



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**Aislamiento y caracterización de un nuevo bacteriófago de
Streptomyces coelicolor, sensible al sistema de restricción Pgl y
dependiente de glicosilación de proteínas.**

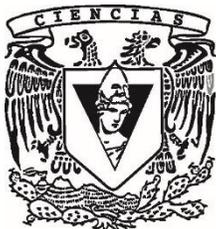
T E S I S

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGO

P R E S E N T A :

Yersain Ely Keller de la Rosa



**DIRECTOR DE TESIS:
Dr. Luis Servín González**

Ciudad Universitaria, Cd. Mx., 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Hoja de datos del jurado

- I. Datos del alumno
Yersaín Ely
Keller
de la Rosa
Universidad Nacional Autónoma de México
Facultad de Ciencias
Biología
310123598
- II. Datos del Tutor
Dr.
Luis
Servín
González
- III. Datos del sinodal 1
Dra.
Claudia
Segal
Kischinevzky
- IV. Datos del sinodal 2
Dra.
Gloria
Soberón
Chávez
- V. Datos del sinodal 3
Dr.
Sergio
Sánchez
Esquivel
- VI. Datos del Sinodal 4
Dr.
Benito
Estrada
Mena

Datos del trabajo escrito

Aislamiento y caracterización de un nuevo bacteriófago de *Streptomyces coelicolor*, sensible al sistema de restricción Pgl y dependiente de glicosilación de proteínas.

56 páginas.

2018

Agradecimientos

A mi tutor, el Dr. Luis Servín González, por el tiempo y dedicación en enseñarme biología molecular, así como la confianza que me ha tenido.

A la M. en C. Gabriela González Cerón, por el apoyo técnico y los consejos en los procedimientos del laboratorio.

A la Dra. Claudia Segal Kischinevzky por la dedicación en la lectura y las correcciones sobre este manuscrito.

A la Dra. Gloria Soberón Chávez por sus comentarios sobre el proyecto realizado, durante la revisión y los varios seminarios de grupo.

Al Dr. Benito Estrada Mena por la clase en la Facultad de Ciencias, así como a sus valiosos comentarios sobre el proyecto.

Al Dr. Sergio Sánchez Esquivel por las correcciones realizadas a este trabajo.

A mi familia por la confianza, el apoyo incondicional y su amor.

A mis compañeros de laboratorio, por los seminarios donde desarrollamos y discutimos los proyectos.

A mis amigos de la Facultad.

Y finalmente a Claudia Carrillo por ser mi compañera de vida, animarme, apoyarme y hacer de mí, una mejor persona.

Índice

Resumen	1
Introducción	2
• Generalidades de las actinobacterias y del género <i>Streptomyces</i>	3
• Bacteriófagos de <i>Streptomyces</i>	5
• Generalidades del fago ϕ C31.....	5
• Estructura de ϕ C31 y su genoma.....	6
• Fago ϕ BT1.....	7
• Ciclo de vida de los bacteriófagos.....	8
• Sistemas de modificación-restricción.....	8
• Sistema Pgl.....	9
• Glicosilación de proteínas en actinomicetos.....	11
Justificación	13
Objetivos	14
Material y Métodos	
• Cepas y bacteriófagos utilizados.....	15
• Aislamiento y purificación de los fagos.....	16
• Propagación de fagos en cepas <i>S. coelicolor</i> J1928, <i>S. coelicolor</i> M145, <i>S.lividans</i> <i>Δpmt</i> , <i>S.lividans</i> 1326	17
• Obtención de cepas lisogénicas	17
• Purificación de DNA genómico del fago ϕ M1.....	18
• Purificación del DNA de la cepa de <i>S. coelicolor</i> lisogénica de ϕ M1.....	19
• Secuenciación del DNA y análisis bioinformático.....	19
Resultados y discusión	21
Conclusiones	45
Perspectivas	46
Referencias	47
Protocolos	51

Resumen

El fago mejor estudiado de la especie bacteriana *Streptomyces coelicolor* es ϕ C31, el cual posee un intervalo relativamente amplio de hospederos, ya que es capaz de infectar a una tercera parte de las especies de *Streptomyces* conocidas; ϕ C31 ha sido empleado para el desarrollo de vectores para la manipulación genética de las diversas especies de *Streptomyces*, muchas de las cuales son importantes productores de antibióticos y otros metabolitos secundarios.

Se ha descrito que la proteína(s) receptora(s) del fago se encuentra glicosilada y sólo así se realiza la infección por las partículas de ϕ C31, ya que las cepas mutantes de *S. coelicolor* afectadas en la capacidad de glicosilar proteínas son resistentes a la infección por este fago (Cowlshaw & Smith, 2001). La glicosilación es una modificación importante de proteínas en actinobacterias y un sistema general de O-manosilación de proteínas está presente en los actinomicetos (Schmidt *et al.*, 2003). Adicionalmente al requerimiento de la glicosilación de proteínas para llevar a cabo la infección, la propagación de ϕ C31 en *S. coelicolor* solamente puede ocurrir en la ausencia del sistema denominado *phage growth limitation* o Pgl, por sus siglas en inglés (Hoskisson *et al.* 2015). Desde el descubriendo de la sensibilidad del fago ϕ C31 al sistema de restricción Pgl, se han tratado de obtener mutantes del fago que presenten resistencia al sistema Pgl, sin embargo, hasta el presente no se han obtenido éstas variantes, por lo cual, otro método de estudio es aislar y caracterizar nuevos fagos sensibles al sistema, y/o dependientes de la glicosilación de proteínas de *Streptomyces*.

En este trabajo se aisló y se caracterizó genómicamente un nuevo fago, llamado ϕ M1, dependiente de glicosilación de proteínas para infectar y sensible al sistema de restricción Pgl. El genoma de este fago consiste de 41,538 pb, y contiene 53 ORFs, de los cuales a 22 se les asignó una función putativa. En 50 ORFs se encuentra identidad con genes de ϕ C31. La identidad global a nivel de nucleótidos con ϕ C31 es de 90%. Se observó una estructura de mosaico entre ambos genomas. Adicionalmente se identificó el sitio attP del fago ϕ M1 y el sitio de integración en el cromosoma en *S. coelicolor*. Además, se discute un propone sitio de reconocimiento del sistema de restricción Pgl.

Introducción

Generalidades de las actinobacterias y del género *Streptomyces*

Los actinomicetos o actinobacterias son bacterias Gram-positivas, aerobias, formadoras de esporas, pertenecientes al orden Actinomycetales, que se caracterizan por poseer en su genoma un contenido alto de guanina y citosina (G+C), así como un ciclo de vida complejo. En cuanto a su morfología pueden variar de cocos a formas miceliales diferenciadas y entre ellas se encuentran algunos de los modelos bacterianos más estudiados, como *Mycobacterium*, *Corynebacterium* y *Streptomyces*. El género *Streptomyces* consiste de actinobacterias miceliales que se encuentran entre las más numerosas y ubicuas del suelo y la especie *Streptomyces coelicolor* ha sido empleada como uno de los más importantes modelos de estudio bacterianos (Chandra, 2014).

El ciclo de vida de *Streptomyces* comienza con la germinación de una espora, lo cual da lugar al crecimiento de un micelio vegetativo ramificado que posee pocos septos y en el que cada compartimento es multinucleoide. Posteriormente al crecimiento del micelio vegetativo, emergen de la superficie de la colonia estructuras especializadas llamadas hifas aéreas, que constituyen al micelio aéreo, también multinucleoide. Estas hifas dan lugar, mediante septación sincrónica, a hifas aéreas con compartimentos uninucleoideos, que posteriormente se diferencian en cadenas de esporas; este proceso se esquematiza en la figura 1 (Kwak, 2001; Yagüe, 2012;).

