiciones de PCR, usando para su estandarización solamente con controles positivos para este sistema en bacterias del genero *Streptomyces*.

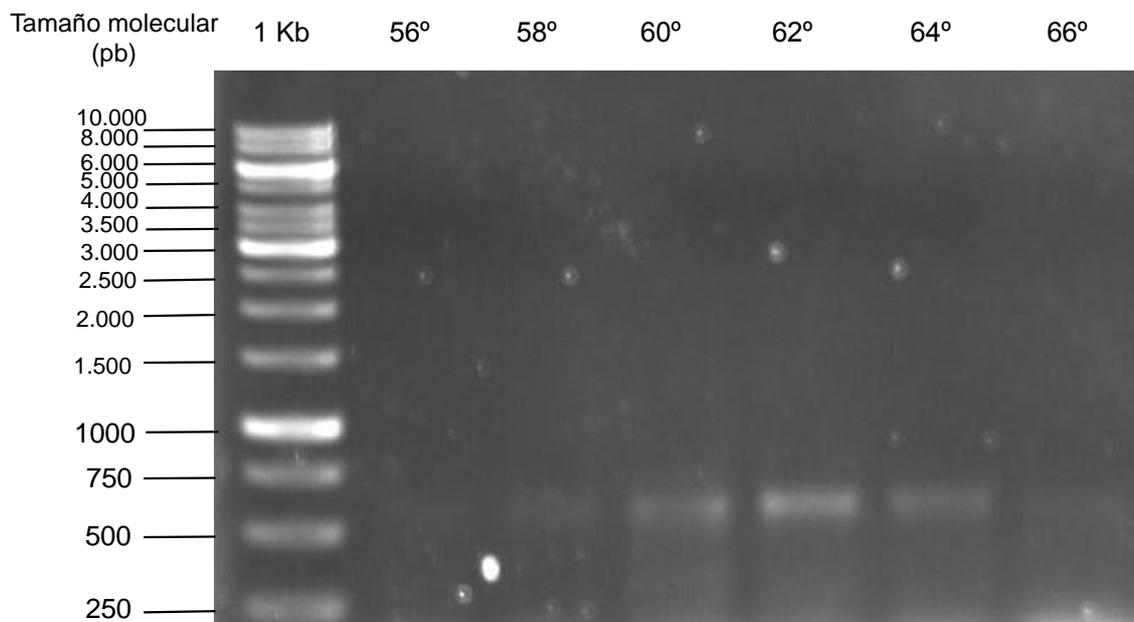
Se utilizaron los controles positivos *S. olindensis* y *S. coelicolor* para la estandarización por PCR. Se realizo un gradiente con 6 temperaturas de anillamiento

(Figura 20); la desnaturalización inicial (*hot start*) se modifico a 5 minutos, ademas el mix de reacción incluyo 5% de DMSO.

Tabla 8. Protocolo estandarizado para la detección de PKS tipo II en bibliotecas de insertos cercanos a 50 kb .

PROGRAMA TERMOCICLADO	
DESNATURALIZACIÓN INICIAL	95°C por 5 minutos
DESNATURALIZACIÓN CÍCLICA	95°C por 1 minuto
ANILLAMIENTO CÍCLICO DE LOS PRIMERS	anillamiento de 1 minuto con temperaturas entre 56° C a 64 °C
EXTENSIÓN CÍCLICA	1 minuto a 72°C
EXTENSIÓN FINAL	1 ciclo de extensión final de 8,5 minutos.
CONCENTRACIÓN DE LOS REACTIVOS	
Concentración inicial ( <i>Stock</i> )	Concentración final de reacción
DNTP's ( <i>Thermo Scientific 25mM</i> )	1 mM de dNTP' ( <i>Thermo scientific</i> )
Primers ( <i>Exxtend-100 mM</i> )	100 pM de cada <i>primer</i>
DNA Polimerasa ( <i>Thermo Scientific 5 U/μl</i> )	2 U/uL de <i>Taq</i> polimerasa ( <i>Thermo scientific</i> )
Buffer Para PCR ( <i>Thermo Scientific</i> )	10 X (20 mM Tris-HCL pH 8,4; 50 mM KCL)
MgCl <sub>2</sub> ( <i>Thermo Scientific</i> )	2,5 mM de MgCl
Volumen Final	25 μL
Ciclos De Repetición	30

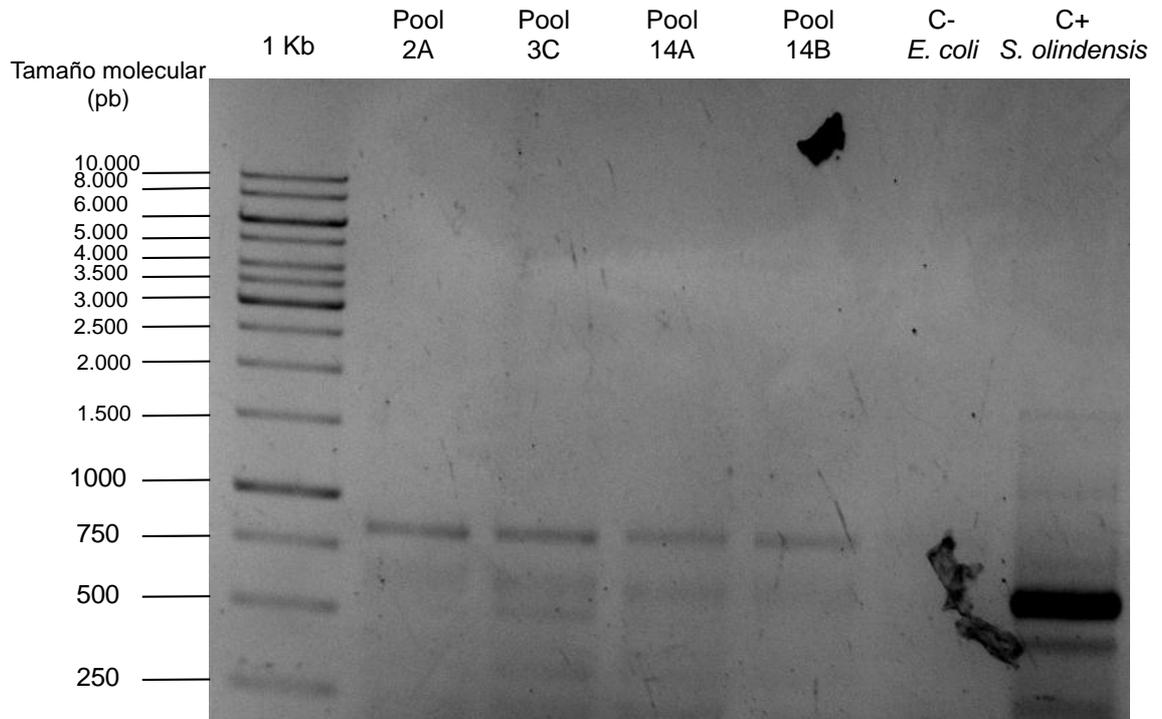
Figura 21. Amplificación PKS II *Streptomyces coelicolor*. 1. Marcador 1Kb (*Thermo Scientific*); 2-7. Gradiente Temperatura 56° a 66°. Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciences*) teñido en solución de 5µg/mL de bromuro de etidio.



En las condiciones presentadas ocurrió amplificación de DNA genómico de los controles positivos *S. coelicolor* y *S. olindensis*, se concluyó que la mejor temperatura de anillamiento es 62°C, temperatura en la cual la el DNA presenta mayor número de copias y por lo tanto óptima para evaluar los clones de la biblioteca metagenómica.

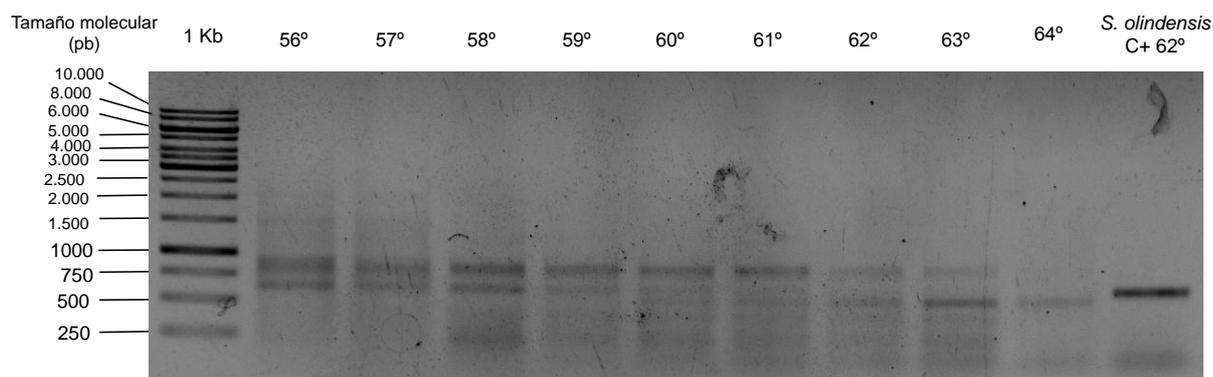
Se utilizaron las condiciones de PCR estandarizadas y optimizadas para *S. coelicolor* y *S. olindensis*. Para realizar el *screening* en la biblioteca se amplificó el DNA extraído en conjuntos de 96 clones. En el *screening* inicial fueron detectadas un total de tres placas de 96 clones positivas (2, 3, 14), la cual se dividió en tres partes formado *pools* de 32 clones (Figura 15), se obtuvieron un total de cuatro *pools* que en primera instancia se consideraron positivos (2A, 3C, 14A y 14B) (Figura 21).