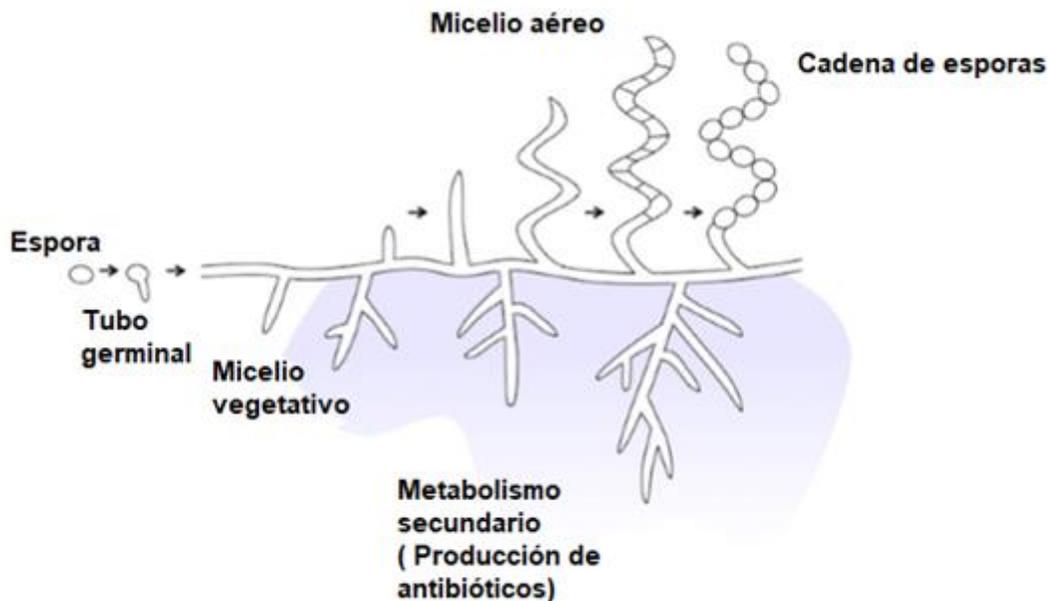


Figura 1. Ciclo de vida de *Streptomyces*. Modificado de Swiercz, 2012.

Los genomas completos de algunas especies de *Streptomyces* han sido secuenciados completamente y publicados, por ejemplo, *S. coelicolor* A3(2) (Bentley et al., 2002), *Streptomyces avermitilis* MA-4680 (Ikeda et al. 2003),

Streptomyces lividans 66 (Cruz-Morales et al. 2013) y *Streptomyces albus* J1074 (Zaburannyi, N et al. 2013). El cromosoma dentro del género es típicamente de topología lineal, contiene una región central o “core”, en la cual se localizan los genes esenciales (Hopwood, 2006).

S. coelicolor A3(2) es una de las especies más estudiadas, y es la mejor caracterizada genéticamente (Hopwood,1999). *S. coelicolor* A3(2) y *S. lividans* 66, son organismos muy cercanamente relacionados, con 99.5% de identidad en la secuencia de rRNA 16S. Ambos pertenecen al sub-clado *Streptomyces violaceouruber* (Jayapal, 2007). *S. coelicolor* tiene un cromosoma de 8,667,507 pb con 7,825 genes predichos mientras que el de *S. lividans* cuenta con 8,496,762 pb, y 7,748 genes.

En el genoma de *S. lividans* se encuentran presentes 299 genes que no están en *S. coelicolor* los cuales se encuentran contenidos en 4 grupos denominados islas genómicas sliGI-1 a sliGI-4. Por otro lado *S. coelicolor* también posee numerosas islas e islotes genómicos que contienen genes ausentes en *S. lividans* (Jayapal, 2007) entre los que se encuentran los sistemas de restricción de DNA metilado (González-Cerón et al., 2009). Además, otra diferencia se encuentra en que *S. lividans* presenta una disminución de la actividad de proteasas.

S. albus J1074 con 6,841,649pb posee unos de los genomas más pequeños del género. Es una cepa derivada de *S. albus* G1 deficiente en el sistema modificación-restricción *Sall* (*SalGI*) y además posee un crecimiento rápido (Myronovsky, 2014).

Bacteriófagos de *Streptomyces*

Todos los bacteriófagos (también nombrados fagos) de *Streptomyces* aislados hasta el momento, presentan DNA de doble cadena, con morfología sifoviral: cabeza polihédrica, con una cola flexible, larga y no contráctil. De los análisis de sus genomas, se conoce que cuentan con una topología lineal con extremos *cos*. El tamaño de sus genomas se encuentra entre 40 kpb y 121 kpb. El rango de hospederos puede ser muy amplio o muy específico, dependiendo del fago (Cowlshaw, 2001).

Los fagos de *Streptomyces* han sido empleados como herramientas genéticas para el estudio de *Streptomyces*, ya sea como vectores de clonación, para la generación de vectores de integración sitio-específica o como cósmidos (Kuhstoss, 1991).

En 2013 Smith y col. analizaron 8 genomas de fagos, de los cuales 4 ya se han empleado como herramientas genéticas en diferentes especies de *Streptomyces* (R4, Hau3, TG1 y SV1) y 4 fagos nuevos (ELB20, Zemlya, Lika, y Sujidade,). Las secuencias generadas se añaden a los 6 genomas anteriormente secuenciados. (ϕ C31, ϕ BT1, Mu1/6, SASD1, SV1 y VWB). En la tabla 1 se resumen las características de estos fagos, así como la pertenencia a un clúster, los cuales son grupos que presentan similitud sinténica mayor al 50% del genoma entre los elementos del clúster, y substancialmente una similitud menor a otro clúster.

Posteriormente se publicaron las secuencias de otros fagos ϕ CAM (Monson,2013), Nanodon (Erill,2016), Verse y Amela (Layton, 2016), Maih, TP1604, Xkcd426, YDN12, BabyGotBac (Donegan-Quick, 2017).

Tabla 1. Características de los genomas de los fagos de *Streptomyces* (Smith, 2013).

Fago	Número de Nucleótidos	Núm. ORFs	%GC	Clúster	Hospedero del cual se aisló
ELB20	51,160	82	67.0	R4	<i>S. lividans</i>
Lika	51,252	82	65.8	R4	<i>S. lividans</i>
Mu1/6	38,196	55	71.2		<i>S. aureofaciens</i> B96
ϕ BT1	41,831	55	62.8	ϕ C31	<i>S. lividans</i>
ϕ C31	41,491	53	63.6	ϕ C31	<i>S. lividans</i>
ϕ Hau3	50,255	77	67.8	R4	<i>S. hygroscopicus</i>
SASD1	37,068	43	66.2		<i>S. avermitilis</i> ATCC 31267
pZL12	90,435	112	69.5		<i>Streptomyces</i> <i>sp. cepa 9R-2</i>
R4	51,071	86	67.0	R4	<i>S. albus G</i>
Sujidade	51,552	81	65.7	R4	<i>S. lividans</i>
SV1	37,612	55	72.7	SV1	<i>S. venezuelae</i>
TG1	40,474	54	64.7	ϕ C31	<i>S. cattleya</i>
VWB	49,220	74	71.1	VWB	<i>S. venezuelae</i> ETH14630

Zemlya	51,077	79	65.7	R4	<i>S. lividans</i>
Nanodon	50,082	75	65.7	R4	<i>S. griseus</i>
φCAM	50,348	72	63.8	R4	<i>S. coelicolor</i>
Verse	49,483,	75	65.6	R4	<i>S. venezuelae</i>
Amela	49,452	75	65.6	R4	<i>S. venezuelae</i>
BabyGot Bac	57,165	72	69.2		<i>S. griseus</i>
Maih	57,256	70	69.3		<i>S. griseus</i>
TP1604	57,168	71	69.2		<i>S. griseus</i>
Xkcd426	64,477	78	68.8		<i>S. griseus</i>
YDN12	56,528	70	69.2		<i>S. griseus</i>

Continuación de Tabla 1. Características de los genomas de los fagos de *Streptomyces* (Smith, 2013).

El clúster que contiene más miembros está representado por R4 y se integra por R4, φHau3, ELB20, Zemlya, Lika, y Sujidade.