Figura 22. Amplificación PKS II . 1. Marcador 1Kb (*Thermo scientific*); 2. Pool 2A; 3. Pool 3C; 4. Pool 14A; 5. Pool 14B; 6. Control negativo *E. coli* 7. Control Positivo *S. olindensis*. Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciences*) teñido en solución de 5µg/mL de bromuro de etídio.



Para mejorar la amplificación de la banda de 600 pb, e intentar eliminar inespecificidad en la amplificación, se realizó un gradiente con variante de 4°C a partir de la T<sub>m</sub> de los primers de PKS II (Figura 22).

Figura 23. Amplificación PKS II clon 2A4F. 1. Marcador 1Kb (*Thermo scientific*); 2-7. Gradiente Temperatura 56° a 64°C. Gel agarosa 0,8% p/v, corrido en buffer TAE 1X a 5v/cm<sup>2</sup>; en una cámara de electroforesis (*Amersham biosciencias*) teñido en solución de 5µg/mL de bromuro de etidio.



La menor cantidad de bandas se obtuvo en la temperatura de 58°C, para los análisis en clones individuales, esta temperatura fue establecida como padrón para realizar las amplificaciones en la siguiente parte de la biblioteca. De los *pools* 2A, 3C, 14A, y 14B que inicialmente se consideraron positivos, únicamente el pool 2A con la modificación de temperatura aplicada a los clones, concuerda con el tamaño molecular del control positivo *S. olindensis* y *S. avermitilis*. Se realizó el *screening* individual de este *pool* y se encontraron 3 clones positivos denominados 2A3F, 2A4E, 2A4F. Para realizar el proceso de secuenciación de los amplificados, se realizó la remoción de la banda con el tamaño molecular deseado, por medio del kit de extracción para gel de agarosa (illustra GFX PCR DNA and Gel Band Purification Kit).

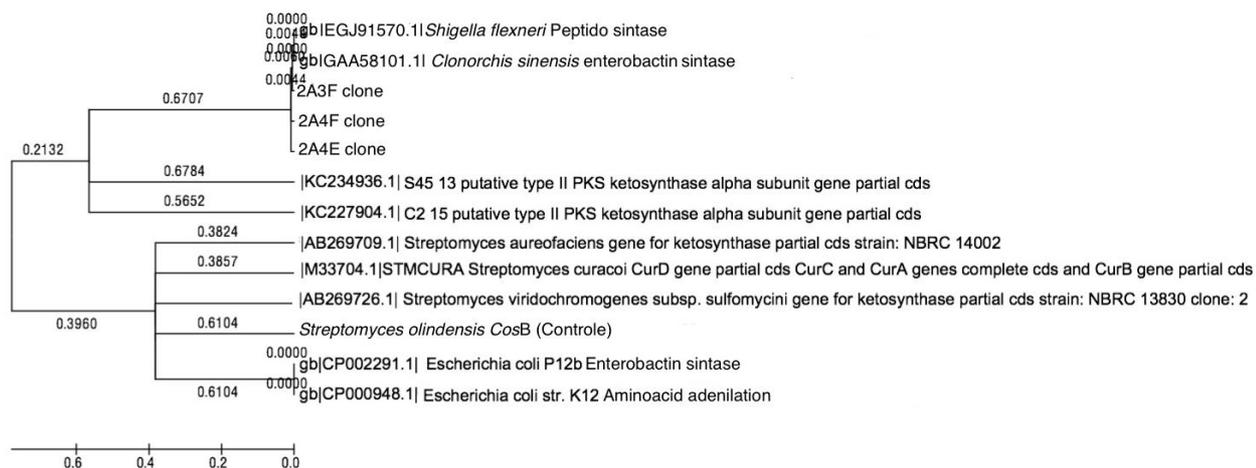
La secuencia control *S. olindensis* presentó una similaridad del 100% de una Cetosintasa *CosB* de *S. olindensis* por lo cual se demuestra que el control positivo está correcto. Las secuencias de los amplificados en los clones tienen una similaridad del 99% con el componente F de una enterobactin sintasa de *Clonorchis sinensis*, también presentaron una similaridad del 99% con una péptido sintetasa de *Shigella flexneri*. Las péptido sintetasa fueron descubiertas en Actinomicetos que incluye vías relacionadas con la biosíntesis de los grupos de antibióticos como: ansatrienina (Chen *et al.*, 1999) y los de la vancomicina, cloroeremomicina bleomicina (Shen *et al.*, 2001) y Estreptomicina (Fernández-Moreno *et al.*, 1997).

Estas enzimas pertenecen a un grupo de moléculas denominados péptidos no ribosomales y son catalizados por enzimas sintetetasas (NRPSs), estas moléculas generalmente presentan una gama extremadamente amplia de actividades biológicas, en ocasiones con aplicación farmacológicas (Turgay *et al.*, 1992 ; Stauton & Weissman, 2001) A menudo son: toxinas, sideróforos, pigmentos, antibióticos, citostáticos, inmunosupresores (la mayoría son de uso comercial). Los Péptidos no ribosomales, son una clase de moléculas, por lo general producidas por microorganismos y en ocasiones se encuentran en organismos superiores, tales como nudibranchios, pero se cree que están hechas por bacterias dentro de estos organismos (Neilands, 1995).

Los péptidos no ribosomales (NRPSs) y los policétidos (PKS) comparten el mismo patrón biosintético por medio de las enzimas sintetetasas y sintetasas (liasas), que constan de módulos con ciertos dominios obligatorios o núcleo (suplementario para la adición de cada unidad de péptidos o malonil CoA y un número variable de dominios opcionales y enzimas responsables de la modificación del péptido o la columna vertebral del policétido. El módulo de núcleo consiste, en un dominio de aciltransferasa (AT) para la selección y extensión de unidad de transferencia, una proteína portadora de acilo (ACP) y un extensor de la enzima cetoacil sintasa (KS) en el dominio que tiene la función de condensación descarboxilativas. Durante la biosíntesis, la cadena en crecimiento permanece unida de forma covalente a la enzima y al llegar a su longitud completa, una tioesterasa (TE) cataliza el dominio de liberación de los NRPS y PKS como productos. Los segmentos de la cadena polipeptídica de todos estos dominios se conocen como enlazadores y se ha demostrado que establecen la comunicación funcional entre y dentro de los módulos (Gokhale & Khosla, 2000; Du & Chen, 2001).

Para establecer cuál es la relación que tienen las secuencias obtenidas de los clones de origen ambiental con las obtenidas en las bases de datos, se realizó un análisis filogenético usando el parámetro de Neighbor joining:

Figura 24. Árbol de las secuencias genicas proveniente de los clones, control positivo (*S. olindensis*) comparacion con secuencias de las bases de datos. La escala indica el grado de similaridad. El árbol construido usando análisis Neighbor joining con el programa MEGA 5.



El árbol muestra que las secuencias provenientes de los clones, están relacionadas con la enterobactin sintasa de *Shigella flexneri* y la péptido sintetasa de *Clonorchis sinensis*. Las secuencias que codifican este tipo de enzimas se caracterizan por tener de 15 a 34 residuos de aminoácidos en su secuencia y un considerable número de lisinas o argininas (Piers *et al.*, 1993; Peschel & Sahl, 2006). Ciertos microorganismos llevan a cabo la síntesis de este tipo de péptidos realizando ensamblajes consecutivos de aminoácidos, a través de un complejo enzimático multifuncional tipo sintetasa (Catalizan la unión carbono-nitrogeno). La enzima contiene dominios, que tienen la capacidad de reconocer un residuo, activarlo, modificarlo si es necesario y agregarlo a la cadena peptídica naciente. Este mecanismo básico da como resultado una gran variedad de productos peptídicos y puede llegar a contener residuos inusuales L-D aminoácidos (Hancock, 1997; Oppermann-Sanio & Steinbüchel, 2002).

Debido a que las conformaciones moleculares que codifican estas secuencias de origen metagenómico, pueden variar con respecto a la secuencia control PKS II, se realizó un análisis bioinformático para evaluar si existen variaciones en su estructura, composición y conformación:

Figura 25. Composición de aminoácidos enzima control *S. olindensis* (aminoácidos: Anexo 3) Fuente: Predict Protein (<https://www.predictprotein.org/> acceso 5-09-2013).