Un segundo grupo, representado por φC31, contiene a φC31, φBT1 y TG1. φSASD1 no presenta una similitud en nucleótidos suficiente para estar incluido en el clúster.

Los fagos VWB, SV1, Mu1/6, y pZL12 no muestran una similitud considerable con otros fagos.

Las regiones más conservadas entre los fagos corresponden a los genes cuyas proteínas se involucran en el empaquetamiento y cápside y las regiones de expresión temprana. Las regiones más variables se encuentran entre los genes que codifican las proteínas de la cola o las fibras de la cola, además de los involucrados en la regulación.

La organización global de los genomas se encuentra conservada, comprende los genes de la terminasa, cápside, cola, lisis, regulación y los genes de la región temprana y también los genes involucrados en el metabolismo del DNA.

Generalidades del fago φC31

φC31 es un bacteriófago atemperado y desde su aislamiento ha sido ampliamente estudiado, presenta una amplia gama de hospederos. Fue descubierto después de “plaquear” sobrenadantes de cultivos líquidos de *S. coelicolor* A3(2), en *S. anthocianicus* 31 y *S. lividans* 66 y 130.

El fago no forma placas en *S. coelicolor* A3 (2) silvestre, mientras que en la cepa *S. lividans* 66 forma placas grandes; existen variantes de *S. coelicolor* sensibles al fago, que producen placas pequeñas y turbias (Lomovskaya, 1980).

Originalmente se pensó que la resistencia silvestre de *S. coelicolor* A3 (2) se debía a la presencia de un profago defectuoso de ϕ C31 (Lomovskaya et al, 1970). Sin embargo, no existen secuencias homólogas presentes de ϕ C31 en el cromosoma de *S. coelicolor*. Posteriormente se descubrió que la resistencia se debe a un sistema de restricción fago-específico denominado Pgl (phage growth limitation), el cual será detallado posteriormente.

Estructura de ϕ C31 y su genoma

Es un fago sin membrana lipídica, con genoma lineal de DNA de doble cadena con extremos *cos*; la estructura de la cabeza es poliédrica de 53 nm de tamaño, con una larga cola no contráctil de 100 x 5 nm y una placa basal de 15nm en diámetro (Suárez et al., 1984).

En 1999 se obtuvo la secuencia completa de su genoma (Smith,1999). El tamaño del genoma de ϕ C31 es de 41,491 pb, con un promedio de G+C de 63.6% (el cual es más bajo del promedio de los cromosomas de cepas de *Streptomyces* que normalmente es de alrededor del 73%) y tiene 54 genes (figura.2). El arreglo de los genes indica que, con la excepción de uno, todos están en la misma orientación. Como en otros genomas de fagos los genes de expresión temprana se encuentran agrupados y en un extremo del cromosoma fágico y los genes de expresión tardía agrupados en el otro extremo.

El gen regulatorio “c”, produce 3 proteínas isomórficas (74, 54, 42 kDa) y es requerido para el mantenimiento de la lisogenia. La proteína de 74 kDa está solamente presente en cantidades traza en la lisógena y no es requerido para dar inmunidad. Las proteínas de 54 y 42 kDa son producidas en cantidades aproximadamente iguales. Se ha reportado que las interacciones proteína-proteína son importantes en el control del desarrollo lítico (Lomovskaya et al, 1980; Smith, 1991).

Han sido identificados dos tipos de promotores en el cromosoma del fago. Los promotores de “expresión temprana” como son, promotores del gen represor, *cp1*, *cp2* y el promotor para el gen de la integrasa, *intp* (Smith, 1991; Kuhstoss, 1991). El segundo tipo de promotores son fago-específicos (Ingham, 1993).

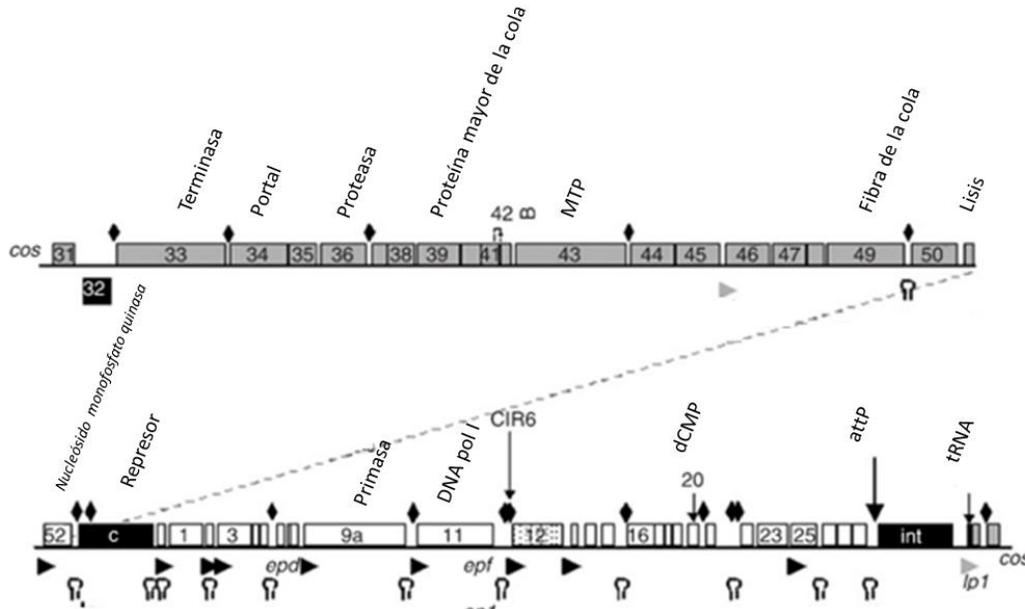


Figura 2. Organización genómica de ϕ C31. El genoma está representado por una línea, separada en el medio, con el brazo izquierdo (sección superior) y el brazo derecho (sección inferior). Los extremos representan los extremos cohesivos (*cos*) del genoma lineal. Los genes están representados por recuadros, algunos de los cuales están numerados, con genes tardíos en gris, genes tempranos en blanco y genes tempranos inmediatos en color negro (represor, *c*; integrasa, *int*; gen 32). También se representa el gen de tRNA, con una barra de color negro. Los diamantes representan los sitios de unión del represor o CIR. Terminadores (iconos de asa-bucle) (Modificado de Smith, 2006).

Fago ϕ BT1

El genoma del fago ϕ BT1 presenta una gran similitud con el fago ϕ C31. La mayoría de los genes tardíos codifican proteínas con identidades de 73% al 96% a sus homólogos de ϕ C31. Las proteínas de los genes de la región temprana son más variables con identidades que van de <30% a 92% en su secuencia. (Smith, 2006).

Existe evidencia de mosaicismos entre los dos genomas, lo que significa que regiones de DNA han sido insertadas y/o eliminadas en un genoma, pero no en el otro. La integrasa de ϕ BT1 exhibe una identidad del 26% con su homólogo de ϕ C31, mientras que el sitio de "attachment" es diferente entre estos fagos (Gregory, 2003), insertándose cada uno en una diferente región del cromosoma. Esto ha permitido el desarrollo de vectores integrativos con el gen *int* de ϕ BT1 compatibles con los desarrollados a partir del fago ϕ C31.

Ciclo de vida de los bacteriófagos

La proliferación de fagos ocurre en una serie de eventos; empieza por el reconocimiento del receptor, posteriormente la absorción de un virión en la pared celular del hospedero, el siguiente paso es la inyección del genoma viral a través de la membrana. La estrategia de replicación dependerá si el fago es virulento o atemperado. Los fagos virulentos, como el fago VP II de *S. coelicolor* o Sf1 de *S. flavovirens*, sólo se replican a través del ciclo lítico, un proceso involucrado en la formación de nueva progenie viral y la liberación de la célula infectada.

Alternativamente, los fagos atemperados como ϕ C31 o SV1 entran ya sea al ciclo lítico o lisogénico. Bajo condiciones de estrés, los profagos pueden salir del estado lisogénico y producir más viriones que son liberados de la bacteria. Normalmente, la liberación de la progenie del fago da lugar a la muerte bacteriana a través de lisis celular.