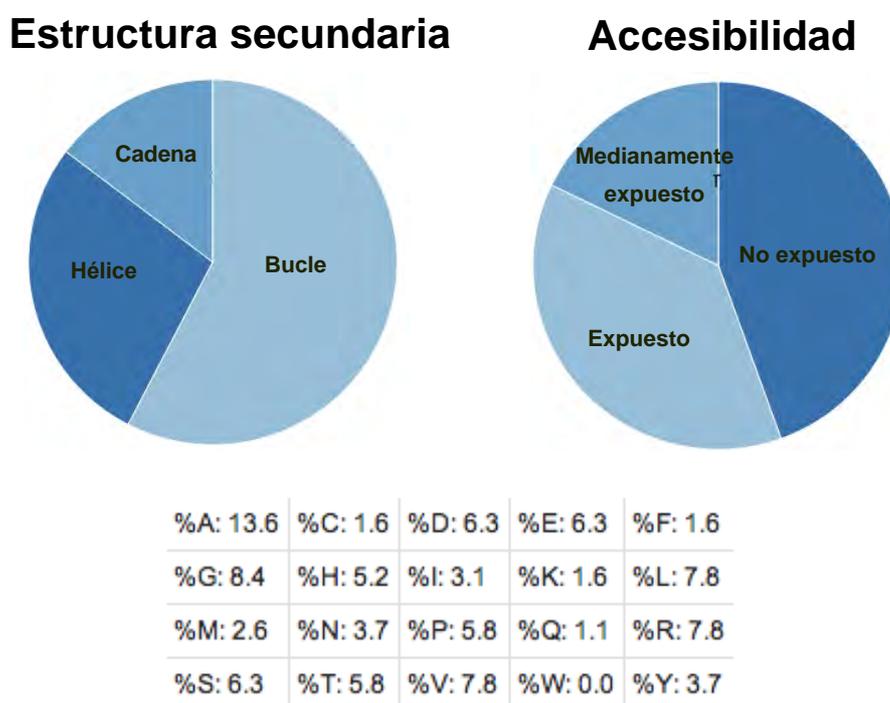
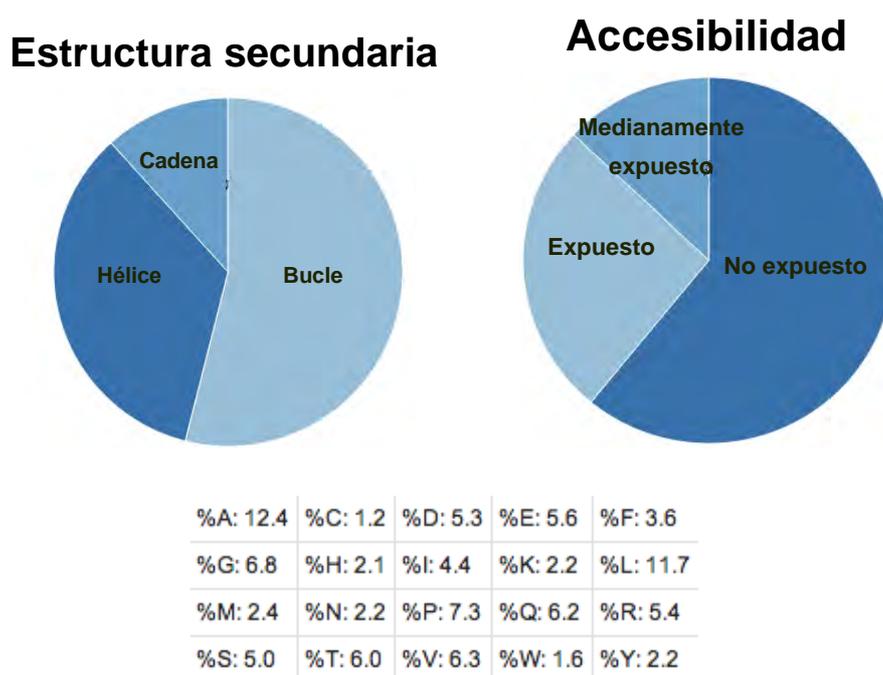


Figura 26. Composición de aminoácidos enzima de origen metagenómico (aminoácidos: Anexo 3) Fuente: Predict Protein (<https://www.predictprotein.org/> acceso 5-09-2013).



Las diferencias de estructura secundaria muestran que existe una variación importante en la solubilidad de la enzima de origen metagenómico, la mayor parte de la molécula es menos soluble con respecto a la cetosintasa del control positivo (*S. olindensis*), una de las probables razones es que la enzima de origen metagenómico posee 6,2% del aminoácido glutamina (Q), frente al 1,1% de la proteína control cetosintasa. Los aminoácidos Serina (S) y glutamina (Q) se comportan como donadores o aceptores de enlaces de hidrogeno; además, si los protones amida o los carbonilos del armazón polipeptídico no están implicados en el enlace peptídico, pueden interaccionar también en este tipo de uniones estabilizadoras (Mathews *et al*, 2002).

### 8.3 Modelamiento molecular y sitios de unión

En la Figura 27 y 28 se observa una clara diferencia en los sitios de unión de las proteínas, esto demuestra que la catálisis y el sustrato que se acopla a cada una de las enzimas es diferente, se puede deducir que efectivamente el gen secuenciado en este estudio codifica una enzima con variaciones importantes y que posiblemente puede estar relacionada en las rutas de síntesis de proteínas de origen no ribosomal como policetidos.

Figura 27. Sitios de unión en la proteína control cetosintasa

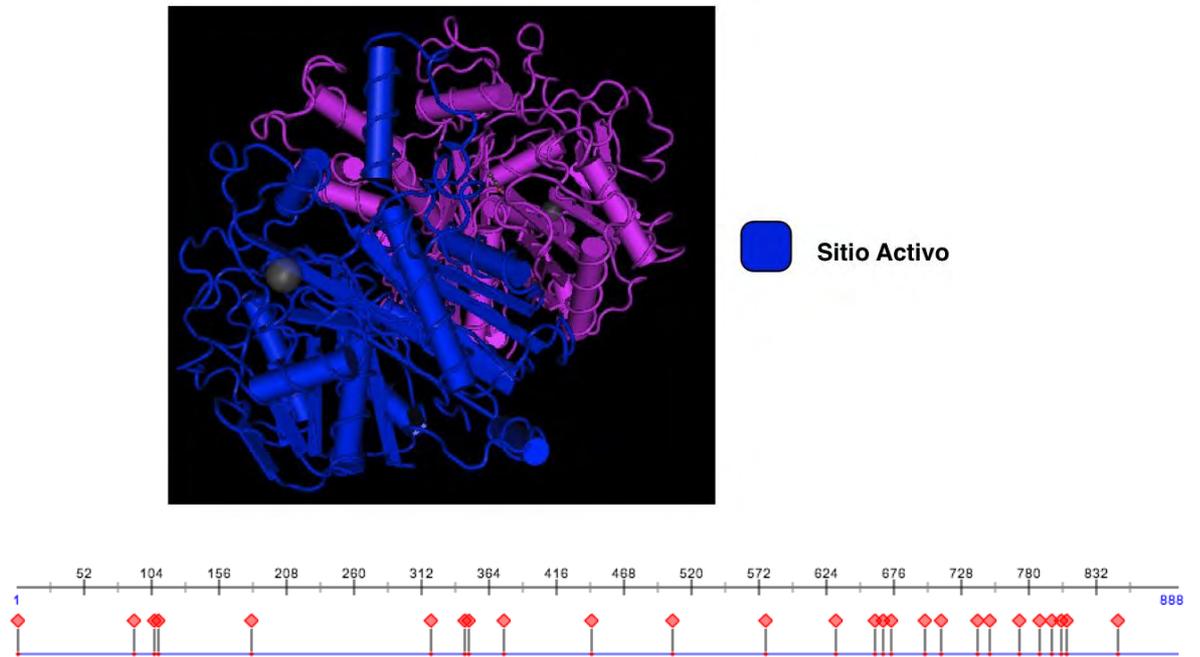
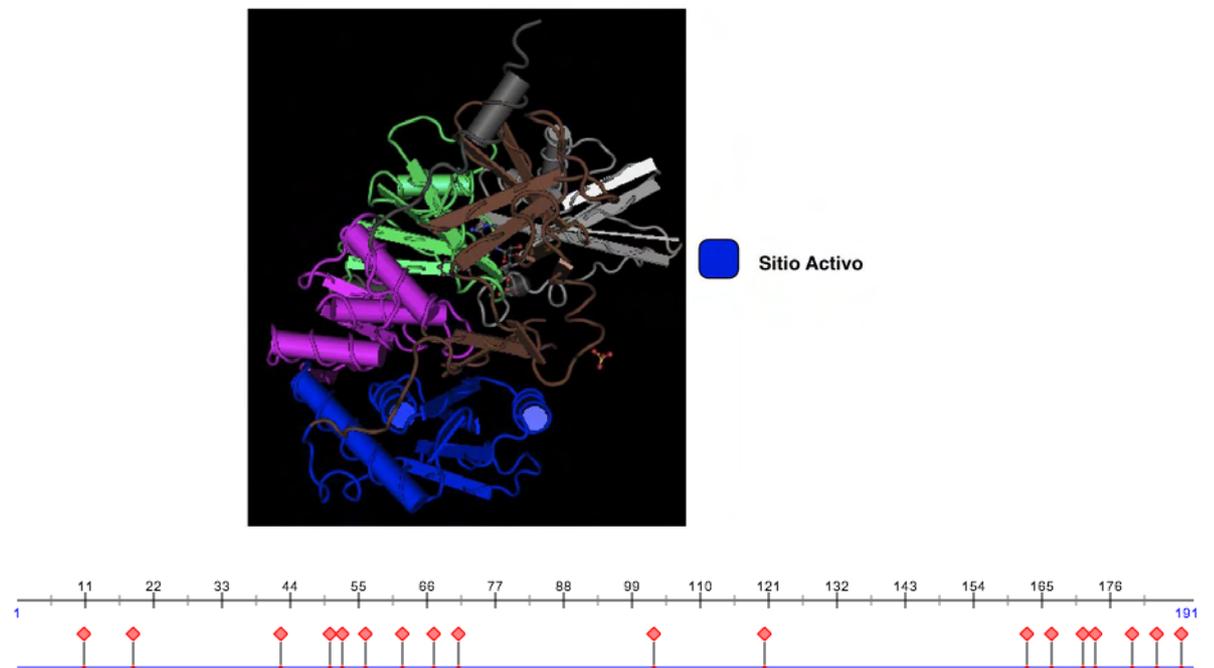
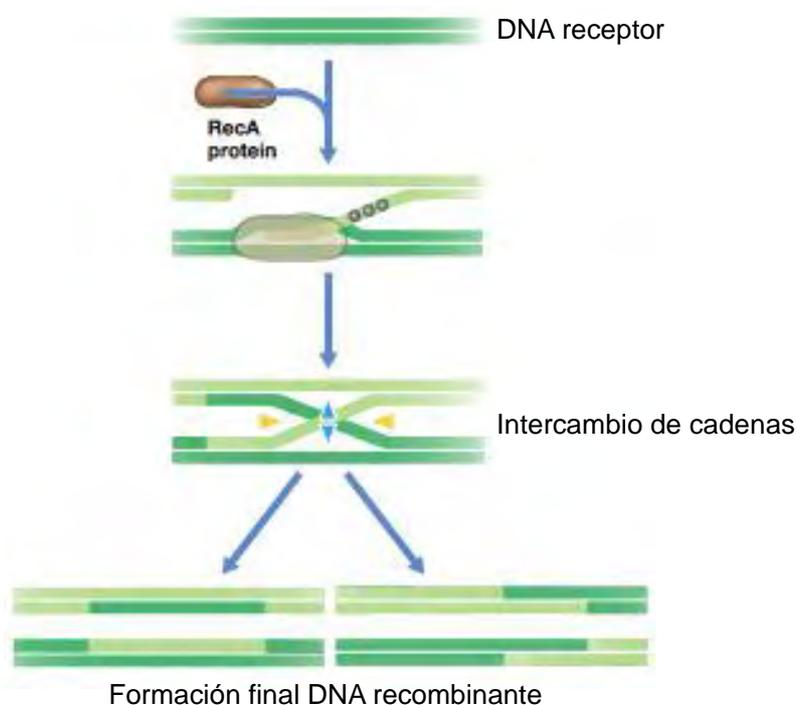