Sistemas de modificación- restricción

Si un bacteriófago se logra absorber e inyectar su genoma dentro de la bacteria, varios sistemas de defensa innatos toman lugar para prevenir la replicación del fago. Los sistemas de modificación y restricción le permiten a la célula hospedera degradar DNA exógeno. Característicamente, los sistemas M-R consisten en una endonucleasa de restricción (REasa) y una metiltransferasa (MTasa). La MTasa modifica el DNA en las adeninas o citocinas en una secuencia específica, y así previene el corte del DNA por acción de la REasa, la cual reconoce un sitio específico de corte, degradando el DNA en fragmentos inocuos. Los sistemas de modificación y restricción han sido clasificados en 4 tipos con base en la composición de las subunidades, el sitio de reconocimiento, los requerimientos de ATP y cofactores y el mecanismo de acción (Roberts, 2003; Tock, 2005).

Tipo I: Se componen de dos subunidades de MTasa, una subunidad S, que puede unirse específicamente al sitio de reconocimiento y dos subunidades de REasa. La subunidad S contiene dos dominios de reconocimiento del blanco y cada uno interactúa con la mitad del sitio de reconocimiento bipartito. La MTasa modifica ambas cadenas en la posición N6. Las REasas requieren S-adenosilmetionina (SAM), ATP y Mg^{2+} . El sitio de escisión se encuentra varios pares de bases lejos del sitio de reconocimiento.

Tipo II: Son usualmente producidos por dos genes que codifican una MTasa y REasa, respectivamente. Su sitio de reconocimiento es a menudo una secuencia palíndroma pequeña (normalmente de 4 a 8 pb). La MTasa es un monómero que metila ambas cadenas de DNA después de la unión, utilizando SAM como donador de grupos metilos. La REasa es un complejo dimétrico el cual puede unirse al sitio de reconocimiento no metilado y cortar al DNA adyacentemente o dentro de la secuencia de reconocimiento. En estos sistemas la MTasa y la REasa no forman un complejo, sino que actúan en forma independiente. Estos sistemas de tipo II se encuentran presentes en algunas especies del género *Streptomyces*, de donde se

han aislado algunas de las más utilizadas para hacer ingeniería genética (por ejemplo *Sall*, *Sacl*, *SaclI*, *SphI* etc), (Roberts, 2003).

Tipo III: Consisten en dos genes localizados cercanamente, *mod* y *res*, cuyas proteínas pueden reconocer, modificar [Mod] e hidrolizar [Res] ciertas secuencias de DNA. El sitio de reconocimiento son secuencias no-palíndromas cortas. La MTasa consiste de dos subunidades y modifica sólo una de las cadenas de DNA después de la unión. La REasa requiere ATP y Mg^{2+} .

Tipo IV: A diferencia de las de tipo I, II y III, las enzimas de tipo IV solamente hidrolizan DNA que ha sido modificado por metilación y por lo mismo contienen sólo un componente REasa. Un ejemplo son las enzimas *McrA* y *McrB* de *E. coli*, que reconocen residuos de metilcitosina y *Mrr* que puede reconocer metilcitosina y metiladenina. Poseen baja especificidad, lo que les permite proteger a la célula hospedera de una amplia gama de DNA exógeno con diferentes patrones de metilación (Pingoud, 2014; Loenen, W 2014).

Streptomyces coelicolor A3(2) cuenta con sistemas de restricción de DNA metilado, al codificar endonucleasas metil-específicas. Estos sistemas funcionan como un mecanismo de defensa para distinguir entre el DNA propio (el cual no está modificado) y cualquier DNA exógeno que ingrese a la célula (metilado). Así el DNA que está modificado por metilación en ciertas bases es reconocido y después degradado por enzimas específicas (González- Cerón, 2009).

Además, posee con un sistema de restricción denominado Pgl (phage growth limitation) y se ha propuesto que el mecanismo metil-especifico se utiliza en este sistema (Cox, 1984; Laity, 1993).

Sistema Pgl

El sistema Pgl de *Streptomyces coelicolor*, es un mecanismo de restricción, que le confiere resistencia contra el bacteriófago Φ C31 y fagos homoinmunes relacionados. El sistema consiste en lo siguiente: primero se produce una infección del fago a una cepa silvestre Pgl⁺, en este momento hay un ciclo de infección que produce progenie; sin embargo, la progenie liberada de este ciclo de infección se encuentra modificada (aparentemente por medio de metilación de adeninas), lo cual bloquea la replicación en el segundo ciclo de infección (figura 3).

La modificación que ocurre en el DNA del fago, a causa de las cepas Pgl⁺, no es letal para el fago, ya que los fagos atenuados liberados en el primer ciclo de infección son capaces de infectar a una cepa Pgl⁻ (Sumby, 2002; Hoskisson, 2015).

El mecanismo molecular de este sistema de restricción aún no es del todo claro, sin embargo se han identificado 4 genes, en dos operones (*pglWX* y *pglYZ*). Que son esenciales para el fenotipo Pgl.

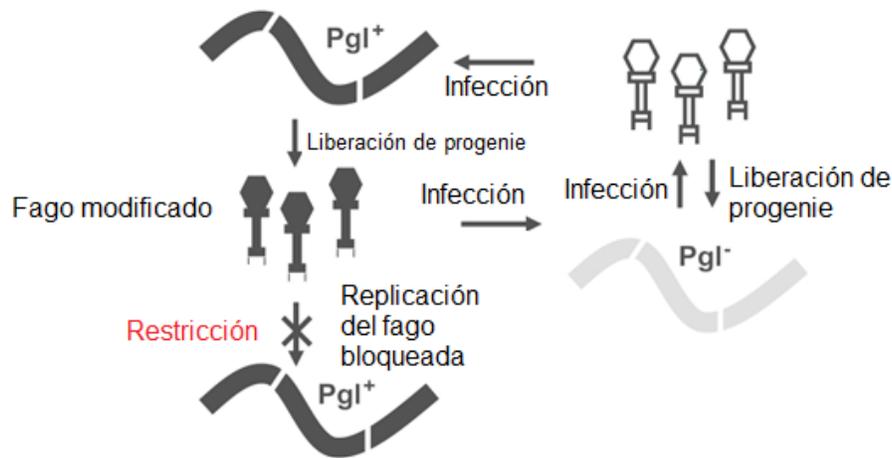


Figura 3. Esquema del fenotipo pgl, basado en el modelo de Chinenova 1982. (Hoskisson, 2015).

Los productos predichos para estos genes son los siguientes:

PglX es una metiltransferasa predicha de 135kDa (1,211 aa) que contiene una “firma” de metiltransferasa de DNA específica de adeninas en la posición 375-381 aa y un sitio de unión a SAM en la posición 237-297.

PglW presenta una actividad de cinasa de proteínas, consiste de 1,557 aminoácidos, con un tamaño de 169 kDa. Los análisis bioinformáticos de la secuencia revelan la presencia de dos dominios putativos de cinasa, un dominio de cinasa de tirosina en 195-490 aa y un dominio de cinasa de serin/treonina en 530-816 aa.

PglY cuenta aparentemente con actividad de ATPasa, ya que muestra la presencia de un motivo Walker A (de unión a ATP (GSFGSGKS)) entre los aminoácidos 75-82.

PglZ se predice que contiene un plegamiento conservado de fosfatasa alcalina.

Se ha propuesto que la metilación del DNA ocurre solamente durante la infección del fago (Hoskisson, 2015). El hecho de que no haya sido posible a la fecha el aislamiento de derivados de ϕ C31 resistentes al sistema Pgl sugiere que el blanco de este sistema se encuentra en múltiples copias en el genoma o bien se encuentra en una región esencial del fago (por ejemplo, en gen(es) codificante(s) de proteína(s) que interactúan con el receptor del hospedero).