Figura 28. Sitios de unión en la proteína de origen metagenómico



Para conocer mejor cuales son las implicaciones a nivel bioquímico y de síntesis a partir del DNA de origen metagenómico, este puede ser expresado directamente en un vector y hospedero adecuado como *Streptomyces*, que permitiría caracterizar a profundidad el producto final; como también implementar estrategias de manipulación genética y biosíntesis combinatoria, para obtener metabolitos secundarios híbridos a partir de la clonación de genes biosintéticos funcionales de orígenes diferentes (Figura 29) (Hopwood, 2004); Esta técnica es de mucha utilidad en la síntesis de compuestos policetónicos análogos como los observados en *S. cellulosum* productor de ambruticuna (Julien *et al.*, 2006).

Figura 29. Recombinación homologa de fragmentos de origen diferente



Fuente: Madigan *et al.*, 2011.

La detección por PCR de sintasas tipo II, fue empleada recientemente para amplificar DNA de poblaciones mixtas obtenidas directamente de suelos (Wawrik *et al.*, 2005) En este estudio los autores revelan alta diversidad de estas enzimas, y presentan similitud a secuencias involucradas en la biosíntesis de antibióticos derivados de anguciclinas, como cimociclinona, pradimicina y jasomicina.

La estrategia de amplificación de genes biosintéticos como cetosinasas pertenecientes a sistemas de policétidos de tipo II, puede ser utilizada como una gran ventaja para la descubierta de nuevas moléculas, a partir de genes obtenidos directamente del ambiente como ocurre en los suelos.

Para las amplificaciones de PCR tipo I y II, fue estandarizado un protocolo dirigido a la detección de este sistema en DNA proveniente de poblaciones microbianas de origen ambiental. Así se demostró que las condiciones de amplificación empleadas en la PCR, fueron adecuadas para el rastreo de estos sistemas en insertos clonados en vectores como los BAC (cromosoma artificial bacteriano).

## 9. CONCLUSIONES

- La biblioteca metagenómica de suelo Atlántico Brasileño, constituida por 5000 clones, usando como vector pBAC oriV e insertos aproximados a los 50 kb, no presenta clones asociados al sistema PKS I, probablemente la pérdida de viabilidad (30%) fue un factor importante de ser considerado en este evento; además los estudios anteriores de *screening* de policétidos tipo I han reportado un número muy bajo de clones con respecto a la totalidad que componen la biblioteca; lo que demuestra que este tipo de genes se encuentra en una proporción baja en los ecosistemas terrestres.
- La biblioteca metagenómica proveniente de suelo Atlántico Brasileño presenta genes relacionados con la biosíntesis de policétidos del grupo de los péptidos no ribosomales, con variantes importantes con respecto a la cetosintasa de *S. olindensis*. Posiblemente estas variaciones en la secuencia de origen metagenómico forman parte de rutas de biosíntesis de otros metabolitos secundarios.

## 10. RECOMENDACIONES

- Construir una biblioteca con un vector de mayor replicación como cósmidos o fosmidos, para optimizar y agilizar la extracción de DNA plasmidial; debido a que los cromosomas artificiales generalmente presentan un numero bajo de copias por célula hospedera.
- La construcción de una biblioteca metagenómica debe ser realizada siguiendo un estricto protocolo de monitoreo permanente y sistemático de la viabilidad celular de los clones.
- Complementar el *Screening* de genes basado en amplificación por PCR, usando tecnologías de hibridación como Microarrays, Macroarrays o Southern blot.
- Caracterizar totalmente los clones positivos, para establecer si existen regiones próximas a las detectadas que estén asociadas con rutas de biosíntesis completas.

## 11. REFERENCIAS BIBLIOGRAFICAS

- Amann, R. I., Ludwig, W., & Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews*, 59(1), 143–169.
- Allison, L.A., (2009) Fundamental Molecular Biology Blackwell Publ ISBN-13: 978-1405103794
- Andrielli F. (2006) Diversidade de Bactérias e Amplificações por PCR de Genes de Sintase de Policetônicos (PKS) tipo II em solo de Mata Atlântica [Tesis maestría]. São Paulo: Instituto de Ciências Biomédicas.
- Angelov, A., Mientus, M., Liebl, S., & Liebl, W. (2009). A two-host fosmid system for functional screening of metagenomic libraries from extreme thermophiles. *Systematic and applied microbiology*, 32(3), 177–185.
- Austin, M. B., & Noel, J. P. (2003). The chalcone synthase superfamily of type III polyketide synthases. *Natural Product Reports*, 20(1), 79–110.
- Bonilla-Rosso, G., Souza, V., & Eguiarte, L. E. (2008). Metagenómica, Genómica Y Ecología Molecular: La Nueva Ecología En El Bicentenario De Darwin. *Tip Revista Especializada en Ciencias QuímicoBiológicas*, 11(1), 41–51.
- Borneman, J., & Trjplett, E. W. (1997). Rapid and direct method for extraction of RNA from soil. *Soil Biology and Biochemistry*, 29(11-12), 1621–1624. doi:10.1016/S0038-0717(97)00084-9
- Bronick, C. J., & Lal, R. (2005). Soil structure and management: a review. *Geoderma*, 124(1-2), 3–22. doi:10.1016/j.geoderma.2004.03.005
- Chen, S., von Bamberg, D., Hale, V., Breuer, M., Hardt, B., Müller, R., Leistner, E. (1999). Biosynthesis of ansatrienin (mycotrienin) and naphthomycin. Identification and analysis of two separate biosynthetic gene clusters in *Streptomyces collinus* Tü 1892. *European journal of biochemistry / FEBS*, 261, 98–107.
- Chen, K. & Pachter, L. (2005). «Bioinformatics for Whole-Genome Shotgun Sequencing of Microbial Communities». *PLoS Computational Biology* 1 (2): pp. e24. doi:10.1371/journal.pcbi.0010024.
- Chenu, C, Y Stotzky G. (2002). Interactions between Microorganisms and soil particles: an overview, p. 3-40. In P. M. Huang, Bollag, J.-M., Senesi, N. (ed.) Interactions between Soil Particles and Microorganisms: Impact on the Terrestrial Ecosystems. Wiley, New- York.

- Courtois, S., Cappellano, C. M., Ball, M., Francou, F.-X., Normand, P., Helynck, G., ... Pernodet, J.-L. (2003). Recombinant Environmental Libraries Provide Access to Microbial Diversity for Drug Discovery from Natural Products. *Applied and Environmental Microbiology*, 69(1), 49–55.
- Cowan, D., Meyer, Q., Stafford, W., Muyanga, S., Cameron, R., & Wittwer, P. (2005). Metagenomic gene discovery: past, present and future. *Trends in biotechnology*, 23(6), 321–329.
- Cultek. (2006). Tecnologia del DNA recombinante, 1–28. Disponible en: [www.cultek.com](http://www.cultek.com)
- Curtis, T. P. & W. T. Sloan (2005). "Microbiology. Exploring microbial diversity--a vast below." *Science* 309(5739): 1331-3.
- Daniel, R. (2005). The metagenomics of soil. *Nature Reviews Microbiology*, 3(6), 470–478.
- Demain, A. L. (1992). Microbial secondary metabolism: a new theoretical frontier for academia, a new opportunity for industry. *Ciba Foundation Symposium*, 171(March), 3–16; discussion 16–23.
- Du, L. & Shen, B. (2001) Biosynthesis of hybrid peptide–polyketide natural products. *Curr. Opin. Drug Discov. Devel.*, 4, 215–228.
- Fernández-Moreno, M. A., Vallín, C., & Malpartida, F. (1997). Streptothricin biosynthesis is catalyzed by enzymes related to nonribosomal peptide bond formation. *Journal of bacteriology*, 179, 6929–6936.
- Fu, C., Hu, Y., Xie, F., Guo, H., Ashforth, E. J., Polyak, S. W., ... Zhang, L. (2011). Molecular cloning and characterization of a new cold-active esterase from a deep-sea metagenomic library. *Applied Microbiology and Biotechnology*, 90(3), 961–970.
- Galindo, P. (2011). *Bioprospecção de genes relacionados à biossíntese de polímeros biodegradáveis a partir de uma biblioteca metagenômica de solo de Mata Atlântica*.
- Gillespie, D. E., Brady, S. F., Bettermann, A. D., Cianciotto, N. P., Liles, M. R., Rondon, M. R., Icrobiol, A. P. P. L. E. N. M. (2002). Isolation of Antibiotics Turbomycin A and B from a Metagenomic Library of Soil Microbial DNA. *Society*, 68(9), 4301–4306. doi:10.1128/AEM.68.9.4301
- Gokhale, R.S. & Khosla, C. (2000) Role of linkers in communication between protein modules. *Curr. Opin. Chem. Biol.*, 4, 22–27.
- Ginolhac, A., Jarrin, C., Gillet, B., Robe, P., Pujic, P., Tuphile, K., Nalin, R. (2004). Phylogenetic Analysis of Polyketide Synthase I Domains from Soil Metagenomic Libraries Allows Selection of Promising Clones. *Applied and Environmental Microbiology*, 70(9), 5522–5527.
- Griekspoor G., y Groothuis, T. (2006). software for edit sequences 4Peaks mekentosj.com.

- Green, E. D. (2001). Strategies for the systematic sequencing of complex genomes. *Nature Reviews Genetics*, 2(8), 573–583.
- Hancock, R. E. (1997). IP\_Peptide antibiotics\_Lancet. *Lancet*, 349, 418–422.
- Handelsman, J, Rondon, M. R., Brady, S. F., Clardy, J., & Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology*, 5(10), R245–R249.
- Handelsman, Jo. (2004). Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and Molecular Biology Reviews*, 68(4), 669–685.
- Hertweck, C. (2009). The biosynthetic logic of polyketide diversity. *Angewandte Chemie (International ed. in English)*, 48(26), 4688–716.
- Hertweck, C., Luzhetskyy, A., Rebets, Y., & Bechthold, A. (2007). Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Natural Product Reports*, 24(1), 162–190.
- Hopwood, D. A. (2004). Cracking the Polyketide Code. *PLoS Biology*, 2(2), E35.
- Hutchinson, C. R., & Fujii, I. (1995). Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics. *Annual Review of Microbiology*, 49, 201–238.
- Hugenholtz, P., & B. M. Goebel. (1998). "Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity." *J Bacteriol* 180(18): 4765- 74.
- Jeon, J. H., Kim, J.-T., Kim, Y. J., Kim, H.-K., Lee, H. S., Kang, S. G., ... Lee, J.-H. (2009). Cloning and characterization of a new cold-active lipase from a deep-sea sediment metagenome. *Applied Microbiology and Biotechnology*, 81(5), 865–874.
- Julien, B.; Tian, Z. Q.; Reid, R.; Reeves, C. D. Analysis of the ambruticin and jerangolid gene clusters of *Sorangium cellulosum* reveals unusual mechanisms of polyketide biosynthesis. *Chem. Biol.*, v. 13, p. 1277-1286, 2006.
- Keatinge-Clay, A. T., Maltby, D. A., Medzihradzsky, K. F., Khosla, C., & Stroud, R. M. (2004). An antibiotic factory caught in action. *Nature Structural & Molecular Biology*, 11(9), 888–893.
- Kirk, J. L.; Beaudette, L. A.; Hart, M.; Moutoglis, P.; Klironomos, J. N.; Lee, H.; Trevors, J. T. (2004) Methods of studying microbial diversity. *Journal of Microbiological Methods*, v. 58, p.169-188.
- Khosla, C., & Zawada, R. J. (1996). Generation of polyketide libraries via combinatorial biosynthesis. *Trends in biotechnology*, 14(9), 335–341.
- Kleinkauf, H., & Von Dohren, H. (1996) A nonribosomal system of peptide biosynthesis *European Journal of Biochemistry*, Oxford, vol.236,p.335-351.

- Kluyver, A.J. & Van Niel, C.B. (1956) *The Microbe's Contribution to Biology* (Harvard University Press, Cambridge, MA).
- Kim, B. S., Oh, H., Kim, S. Y., Park, J. A., Yoon, Y. J., Lee, S. K., Ahn, J. S. (2005). Isolation and identification of new anthracycline antibiotics, rubomycins F and H. *The Journal of antibiotics*, 58(3), 1185–1186.
- Ludwig, W., Amann, R.I.L., Schleifer, K.H. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiological Reviews*, Washington, v. 59, 143-169
- Lefevre, F., Robe, P., Jarrin, C., Ginolhac, A., Zago, C., Auriol, D., ... Nalin, R. (2008). Drugs from hidden bugs: their discovery via untapped resources. *Research in Microbiology*, 159(3), 153–161.
- Lorenz, P., Liebeton, K., Niehaus, F., & Eck, J. (2002). *Screening* for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Current Opinion in Biotechnology*, 13(6), 572–577.
- Michael T, M., Martinko M, J., Stahl, A.D., Clark, P.D. (2011) *Brock Biology of Microorganisms* Pearson Education, Inc. ISBN-13: 978-0-321-64963-8. 1155.p
- Mariano, A. (2009). *Herramientas para el mejoramiento genético de S violaceusniger YCED-9 enfocadas a la producción de metabolitos secundarios*. [Tesis maestría]. Instituto Politecnico Nacional Tlaxcala, Mexico.
- Massini, K. C. (2009). *Karen Cristina Massini Bioprospección de Genes Biosintéticos de Policetídicos em DNA Metagenômico de Solo de Mata Atlântica de Solo de Mata Atlântica* [Tesis Doctorado]. São Paulo: Instituto de Ciências Biomédicas.
- Metsä-Ketelä, M., Salo, V., Halo, L., Hautala, A., Hakala, J., Mäntsälä, P., & Ylihonko, K. (1999). An efficient approach for screening minimal PKS genes from *Streptomyces*. *FEMS Microbiology Letters*, 180(1), 1–6. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10547437>
- Mathews, C. K.; Van Holde, K.E et Ahern, K.G (2002). «6». *Bioquímica* (3ª edición). pp. 204. [ISBN 84-7892-053-2](#). 1110.p
- Neilands, J. B. (1995). Siderophores: structure and function of microbial iron transport compounds. *The Journal of biological chemistry*, 270, 26723–26726.
- Olsen, G. J., Woese, C. R., & Overbeek, R. (1994). The winds of (evolutionary) change: breathing new life into microbiology. *Journal Of Bacteriology*, 176(1), 1–6.
- Ofran Y., & Rost B. (2007) interaction sites identified from sequence." *Bioinformatics*. 23: e13-e16 .
- Oppermann-Sanio, F. B., & Steinbüchel, A. (2002). Occurrence, functions and biosynthesis of polyamides in microorganisms and biotechnological production. *Die Naturwissenschaften*, 89, 11–22.

- Osburne, M., Grossman, T., August, P., Macneil, I. (2002). Tapping into microbial diversity for natural products drug discovery. *ASM News.*; 66 (7): 411-417.
- Parachin, N. S., & Gorwa-Grauslund, M. F. (2011). Isolation of xylose isomerases by sequence- and function-based screening from a soil metagenomic library. *Biotechnology for biofuels*, 4(1), 9.
- Parsley, L., Linneman, J., Goode, A., Becklund, K., George I., Goodman, R., Lopanik, N & Liles, M (2010) Polyketide synthase pathways identified from a metagenomic library are derived from soil Acidobacteria *FEMS Microbiol Ecol*78(2011)176–187
- Park, S.-Y., Shin, H.-J., & Kim, G.-J. (2011). Screening and identification of a novel esterase EstPE from a metagenomic DNA library. *Journal of microbiology Seoul Korea*, 49(1), 7–14.
- Perfect, E., & Grove, J. H. (2002). Disruptive methods for assessing soil structure. *Soil and Tillage Research*, 64(1-2), 3–22.
- Peschel, A., & Sahl, H.-G. (2006). The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nature reviews. Microbiology*, 4, 529–536.
- Piers, K. L., Brown, M. H., & Hancock, R. E. (1993). Recombinant DNA procedures for producing small antimicrobial cationic peptides in bacteria. *Gene*, 134, 7–13.
- <https://www.predictprotein.org>
- Raes, J., Foerstner, K.U. & Bork, P.B (2007). Get the most out of your metagenome: computational analysis of environmental sequence data. *Current opinion in microbiology* 10(5), 490-498.
- Rees, H. C., Grant, S., Jones, B. E., Grant, W. D., & Heaphy, S. (2003). Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries. *Extremophiles life under extreme conditions*, 7(5), 415–421.
- Reeves, C. D. (2003). The enzymology of combinatorial biosynthesis. *Critical Reviews in Biotechnology*, 23(2), 95–147.
- Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., ... Goodman, R. M. (2000). Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Applied and Environmental Microbiology*, 66(6), 2541–2547.
- Robe P, Nalin R, Capellano C, Voget T. (2003) Extraction DNA from soil. *European Journal of Soil Biology*, v. 39, p.183-190.
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor laboratory press. New York* (pp. 931–957). Cold Spring Harbor.
- Silva LF, Gomez JGC, Oliveira MS, Alterthum F (1992). Freeze-drying of industrial yeast strains: influence of growth conditions, cooling rates and suspending media on