Debido a que sólo fagos del mismo grupo de inmunidad de ϕ C31 son sensibles al sistema Pgl también se ha propuesto que el blanco para el sistema Pgl pudiera encontrarse en el sistema represor-operador, posiblemente en los operadores o sitios CIR. Se han localizado 18 sitios CIR a lo largo del genoma, y los sitios se encuentran perfectamente conservados con una secuencia núcleo de 17 bp (Ingham et al., 1993).

Otra característica del sistema Pgl es una variación de fase de alta frecuencia en la que una cepa Pgl⁺ da lugar espontáneamente a cepas Pgl⁻ con una frecuencia de 10⁻² a 10⁻³ y Pgl⁻ a Pgl⁺ con una frecuencia de 10⁻³ a 10⁻⁴ (Chinenova et al., 1982, Laity et al., 1993). La variación de fase se ha atribuido a una variación en la longitud de una secuencia repetida de guaninas dentro del gen *pglX* (Sumbly, 2003).

S. lividans no cuenta con los genes del sistema Pgl y por lo tanto es naturalmente Pgl⁻, sin embargo, puede convertirse a Pgl⁺ mediante la transferencia de los cuatro genes del sistema Pgl (Sumbly et al., 2003).

Glicosilación de proteínas en actinomicetos.

Las glicoproteínas pueden dividirse en dos grandes grupos dependiendo del sitio de unión del carbohidrato a la proteína. La N-glicosilación corresponde a la unión de un oligosacárido a la amida de la cadena lateral de un residuo de asparagina de la proteína y es común encontrarla en la secuencia consenso N-X-S(T); y la O-glicosilación se presenta cuando el oligosacárido se une covalentemente en un enlace glucosídico al grupo hidroxilo de un residuo de serina o de treonina de la proteína en secuencias ricas en glicina, prolina y valina (Lommel, 2009; Smith, et al., 2014). Un sistema general de O-glicosilación, en particular O-manosilación, de proteínas está presente en las actinobacterias y se ha demostrado su funcionalidad en los géneros *Mycobacterium*, *Corynebacterium* y *Streptomyces* los cuales glicosilan varios productos bioactivos naturales y antígenos secretados (Schmidt, 2003).

Las glicoproteínas cumplen diversas funciones, dependiendo del tipo de proteína y el azúcar conjugado que contienen, entre éstas destacan la adhesión celular, protección contra la proteólisis, incremento de la solubilidad y virulencia de la célula (Córdova-Dávalos, 2009)

La biogénesis de la O-manosilación inicia cuando la fructosa-6-fosfato genera GDP-manosa por la acción de las enzimas, fosfomanosa isomerasa (PMI), fosfomanomutasa (PMM) y GDP-manosapirofosforilasa (GMPP). La manosa es entonces activada mediante su transferencia a una molécula de poliprenol fosforilado, que se localiza en el lado citoplásmico de la membrana; la enzima poliprenol-fosfato-manosa sintasa (Ppm sintasa), es la encargada de unir covalentemente la molécula de manosa, proveniente de GDP manosa, al poliprenol fosforilado. Posteriormente, el poliprenol fosfato manosilado expone la manosa en el sitio extracelular de la membrana por un movimiento de “voltereta” o “flipping” a través de la membrana. Entonces la manosa puede ser utilizada como sustrato por la proteína de membrana manosil-transferasa de proteínas (Pmt), la cual unirá covalentemente el residuo de manosa a los grupos hidroxilo de las cadenas laterales de serinas y treoninas de las proteínas que se están secretando, acción que ocurre en el lado extracelular de la membrana (Vanderven, 2005; Lommel 2009;

Espitia et al., 2010). A diferencia de la *N*-glicosilación no existe un motivo de aminoácidos específico que pueda determinar con precisión los sitios de glicosilación, aunque sí se sabe que la *O*-manosilación ocurre preferentemente en regiones donde las serinas y treoninas se encuentran en regiones ricas en prolinas, alaninas y glicinas (Smith et al., 2014).

Se han descrito algunas glicoproteínas que contienen *O*-manosil-glicanos en bacterias. Inicialmente se describieron en *M. tuberculosis* por Espitia y Mancilla (1989), y a partir de entonces se han descrito un gran número de proteínas antigénicas glicosiladas de las microbacterias, incluyendo la proteína secretada Apa, y un gran número de lipoproteínas glicosiladas, tales como LpqH, MPT83, PstS1, etc (Hewinson, 1996; Lara, 2004), así como una $\beta(1,4)$ -glucanasa de *Cellomonas fimi* y *S. lividans* (Ong et al., 1994) y la proteína de unión a fosfato PstS de *S. coelicolor* (Wehmeier et al. 2009).

El reconocimiento de la existencia de la glicosilación de proteínas glicosiladas en *Streptomyces* tiene su origen en la búsqueda del receptor del fago ϕ C31. La búsqueda del receptor del fago se realizó mediante el aislamiento de cepas mutantes de *S. coelicolor* resistentes a ϕ C31, pero formadoras de placas por transfección obtenidas por mutagénesis aleatoria con luz UV. Posteriormente el análisis de las mutaciones de estas cepas reveló que los genes afectados por las mutaciones (*sco1423* y *sco3154*) codificaban para enzimas homólogas a DPM sintasa y PMT, respectivamente, ambas proteínas de *Saccharomyces cerevisiae* (Cowlshaw D & Smith, M., 2001; Cowlshaw & Smith, 2002).

Dado que tanto mutaciones en *pmt* (*sco3154*) como en *ppm* (*sco1423*) de *S. coelicolor*, resultan en la resistencia al bacteriófago ϕ C31, se sugiere que el receptor de ϕ C31 es una glicoproteína de superficie celular *O*-manosilada.

Hasta el momento no se ha descrito otro bacteriófago aparte de ϕ C31 que requiera de la glicosilación del receptor para poder infectar, o que presente sensibilidad al sistema Pgl de *S. coelicolor*. Por lo anterior, el objetivo del presente trabajo es el aislamiento y caracterización de bacteriófagos capaces de infectar a *S. lividans* y a *S. coelicolor* que presenten una o ambas características.

Justificación

El bacteriófago ϕ C31 es el fago mejor estudiado del género *Streptomyces*. Desde su aislamiento en la década de los 60s del siglo pasado, se ha estudiado su absorción, estabilidad, curvas de crecimiento, entre otras características biológicas, así como el estudio de su genoma completo (Smith, 1999); sin embargo, desde el descubrimiento de la resistencia natural de *S. coelicolor* a ϕ C31 por el sistema Pgl (Chinenova, 1982), no se ha determinado el sitio de reconocimiento de las proteínas de este sistema dentro del genoma del fago.

Otra característica del fago ϕ C31 es el usar como receptor a una proteína glicosilada de *Streptomyces*, lo cual es poco común en la naturaleza; aunque hasta el momento el receptor no ha sido identificado, sí se sabe que debe estar glicosilado.

El aislamiento de nuevos fagos con uno o ambos fenotipos nos permitirá comparar y proponer por homología de secuencias los posibles sitios de reconocimiento y una posible relación entre ambos fenotipos.

Objetivo general:

Aislar y caracterizar nuevos fagos de *S. coelicolor* sensibles al sistema Pgl y/o dependientes de glicosilación de proteínas para llevar a cabo la infección.

Objetivos particulares

1. Aislar actinofagos de diversos suelos de México.
2. Determinar sensibilidad al sistema Pgl mediante la utilización de cepas Pgl⁺ y Pgl⁻
3. Determinar si requieren de un receptor glicosilado, mediante la utilización de cepas afectadas en la glicosilación de proteínas.
4. Identificar el grupo de inmunidad de los fagos que presenten el fenotipo esperado.
5. Secuenciar y analizar el genoma de cualquier fago nuevo encontrado.
6. Comparar el genoma de los actinofagos nuevos encontrados respecto a ϕ C31 y homólogos relacionados.

Materiales y métodos

1. Cepas y bacteriófagos utilizados:

En este trabajo se utilizaron cepas del género *Streptomyces* (tabla 1) con el objetivo de la propagación de los fagos. Todos los cultivos se crecieron a 30°C.