- the viability of the recovered cells. ISSN 1517- 8382 Brazilian Journal of Microbiology 23: 117 - 122.
- Schloss, P. D., & Handelsman, J. (2003). Biotechnological prospects from metagenomics. *Current Opinion in Biotechnology*, 14(3), 303–310.
- Schloss, P. D., & Handelsman, J. (2006). Toward a Census of Bacteria in Soil. (D. Relman, Ed.) *PLoS Computational Biology*, 2(7), 8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16848637>
- Shen, B., Du, L., Sanchez, C., Edwards, D. J., Chen, M., & Murrell, J. M. (2001). The biosynthetic gene cluster for the anticancer drug bleomycin from *Streptomyces verticillus* ATCC15003 as a model for hybrid peptide-polyketide natural product biosynthesis. *Journal of industrial microbiology & biotechnology*, 27, 378–385.
- Staunton, J., & Weissman, K. J. (2001). Polyketide biosynthesis: a millennium review. *Natural Product Reports*, 18(4), 380–416.
- Streit wolfgang R.& Daniel R (2010), *Metagenomics: Methods and Protocols*, Methods in Molecular Biology, vol. 668, DOI 10.1007/978-1-60761-823-2\_15, © Springer Science+Business Media, LLC.
- Strobel, G., & Daisy, B. (2003). Bioprospecting for Microbial Endophytes and Their Natural Products. *Microbiology and Molecular Biology Reviews*, 67(4), 491–502.
- Tamura K, Dudley J, Nei M and Kumar S MEGA 5 (2011). Molecular Evolutionary Genetics Analysis (MEGA) software version 5.0. *Molecular Biology and Evolution*. 24: 1596-1599
- Theron, J.; Cloete, T. E (2000). Molecular techniques for determining microbial diversity and community structure in natural environments. *Crit. Rev. Microbiol.*, v. 26, p. 37-57.
- Tisdall, J. M., & Oades, J. M. (1982). Organic matter and water-stable aggregates in soils. *Journal of Soil Science*, 33(2), 141–163.
- Tringe, S. G.; Rubin, E. M. (2005) Metagenomics: DNA sequencing of environmental samples. *Nature Reviews*, v.6, p. 805-814.
- Torsvik, V, Daae, F. L., Sandaa, R. A., & Øvreås, L. (1998). Novel techniques for analysing microbial diversity in natural and perturbed environments. *Journal of Biotechnology*, 64(1), 53–62.
- Torsvik, V, Goksøyr, J., & Daae, F. L. (1990). High diversity in DNA of soil bacteria. *Applied and Environmental Microbiology*, 56(3), 782–787.
- Torsvik, Vigdis, Øvreås, L., & Thingstad, T. F. (2002). Prokaryotic diversity--magnitude, dynamics, and controlling factors. *Science*, 296(5570), 1064–1066.

- Turgay, K., Krause, M., & Marahiel, M. A. (1992). Four homologous domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes. *Molecular microbiology*, *6*, 529–546.
- Van Elsas J.D., Speksnijder, A.J., Van Overbeek L.S (2008) A procedure for the metagenomics exploration of disease-suppressive soils *Journal of Microbiological Methods* *75* (2008) 515–522
- Watrous, J., Roach, P., Alexandrov, T., Heath, B. S., Yang, J. Y., Kersten, R. D., ... Dorrestein, P. C. (2012). Mass spectral molecular networking of living microbial colonies. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(26), E1743–52.
- Waugh, A. C. W., & Long, P. F. (2002). Prospects for generating new Antibiotics. *Science Progress*, *85*(Pt 1), 73–88.
- Wawrik, B., Kerkhof, L., Zylstra, G. J., & Kukor, J. J. (2005). Identification of unique type II polyketide synthase genes in soil. *Applied and environmental microbiology*, *71*, 2232–2238.
- Wild, J., Hradecna, Z., & Szybalski, W. (2002a). Conditionally amplifiable BACs: switching from single-copy to high-copy vectors and genomic clones. *Genome research*, (1), 1434–1444.
- Yadav, G., Gokhale, R. S., & Mohanty, D. (2003). Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *Journal of Molecular Biology*, *328*(2), 335–363.
- Zhou J, Bruns MA, Tiedje J. (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol.* *62*(2):316-322.

## 12. ANEXOS

### Anexo A

Protocolos de extracción de DNA usados por Massini 2004, para la construcción de la librería

Método de Extracción	Resultados de la Extracción	Método de Purificación	Test de amplificación
Utilización de PVPP + Bead-beater velocidad de. 4, 30 s, con perlas de vidrio 0,5 mM.	Fragmentos de DNA de todas las muestras de suelo entre ~15 a 20kb.	Las muestras presentaron coloración marrón y no fueron purificadas.	
Lisozima + bead-beater: velocidad de 5, 30 s, perlas de vidrio 1,0 mM.	Fragmentos de DNA entre 6 y 20 kb, aproximadamente.	Las muestras presentaron coloración marrón y no fueron purificadas	
PVPP + Lisozima + bead-beater velocidad de 4, 20 s, perlas de vidrio 1,0 mM.	DNA de menor calidad que los métodos anteriores	Recuperación de DNA a partir de escisión de la banda en gel de agarosa 0,8%.	
CT AB + proteinasa K+ bead-beater o shaker.	Poca cantidad de DNA, fue mayor en la extracción con shaker, sin embargo se degradó	Las muestras no fueron purificadas.	
Lavado con solución y tampón fosfato+ CTAB + proteinasa K (shaker) protocolo según Zhou	El DNA estaba degradado		

Protocolo según Zhou + Nycodens	DNA de bajo tamaño molecular entre ~15 y 20kb.	Las muestras presentaron coloración marrón y no fueron purificadas	
Método de Extracción	Resultados de la Extracción	Método de Purificación	Test de amplificación
Protocolo modificado de Zhou	DNA ~25kb.	El DNA fue purificado utilizado el kit GFX PCR: <i>DNA and gel band purification</i> (Invitrogen). El DNA de bajo tamaño molecular se fragmentó	No fueron amplificadas
Protocolo modificado según Zhou	DNA ~25kb.	Fue realizado el protocolo de electroelución en gel de agarosa. Poca cantidad de DNA y algunas muestras se degradaron.	No fueron amplificadas
Protocolo según Zhou modificado	DNA ~25kb.	Dos purificaciones sucesivas, siendo una con protocolo de electroelución y modificación en el proceso de precipitación de DNA.	No fueron amplificadas
Protocolo según Zhou modificado + adición de guanidina 0,01M en el tampón de extracción .	DNA en gran cantidad. DNA con fragmentos de 30kb.	Dos purificaciones sucesivas, siendo una con protocolo de electroelución y modificación en el proceso de precipitación de DNA. El DNA fue	Las muestras fueron digeridas con <i>Hind III</i> . No hubo digestión

		purificado utilizado el kit GFX PCR: <i>DNA and gel band purification</i> (Invitrogen).	
Método de Extracción	Resultados de la Extracción	Método de Purificación	Test de amplificación
Protocolo según Zhou con modificaciones en el proceso de precipitación de DNA	Ocurrió extracción de DNA en gran cantidad. DNA con fragmentos de 30kb.	Fue realizado el protocolo de electroelución en gel de agarosa. El DNA de bajo tamaño molecular se fragmentó	Las muestras fueron digeridas con la enzima <i>BamHI</i> . No hubo digestión
Gradiente de nycodenz	No hubo extracción de DNA		
Extracción de DNA utilizando el kit Ultra clean Mega Soil DNA de MoBio	DNA de baja calidad		Fueron realizadas digestiones: solamente con la enzima <i>BamHI</i> y con la enzima <i>EcoRI</i>  No hubo digestión con <i>EcoRI</i>
Extracción de DNA utilizando el kit Ultra clean Mega Soil DNA de MoBio con modificaciones en los pasos de lavado y paso por la columna (2X)	DNA con coloración más clara. DNA con tamaño de ~ 50kb.		Amplificado con primers 16S DNAr y hubo digestión con la enzima <i>HindIII</i> .

Fuente: Adaptado de Massini, 2009

## Anexo B

Protocolo de *miniprep* de DNA plasmídico usando el protocolo modificado de lisis alcalina

Crear la bacteria de interés, en 2 a 5 ml medio LB con el antibiótico de selección por toda la noche con agitación a 37°C.

1. Transferir la solución de bacterias para un microtubo, centrifugar en microcentrífuga (13000-16000) por 10 minutos, y descartar la solución sobrenadante.
2. Resuspender el *pellet* en 200 µl de la Solución I, (vortex). Mantener por 10 minutos a temperatura ambiente.
3. Adicionar 200 µl de la Solución II. Mezclar por inversión. Incubar en hielo por 5 minutos.
4. Adicionar 150 µl de la solución III. Mezclar por inversión. Incubar en hielo por 5 minutos.
5. Microcentrifugar a 4°C por 5-10 minutos (13000-16000) (hasta que el *pellet* quede sólido).
6. Tomar la solución sobrenadante (contiendo el DNA plasmidial) para otro microtubo Eppendorf.
7. Adicionar 500 µl de fenol: cloroformo: alcohol isoamílico (25:24:1), (vortex), y centrifugar por 5 minutos (13000-16000) 4°C .
8. Colectar la fase superior acuosa y transferir para un tubo de microcentrífuga nuevo.
9. Adicionar al tubo 500 µl de cloroformo: alcohol isoamílico (24:1), mezclar con vortex, y centrifugar (13000-16000) 4°C por 10 minutos. Colectar la fase acuosa superior e transferir para un tubo de microcentrífuga.
10. Adicionar 1 ml de etanol absoluto (100%), Mezclar por inversión, e incubar a -20°C por lo menos 1 hora.
11. Centrifugar el DNA por 5 minutos en una microcentrífuga (13000-16000) 4°C. Decantar el etanol y adicionar 500 µl de etanol 70%. (Vortex) y centrifugar por 5 minutos. Decantar el etanol y secar el *pellet* al aire, desecador o concentrador
12. Resuspender el DNA en 50 µl de 1 mM Tris-HCl, 0,1 mM Na<sub>2</sub>-EDTA (TE), o Agua miliq con 50 µg/ml RNAsa A.