La descripción de las cepas y fagos se incluye a continuación en la Tabla 1.

Cepa	Descripción	Referencia
<i>Streptomyces coelicolor</i> M145	A3(2) SCP1 ⁻ SCP2 ⁻ , Pgl ⁺	Hopwood, 1985
<i>Streptomyces coelicolor</i> J1928	M145 pglY::aadA, Pgl ⁻	Bedford et al. 1995
<i>Streptomyces lividans</i> 1326	Silvestre	Kieser et al. 2000.
<i>Streptomyces lividans</i> Δ pmt	Mutante sin capacidad de glicosilar proteínas.	Arista-Carrera, 2016
<i>Streptomyces coelicolor</i> IB86	Mutante en enzima responsable de adicionar manosas a glicoproteínas	Arroyo-Peréz.E, 2017
<i>Streptomyces albus</i> J1074	Deficiente en el sistema de modificación-restricción SalG1 (r ⁻ m ⁻)	Chater & Wilde, 1980
Lisógenas		
<i>Streptomyces coelicolor</i> J1928 (ϕ M1)	Lisógena del fago ϕ M1	Este trabajo
<i>Streptomyces coelicolor</i> J1928 (KC515)	Lisógena del fago KC515	Este trabajo
<i>Streptomyces coelicolor</i> J1928 (KC516)	Lisógena del fago KC516	Este trabajo
<i>Streptomyces coelicolor</i> J1928 (KC857)	Lisógena del fago KC857	Este trabajo

--	--	--

Fago	Descripción	Referencia
φC31	Sensible a sistema de restricción Pgl, dependiente de glicosilación de proteínas para infectar.	Lomovskaya,1972
KC515	Derivado de φC31, c+ attP ⁻ tsr+	Rodicio, 1985
KC516	Derivado de φC31, att- c+ tsr+	Rodicio, 1985
KC519	Derivado de φC31, c ⁻ attP ⁻	Rodicio & Chater, 1988
KC857	Derivado de φC31, att- c+, tsr+	Bruton et al. 1991
φM1	Sensible al sistema de restricción Pgl, dependiente de glicosilación de proteínas	Este trabajo

Tabla 1. Cepas utilizadas de *Streptomyces* y bacteriófagos.

2.- Aislamiento y purificación de los fagos

Se utilizaron muestras de diferentes suelos (cada una con 10 g de tierra) y, mediante el protocolo para el aislamiento de fagos nuevos de *Streptomyces* (protocolo 1), se aislaron 70 fagos candidatos propagados con *S. lividans* 1326 como cepa indicadora.

Se realizaron ensayos de formación de placas, mezclando 20 µL de esporas concentradas con 100 µL de diferentes diluciones de una suspensión de fagos en buffer SM (Tris-HCl 50 mM pH 7.5, Mg₂SO₄ 10 mM, NaCl 100 mM, Gelatina 0.1 %) obtenidos anteriormente y 2.5 mL de medio SNA (consistente de Difco Nutrient Broth 8 g/L, agar 5 g/L), y se plaquearon en una caja de Petri con medio DNA (Difco Nutrient Agar 23 g/L) enriquecido con Ca(NO₃)₂ 8 mM, MgSO₄ 10 mM y glucosa 0.5%.

En las muestras de suelo comúnmente se encontraban más de un tipo de placas (placas turbias y claras, por ejemplo) de fagos, por lo que los diferentes fagos se

aislaron (protocolo 2), picando una placa de cada tipo con una pipeta Pasteur y dejándola en 1mL de buffer SM durante 2 horas a temperatura ambiente, para posteriormente realizar diluciones y hacer de nuevo ensayos de formación de placas, hasta que la morfología de las placas fuera uniforme. De estas placas se realizaron los “stocks” de los fagos con un alto título (protocolo 3).

2.1 Propagación de fagos en cepas *S. coelicolor* J1928, *S. coelicolor* M145, *S.lividans* Δ pmt, *S.lividans* 1326

Las suspensiones de fagos con alto título fueron probadas para formación de placas utilizando esporas de las diferentes cepas: *S. lividans* 1326, *S. lividans* Δ pmt, *S. coelicolor* J1928 y *S. coelicolor* M145, con el objetivo de identificar los fagos sensibles al sistema de restricción Pgl de la cepa *S. coelicolor* M145 y los fagos dependientes de glicosilación. Los primeros no formarían placas en la cepa *S. coelicolor* M145 pero sí en la *S. coelicolor* J1928, y los segundos no formarían placas en la cepa *S. lividans* Δ pmt.

De los 70 fagos aislados sólo uno presentó los dos fenotipos buscados, el cual fue nombrado ϕ M1.

3. Obtención de cepas lisogénicas

Con el objetivo de conocer si el nuevo fago aislado era un fago homoinmune a ϕ C31, se obtuvieron cepas lisogénicas de los fagos.

En medio MS (Manitol 20g/L, harina de soya 20g/L, agar 20g/L) se plaqueó 0.1 mL de una suspensión densa de esporas de *S. lividans* y se dejó secar completamente. Posteriormente, se marcó una zona en la base de la caja Petri, y sobre el medio se colocaron 20 μ L de una suspensión del fago (M1, KC515, KC516 y KC857) y se dejó secar completamente.

Se incubaron a 30°C, hasta que se observó esporulación en la zona previamente marcada en la caja (4-6 días). Se cosecharon las esporas de la zona marcada y se resuspendieron en 1 mL de pirofosfato tetrasódico (25 mM) y se incubó a temperatura ambiente 30 minutos.

Se mezcló en vórtex la suspensión 30 segundos y se filtró a través de un filtro de algodón montado en una pipeta Pasteur. Posteriormente se filtró de nuevo por un filtro con membrana de 0.45 μ m, con el objetivo de coleccionar las esporas. Se lavó la membrana pasando 5 mL de pirofosfato tetrasódico (25 mM) por 2 ocasiones más. Se tomó la membrana del filtro y se transfirió (en esterilidad) a un tubo de 5 mL y se le añadió 1 mL de H₂O destilada y se mezcló en vórtex.

Se realizaron diluciones de la suspensión de esporas, las cuales se plaquearon y se incubaron hasta observar crecimiento y esporulación de colonias (más de 4 días). En el caso de las cepas que fueron inoculadas con el fago M1 se realizaron réplicas de las colonias sobre cajas de medio DNA a las que se les vertió medio SNA con

esporas de *S. lividans* 1326 como cepa indicadora. Se incubaron una noche y se identificaron las colonias lisogénicas como aquellas que fueron capaces de liberar fagos, lo cual se evidenció como placas de lisis en el tapete de *S. lividans* alrededor de la réplica de las colonias.

Para obtener lisógenas con los fagos derivados de ϕ C31, se realizaron primeramente réplicas en medio MS con tioestreptón ($50 \mu\text{g mL}^{-1}$), se incubaron a 30°C y las colonias resistentes fueron purificadas. Las lisógenas se confirmaron por la prueba de liberación de fagos al medio. De las colonias lisogénicas de los fagos se realizaron “stocks” y se conservaron a -20°C para su posterior uso.

4. Purificación de DNA genómico del fago ϕ M1.

La purificación de DNA genómico del fago ϕ M1 presentó problemas al utilizar la cepa de *S. lividans* 1326, ya que se obtenían muy bajos rendimientos, por lo cual se empleó a la cepa *S. albus* J1074, en la cual el fago daba placas un poco más grandes y cuenta con un crecimiento rápido. Se inocularon 6 cajas Petri (8 cm) con medio DNA enriquecido, con 2.5 mL de medio SNA vertido en la superficie por cada caja, el cual contenía las esporas de *S. albus* J1074 y una concentración de fagos $\sim 2 \times 10^4$ pfu/mL. Se incubaron a 30°C durante una noche.