## SOLUCIONES

### Solución I:

Componente	Concentración Final
Glucosa	50 mM glucosa
Tris-HCl 1 M, pH 8,0	25 mM Tris-HCl, pH 8,0
EDTA, 0,5 M, pH 8,0	10 mM Na <sub>2</sub> ·EDTA
Água mQ destilada	
Autoclavar por 20 minutos a 121°C	
Almacenar a 4°C	

### Solución II:

Componente	Concentración Final
NaOH 1N	0,2 M NaOH
SDS 10% p/v	1,0% SDS
Agua mQ destilada	
Solución de uso inmediato; <u>no autoclavar</u>	

### Solución III:

3 M ácido acético, pH 5,5  
10 mM Tris-HCl, pH 8,0  
  
1 mM Na<sub>2</sub>·EDTA, pH 8,0

Autoclavar

Adicionar RNAsa A concentración final de 50 µg/ml.

RNAsa:

Tradicional (Sambrook et al., 1989)

RNAse pancreática (RNAse A) comercial debe ser disuelta a una concentración de 10 mg ml<sup>-1</sup> en acetato de sodio 0,01 M (pH 5,2), seguido de incremento de temperatura a 100°C por 15 minutos. Después bajar lentamente la temperatura, ajustar el pH con 0,1 volúmenes de 1 M Tris-HCl pH 7,4. Mantener congelado el *stock* a -20°C.

RNAse A Sigma (R6513)

Disolver la RNAsa a una concentración de 10 mg ml<sup>-1</sup> en 10 mM Tris-HCl, pH 7,5, 15 mM NaCl.

MEDIOS DE CULTIVO

LB (L broth)

10 g Bactotripton

10 g NaCl

5 g extracto de levadura

Completar el volumen para 1 litro con agua destilada. Ajustar el pH para 7,0 con NaOH.

## Anexo C

Secuencia (Query) (predict protein database (<https://www.predictprotein.org/> acceso 5-09-2013).

### Cetosintase

```
>query
RATSPNNDP EHASRPDAH RDGFVMEGS AVLVLEELEH ARARGAHVYC
EIRGYATFGN AYHMTGLTSE GLEMARAIDT ALGQARLDPS LIDYVNAHGS
GTRQNRHET AAVKRSLGAH AYDTPMSSIK SMVGHSLGAI GAIEVVACVL
ALAHRVVPT ANYETPDPEC DLDYVPRAR PRKLANVLSV G
```

### Secuencia origen metagenómico

```
>query
MSIEILANKQ RYDEPTLIQH AERLKMLIAQ FAADPALLCG DVDIMLPGEY
AQLAQINATQ VEIPETTLA LVAEQAAKTP DAPALADARY LFSYREMHEQ
VVALANLLRE RGVKPGDSVA VALPRSVFLT LALHAIVEAG AAWLPLDTGY
PDDRKMMLE DARPSLLITT DDQLPRFSDV PNLTNLCYNA PLTPQGSAPL
QLSQPHHTAY IIFTSGSTGR PKGVMVGQTA IVNRLLWMQN HYPLTGEDVV
AQKTPCSFDV SVWEFFWPF I AGAKLVMAEP EAHRDPLAMQ QFFAEYGVTT
THFVPSMLAA FVASLTPQTA RQSCVTLKQV FCSGEALPAD LCREWQQLTG
APLHNLYGPT EAAVDVSWYP AFGEELAQVR GSSVPIGYPV WNTGLRILDA
MMHPVPPGVA GDLYLTGIQL AQGYLGRPDL TASRFIADPF APGERMYRTG
DVARWLDNGA VEYLGRSDDQ LKIRGQRIEL GEIDRVMQAL PDVKQAVTHA
CVINQAAATG GDARQLVGYL VSQSGPLDTSALQAQLRET LPPHMVPVVL
LQLPQLPLSA NGKLDKALP LPELKAQAPG RAPKAGSETI IAAAFASLLG
CDVQDADADF FALGGHSLA MKLAAQLSRQ FARQVTPGQV MVAHVAKLA
TIIDGEESS RRMGFETILP LREGNGPTLF CFHPASGFAW QFSVLSRYLD
PQWSIIGIQS PRPHGPMQTA TNLDEVCEAH LATLLEQQPH GIAARLRARG
EQVAFLGLLD TWPPETQNWQ EKEANGLDPE VLAEINRERE AFLAAQQGST
STELFTTIEG NYADAVRLLT TAHSVPFDGK ATLFVAERTL QEGMSPERAW
SPWIAELDIY RQDCAHVDII SPGAFVKIGP IIRATLNR
```

PSI BLAST (Predict protein database (<https://www.predictprotein.org/> acceso 5-09-2013)).

Secuencia Cetosintasa *S. olindensis*

```

-----
--- PSI-BLAST multiple sequence alignment
-----
---
--- PSI-BLAST ALIGNMENT HEADER: ABBREVIATIONS FOR SUMMARY
--- SEQLength      : 191
--- ID             : identifier of aligned (homologous) protein
--- LSEQ2          : length of aligned sequence
--- IDE            : percentage of pairwise sequence identity
--- SIM            : percentage of similarity
--- LALI           : number of residues aligned
--- LGAP           : number of residues in all indels
--- BSCORE         : blast score (bits)
--- BEXPECT        : blast expectation value
--- OMIM           : OMIM (Online Mendelian Inheritance in Man) ID
--- PROTEIN        : one-line description of aligned protein
--- '!'            : indicates lower scoring alignment that is combined
---                  with the higher scoring adjacent one
---
ID             LSEQ2  IDE   SIM  LALI  LGAP  BSCORE  BEXPECT  PROTEIN
trIA1IJL2IA1IJL2_9ACTO      191   71   82   191   0     340     2e-91  Ketosynthase (Fragment) 0
trIA1IJR0IA1IJR0_STRVR     191  100  100   191   0     337     1e-90  Ketosynthase (Fragment) 0
trIA1IJR4IA1IJR4_STRVR     191   90   94   191   0     335     8e-90  Ketosynthase (Fragment) 0
trIA3DDY7IA3DDY7_CLOTH     411   51   63  190   5     358     8e-97  3-oxoacyl-[acyl-carrier-p
trIA3KW01IA3KW01_PSEAI     404   52   63   191   2     344     2e-92  3-oxoacyl-[acyl-carrier-p
trIA3LBT8IA3LBT8_PSEAI     390   52   63   191   2     343     2e-92  Beta-ketoacyl-acyl carrie
trIA4IL61IA4IL61_GEOTN     412   55   67  189   3     357     2e-96  3-oxoacyl-[acyl-carrier-p
trIA4J684IA4J684_DESRM     413   53   67   191   2     356     2e-96  3-oxoacyl-[acyl-carrier-p
trIA4SMJ7IA4SMJ7_AERS4     413   50   63   191   2     340     2e-91  3-oxoacyl-[acyl-carrier-p
trIA4VMR9IA4VMR9_PSEU5     402   50   64   191   2     337     1e-90  3-oxoacyl-[acyl-carrier-p
trIA5D1H7IA5D1H7_PELTS     413   50   66   191   2     352     5e-95  3-oxoacyl-[acyl-carrier-p
trIA5GF64IA5GF64_GEOUR     410   51   64   191   2     349     3e-94  3-oxoacyl-[acyl-carrier-p
trIA5JMD7IA5JMD7_9ACTO     203   71   82   191   0     337     1e-90  Type II PKS (Fragment) 0S
trIA5UZ92IA5UZ92_ROSS1     479   50   65   191   2     344     1e-92  Beta-ketoacyl synthase 0S