Con una espátula se raspó el agar suave y se colectó en un tubo de centrifuga al cual se añadieron 25 mL de medio DNB y se incubó a temperatura ambiente durante 2 horas. Se centrifugó a 10,000 rpm durante 10 minutos (para retirar el agar), y el sobrenadante se pasó a un tubo de ultra-centrifuga para centrifugar a 50,000 rpm durante 30 minutos. La pastilla fue resuspendida en 0.5 mL de solución de RNasa (buffer SM con RNasa $40 \mu\text{g/mL}$) e incubado a 37°C durante 20 minutos. Posteriormente se añadieron 100 μL de una mezcla de SDS (1 vol de Tris.HCl 2 M pH 9.6, 2 vol EDTA 0.5 M pH 7.4, y 1 vol SDS 10%), y se incubó a 70°C durante 30 minutos. A la solución se le añadieron 125 μL de acetato de potasio 8 M y se colocó en hielo 15 minutos. Se centrifugó a 14,000 rpm durante 30 minutos y al sobrenadante se le añadieron 1.75 mL de TE (Tris-HCl 10 mM, pH=8, EDTA 1 mM, pH=8) y se agregaron 4.5 ml de etanol absoluto. Se dejó toda la noche a -20°C .

Posteriormente se centrifugó a 8,000 rpm durante 15 minutos, el sobrenadante se desechó y el pellet se resuspendió en 250 μL de TE, se agregaron 80 μL de acetato de sodio 3 M, 2 volúmenes de etanol absoluto y se invirtió varias veces el tubo hasta observar la precipitación del DNA en forma de malla. El DNA se resuspendió en un volumen mínimo de buffer TE (100 μL) y se le adicionaron 33 μL de acetato de sodio 3 M, un volumen de isopropanol y se precipitó el DNA a temperatura ambiente. El sobrenadante se removió y se lavó el DNA con 250 μL de etanol 70%. Se removió el etanol, y se resuspendió el DNA en 50 μL de buffer TE.

El DNA fue digerido con enzimas de restricción con el objetivo de conocer el patrón de restricción y compararlo con el del fago ϕ C31.

Una muestra de DNA fue empleada para su secuenciación, para lo cual se circularizó mediante la incubación en enzima DNA ligasa de T4.

4. Purificación del DNA de la cepa de *S. coelicolor* lisogénica de ϕ M1

Se inocularon 100 μ L de una preparación densa de esporas en 25 mL de medio LB + 25% sacarosa (bacto triptona 10g/L, extracto de levadura 5 g/L, NaCl 10 g/L, sacarosa 250 g/L), y se incubó 48 horas a 30°C en rotación. El micelio se centrifugó y se resuspendió en 3 mL de buffer TE 25 mM (Tris-HCl 25 mM, pH=8, EDTA 25mM, pH=8), con lisozima a una concentración final de 2.5 mg/mL. Se incubó a 37°C durante 30-60 minutos (hasta que la solución se observara viscosa y transparente, lo que evidencia la lisis del micelio).

Se le añadieron 4 mL de SDS 2% y se mezcló en vórtex durante 1 minuto. Posteriormente se añadió un volumen de fenol/cloroformo y se agitó durante 15 segundos. La mezcla se centrifugó a 14,000 rpm durante 5 minutos. Se transfirió la fase acuosa a un tubo nuevo y se repitió el procedimiento de extracción con fenol/cloroformo al menos 3 veces más. Posteriormente se agregó 1/10 del volumen de acetato de sodio 3 M a la fase acuosa y un volumen de isopropanol, se mezcló y el DNA precipitado se lavó con etanol al 70%.

Se resuspendió el DNA en buffer TE y se añadió RNasa a 40 μ g/mL y se incubó a 37°C durante 30 minutos. Se realizó una extracción con fenol/cloroformo y se precipitó el DNA con 1/10 del volumen total de acetato de sodio 3 M y 1 volumen de isopropanol. El DNA se resuspendió en 100 μ L de buffer TE y se cuantificó.

5. Secuenciación del DNA y análisis bioinformático.

El genoma de ϕ M1 y la cepa lisogénica de ϕ M1 fueron secuenciados por el método de Illumina/Solexa en la Unidad Universitaria de Secuenciación Masiva y Bioinformática de la UNAM, ubicada en el Instituto de Biotecnología en Cuernavaca, Morelos.

Las lecturas-cortas obtenidas de la secuenciación del genoma de ϕ M1 fueron ensambladas preliminarmente en línea utilizando el programa CAP3 (Huang & Madan, 1999. <http://doua.prabi.fr/software/cap3>). Posteriormente se emplearon los programas Velvet (Zerbino & Birney, 2008) y ABySS (Simpson et al., 2009) para ensamblar tanto la secuencia del fago ϕ M1 como la de la cepa lisógena. Las secuencias resultantes de los ensamblajes se analizaron en Artemis (Rutherford K, *et al.*, 2000) con el objetivo de identificar secuencias que alinearan con la secuencia del genoma del fago ϕ C31 el cual se empleó como genoma de referencia.

Los marcos de lectura abiertos (ORFs, por sus siglas en inglés) fueron predichos con el programa GeneMark.hmm para procariontes (Besemer & Borodovsky, 1999), y también se utilizó el servidor RAST (*Rapid Annotations using Subsystems Technology*, <http://rast.theseed.org/FIG/rast.cgi>. Aziz, et al., 2008). Los genes resultantes fueron verificados manualmente bajo los siguientes criterios; la inclusión de la secuencia RBS y una secuencia mínima de 150 nucleótidos, y el uso de

codones. El análisis entre la similitud de nucleótidos se realizó empleando los programas Gepard (Krumsiek et al., 2007) y BLASTn (Altschult et al., 1990). Las proteínas hipotéticas fueron analizadas utilizando los servidores: Pfam, (<http://pfam.sanger.ac.uk/>), InterProScan (<http://www.ebi.ac.uk/Tools/InterProScan/>), HNN y HTH

(http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_server.html).

La búsqueda para genes de tRNA se realizó utilizando el servidor tRNAscan-SE (<http://lowelab.ucsc.edu/tRNAscan-SE/>) y Aragorn (Laslett, D. 2004).

Se realizó el ensamblaje del genoma de la cepa lisogénica de ϕ M1 y una búsqueda del genoma del fago en las secuencias con el objetivo de conocer el sitio de *attachment*, tanto en el cromosoma del fago como en el cromosoma de la cepa bacteriana.

Resultados y discusión

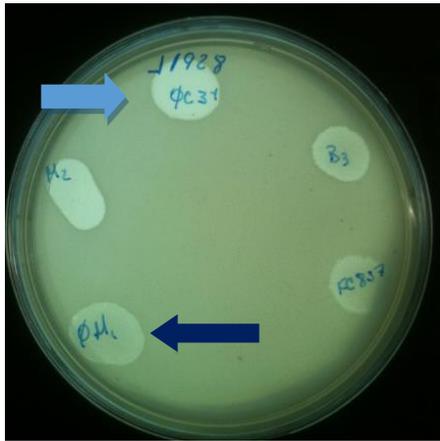
Aislamiento de fagos

Diversas muestras de suelo (10 g) fueron tratadas para buscar la presencia de actinofagos, utilizando *S. lividans* 1326 como cepa hospedera. Las muestras se obtuvieron de diversos sitios, así como zonas del país; Ciudad de México; (Coyoacán, Gustavo A. Madero, Santa Fé, por ejemplo.), Guanajuato, Querétaro, Veracruz, además de muestras de macetas. Se obtuvieron 70 fagos aislados y con un lisado de alto título, con los cuales se realizó el ensayo de formación de placa en las cepas (*S. coelicolor* M145, *S. coelicolor* J1928, *S. lividans* Δpmt). De todos los fagos aislados, sólo uno, presentó el fenotipo buscado, el cual fue llamado ϕ M1, por el sitio del que fue aislado de la tierra de una maceta.