```

tr B4T329 B4T329_SALNS	413	48	62	191	3	335	8e-90	3-oxoacyl-[acyl-carrier-p
tr B4TFH2 B4TFH2_SALHS	413	48	62	191	3	335	8e-90	3-oxoacyl-[acyl-carrier-p
tr B4VT70 B4VT70_9CYAN	418	53	64	191	2	358	1e-96	3-oxoacyl-[acyl-carrier-p
tr B4WHJ7 B4WHJ7_9SYNE	435	55	67	191	2	341	1e-91	3-oxoacyl-[acyl-carrier-p
tr B5E8U2 B5E8U2_GEOBB	410	50	64	191	2	346	3e-93	3-oxoacyl-[acyl-carrier-p
tr B5EJD3 B5EJD3_ACIF5	413	49	59	191	2	339	4e-91	3-oxoacyl-[acyl-carrier-p
tr B5F910 B5F910_SALA4	413	48	62	191	3	335	8e-90	3-oxoacyl-[acyl-carrier-p
tr B5FKB0 B5FKB0_SALDC	413	48	62	191	3	335	8e-90	3-oxoacyl-[acyl-carrier-p
tr B5UHR6 B5UHR6_BACCE	412	52	66	190	5	346	4e-93	3-oxoacyl-[acyl-carrier-p
tr B5VB58 B5VB58_BACCE	412	52	66	190	5	345	6e-93	3-oxoacyl-[acyl-carrier-p
tr B5YDQ4 B5YDQ4_DICT6	417	49	65	191	2	351	7e-95	3-oxoacyl-[acyl-carrier-p
tr B5YHD3 B5YHD3_THEYD	415	51	65	191	2	361	8e-98	3-oxoacyl-[acyl-carrier-p
tr B6SEF1 B6SEF1_9ACTO	421	70	80	191	0	335	5e-90	AlnL ketoacyl synthase al
tr B7A836 B7A836_THEAQ	408	51	63	190	5	343	2e-92	3-oxoacyl-[acyl-carrier-p
tr B7GLZ5 B7GLZ5_ANOFW	414	53	66	190	5	354	1e-95	3-oxoacyl-[acyl-carrier-p
tr B7HGT5 B7HGT5_BACC4	412	52	66	190	5	347	2e-93	3-oxoacyl-[acyl-carrier-p
tr B7HZV1 B7HZV1_BACC7	412	52	66	190	5	345	6e-93	3-oxoacyl-[acyl-carrier-p
tr B7ILB9 B7ILB9_BACC2	412	52	66	190	5	348	9e-94	3-oxoacyl-[acyl-carrier-p
tr B7JC11 B7JC11_ACIF2	413	49	59	191	2	339	4e-91	3-oxoacyl-[acyl-carrier-p
tr B7JDY0 B7JDY0_BACC0	412	52	66	190	5	346	4e-93	3-oxoacyl-[acyl-carrier-p
tr B7V164 B7V164_PSEA8	414	52	63	191	2	343	3e-92	3-oxoacyl-[acyl-carrier-p
tr B8DFA9 B8DFA9_LISMH	413	46	61	190	3	342	7e-92	3-oxoacyl-[acyl-carrier-p
tr B8E198 B8E198_DICTD	413	50	66	191	2	353	4e-95	3-oxoacyl-[acyl-carrier-p
tr B8GSL0 B8GSL0_THISH	412	53	65	191	2	348	1e-93	3-oxoacyl-[acyl-carrier-p
tr B8KZV4 B8KZV4_9GAMM	420	48	64	189	0	348	7e-94	3-oxoacyl-[acyl-carrier-p

## Secuencia de origen metagenómico

ID	LSEQ2	IDE	SIM	LALI	LGAP	BSCORE	BEXPECT	PROTEIN
sp P11454 IENTF_ECOLI	1293	96	97	888	16	936	0.0	Enterobactin synthase com
sp P29698 IENTF_SHIFL	1281	99	100	888	0	972	0.0	Enterobactin synthase com
sp Q8XBV9 IENTF_ECOS7	1293	96	97	888	16	940	0.0	Enterobactin synthase com
tr A1A8L7 A1A8L7_ECOK1	1339	95	97	888	16	942	0.0	Enterobactin synthase mul
tr A7ZIX0 A7ZIX0_EC024	1293	96	97	888	16	939	0.0	Enterobactin synthetase c
tr A7ZXL2 A7ZXL2_ECOHS	1293	96	97	888	16	936	0.0	Enterobactin synthetase c
tr A8AJM1 A8AJM1_CITK8	1294	81	88	888	17	934	0.0	Uncharacterized protein 0
tr B1EKE6 B1EKE6_9ESCH	1293	91	94	888	16	923	0.0	Enterobactin synthetase c
tr B1IYM0 B1IYM0_ECOLC	1293	96	97	888	16	938	0.0	Amino acid adenylation do
tr B1LKH9 B1LKH9_ECOSM	1293	95	97	888	16	936	0.0	Enterobactin synthetase c
tr B1X5Y3 B1X5Y3_ECODH	1293	96	97	888	16	936	0.0	Enterobactin synthase mul
tr B2J682 B2J682_NOSP7	4389	30	47	850	55	924	0.0	Amino acid adenylation do
tr B2N8M8 B2N8M8_ECOLX	1293	96	97	888	16	941	0.0	Enterobactin synthetase c
tr B2NR91 B2NR91_ECOS7	1293	96	97	888	16	940	0.0	Enterobactin synthetase c
tr B2P1Y0 B2P1Y0_ECOS7	1293	96	97	888	16	940	0.0	Enterobactin synthetase c
tr B2PFT2 B2PFT2_ECOS7	1293	96	97	888	16	940	0.0	Enterobactin synthetase c
tr B2TTE5 B2TTE5_SHIB3	1293	95	96	888	16	932	0.0	Enterobactin synthetase c
tr B3A4V2 B3A4V2_ECOS7	1293	96	97	888	16	940	0.0	Enterobactin synthetase c
tr B3ARW6 B3ARW6_ECOS7	1293	96	97	888	16	940	0.0	Enterobactin synthetase c
tr B3AYC1 B3AYC1_ECOS7	1293	96	97	888	16	940	0.0	Enterobactin synthetase c
tr B3BFE6 B3BFE6_ECOS7	1293	96	97	888	16	940	0.0	Enterobactin synthetase c
tr B3BTF0 B3BTF0_ECOS7	1293	96	97	888	16	940	0.0	Enterobactin synthetase c
tr B3HEG2 B3HEG2_ECOLX	1293	97	97	888	16	939	0.0	Enterobactin synthetase c

trIH5IWU8IH5IWU8_ECOLX	1293	96	97	888	16	936	0.0	Amino acid adenylation do
trIH5JEF1IH5JEF1_ECOLX	1293	96	97	888	16	936	0.0	Enterobactin synthase com
trIH5JUR8IH5JUR8_ECOLX	1293	96	97	888	16	936	0.0	Amino acid adenylation do
trIH5K9Z8IH5K9Z8_ECOLX	1293	96	97	888	16	936	0.0	Amino acid adenylation do
trIH5KRF4IH5KRF4_ECOLX	1293	96	97	888	16	941	0.0	Amino acid adenylation do
trIH5L4F4IH5L4F4_ECOLX	1293	96	97	888	16	941	0.0	Amino acid adenylation do
trIH5LHU7IH5LHU7_ECOLX	1293	96	97	888	16	941	0.0	Amino acid adenylation do
trIH5LXX4IH5LXX4_ECOLX	1293	96	97	888	16	941	0.0	Amino acid adenylation do
trIH5MBM3IH5MBM3_ECOLX	1293	96	97	888	16	941	0.0	Amino acid adenylation do
trIH5MR31IH5MR31_ECOLX	1293	97	97	888	16	941	0.0	Enterobactin synthase com
trIH5N6A3IH5N6A3_ECOLX	1293	96	97	888	16	940	0.0	Amino acid adenylation do
trIH5NL55IH5NL55_ECOLX	1293	96	97	888	16	940	0.0	Amino acid adenylation do
trIH5P0T4IH5P0T4_ECOLX	1293	96	97	888	16	940	0.0	Amino acid adenylation do
trIH5PFU8IH5PFU8_ECOLX	1293	97	97	888	16	939	0.0	Amino acid adenylation do
trIH5PV92IH5PV92_ECOLX	1293	97	97	888	16	939	0.0	Amino acid adenylation do
trIH5QAD9IH5QAD9_ECOLX	1293	97	97	888	16	939	0.0	Amino acid adenylation do
trIH5QQ35IH5QQ35_ECOLX	1293	97	97	888	16	939	0.0	Amino acid adenylation do
trIH5R5P4IH5R5P4_ECOLX	1293	97	97	888	16	939	0.0	Amino acid adenylation do
trIH6M8J5IH6M8J5_ECOLX	1293	96	97	888	16	941	0.0	Enterobactin synthase sub
trIH8D659IH8D659_ECOLX	1293	95	97	888	16	941	0.0	Enterobactin synthase sub
trIH9UPK7IH9UPK7_ECOLX	1293	96	97	888	16	934	0.0	Enterobactin synthetase c
trII0MTX6II0MTX6_SALET	1294	77	85	888	17	919	0.0	Enterobactin synthase sub
trII0VAF0II0VAF0_SHIFL	1297	98	98	888	16	941	0.0	Enterobactin synthase sub
trII0VQB1II0VQB1_ECOLX	1278	97	97	887	16	932	0.0	Enterobactin synthase mul

Anexo D Código genético

<b>Aminoácido</b>	<b>Abreviatura (3 letras)</b>	<b>Abreviatura (1letra)</b>	<b>Codones</b>
Acido Aspártico	Asp	D	GAC GAU
Acido Glutámico	Glu	E	GAA GAG
Arginina	Arg	R	CGA CGC CGG CGU AGA AGG
Lisina	Lys	K	AAA AAG
Asparagina	Asn	N	AAC AAU
Histidina	His	H	CAC CAU
Glutamina	Gln	Q	CAA CAG
Serina	Ser	S	UCA UCC UCG UCU AGC AGU
Treonina	Thr	T	ACA ACC ACG ACU
Alanina	Ala	A	GCA GCC GCG GCU
Glicina	Gly	G	GGA GGC GGG GGU
Valina	Val	V	GUA GUC GUG GUU
Prolina	Pro	P	CCA CCC CCG CCU
Leucina	Leu	L	CUA CUC CUG CUU UUA UUG
Fenilalanina	Phe	F	UUC UUU
Tirosina	Tyr	Y	UAC UAU
Isoleucina	Ile	I	Aua auc auu
Metionina	Met	M	AUG
Triptofano	Trp	W	UGG
Cisteína	Cys	C	UGC UGU
Terminación			UAA UAG UGA

Fuente: (Mathews *et al*, 2002).