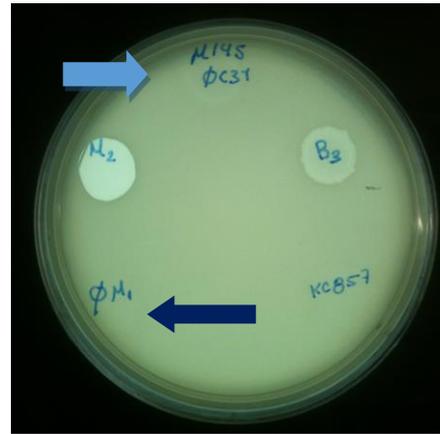
Se intentó purificar el DNA del fago M1 utilizando *S. lividans* 1326 como cepa hospedera, sin embargo, la cantidad de partículas obtenidas no eran suficientes para la purificación de DNA viral, quizá debido al ciclo de vida del fago. Por lo cual el DNA se purificó de la cepa *S. albus* J1074 la cual presenta crecimiento más rápido que *S. lividans* y *S. coelicolor* (Zaburanyi, 2014). Con el DNA purificado se realizaron las digestiones con varias enzimas de restricción con el objetivo de comparar el patrón de restricción con ϕ C31.

Ensayo de formación de placa de fagos aislados.

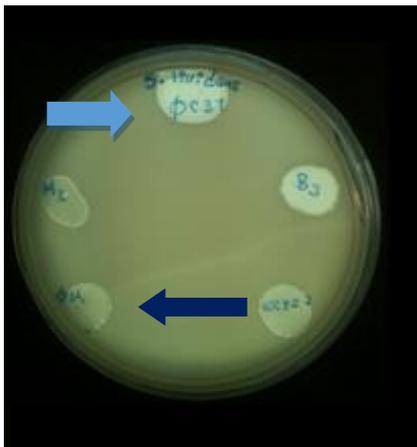
Se ha descrito que el receptor del fago ϕ C31 está glicosilado por la vía de O-manosilación (Cowlshaw & Smith, 2001). En adición ϕ C31 es sensible al sistema de restricción Pgl de *S. coelicolor*. Por lo tanto, se buscaba un fago distinto a ϕ C31 que presentara alguno de los fenotipos mencionados o ambos. *S. lividans* glicosila sus proteínas sin embargo no cuenta con los genes homólogos a *pglYZ*, necesarios para el fenotipo Pgl (Sumby, 2002).



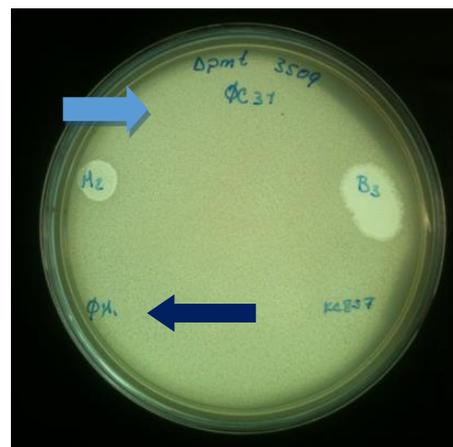
J1928(pgl-, glicosilación+)



M145 (Pgl+, glicosilación +)



S. lividans 1326,
(pgl -, glicosilación +)



S. lividans (Δpmt 3509, pgl -,
glicosilación -)

Figura 4. Ensayo de formación de placa en las cepas *S. coelicolor* J1928, *S. coelicolor* M145, *S. lividans* 1326, *S. lividans* Δpmt . La flecha de color azul claro muestra las placas del fago $\phi C31$ y las flechas de color azul oscuro las placas formadas por $\phi M1$.

En la figura 4 se observa que el fago $\phi M1$ aislado en este trabajo fue capaz de tener su ciclo de replicación normal en la cepa de *S. lividans*, mientras que en la cepa M145 (silvestre) de *S. coelicolor* fue incapaz de formar placas. El mismo resultado se observa con dos variantes del fago $\phi C31$, una virulenta (marcada como $\phi C31$ en la figura 4) y una lisogénica (KC857) que se incluyeron como controles. En la figura 4 también es posible observar como otros dos fagos aislados a partir de muestras de suelo y purificados (marcados como M2 y B3) son insensibles al sistema Pgl, pues son capaces de formar placas tanto en *S. lividans* como en la cepa silvestre M145 de *S. coelicolor*.

El gen SI3509 o *pmt* de *S. lividans* codifica una protein-manosil transferasa (Pmt), la cual une covalentemente el primer residuo de manosa en el proceso de glicosilación de proteínas. Por lo tanto, la presencia de esta proteína es indispensable para la infección de $\phi C31$ (Cowlshaw D & Smith, M., 2001). El fago

ϕ M1 fue incapaz de formar placas en las cepas Δpmt de *S. lividans*, pero sí formó placas en la cepa silvestre, lo que mostró el requerimiento del receptor glicosilado para el reconocimiento del fago ϕ M1.

De todos los fagos aislados a partir de diferentes muestras de suelo de México, sólo este fago, Φ M1, fue el que mostró los dos fenotipos similares al fago ϕ C31.

Purificación de DNA del fago Φ M1

Se purificó el DNA del fago Φ M1, ya que presentaba los dos fenotipos de interés. Con el fin de determinar que este fago no fuera un reisolamiento del mismo fago ϕ C31 se determinó digerir el DNA con enzimas de restricción y así comparar los fragmentos resultantes de la digestión respecto al patrón esperado del fago ϕ C31.

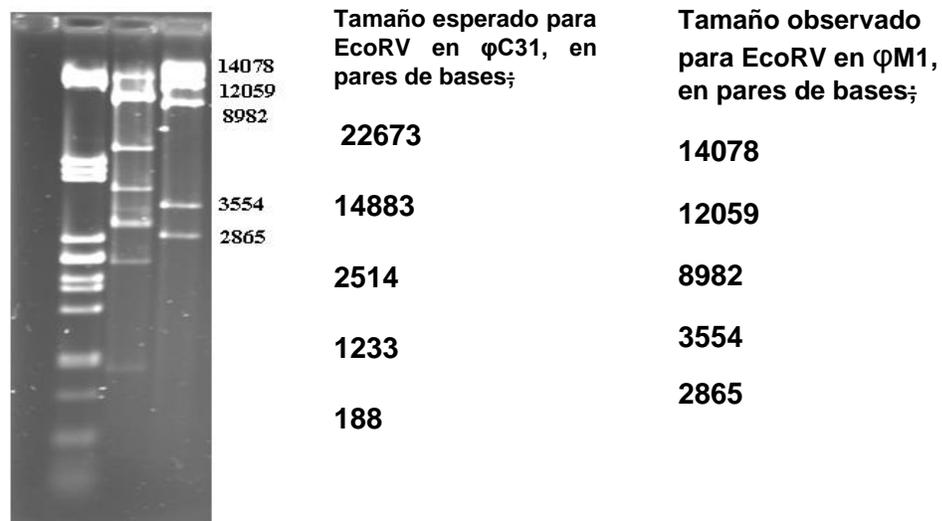


Figura 5. Digestión del DNA del fago Φ M1 con enzimas de restricción. Las digestiones se sometieron a electroforesis en un gel de agarosa al 0.8% teñido con bromuro de etidio. Carril 1, Marcador de peso molecular (DNA del fago lambda digerido con PstI); carril 2, DNA del fago ϕ M1, digerido con EcoRI; Carril 3, DNA del fago ϕ M1, digerido con EcoRV.

Los resultados anteriores muestran que el fago aislado no era igual al fago ϕ C31, ya que los patrones de restricción con las enzimas utilizadas no correspondieron a los fragmentos generados al digerir con diferentes enzimas de restricción (EcoRI y EcoRV).

Obtención de cepas lisógenas para el ϕ M1

Los fagos KC515, KC516 y KC857 son derivados no virulentos de ϕ C31. En la formación de la lisogenia, la proteína represora del fago se une específicamente a operadores, lo cual bloquea la expresión de genes líticos. Además, el represor es responsable de la inmunidad de una lisógena a una superinfección por el mismo fago u otro fago temperado cercanamente relacionado. Para determinar si el Φ M1 y el ϕ C31 se encuentran emparentados, se obtuvieron cepas lisógenas de ambos fagos y se probó la capacidad de infección de los fagos de infectarlas.

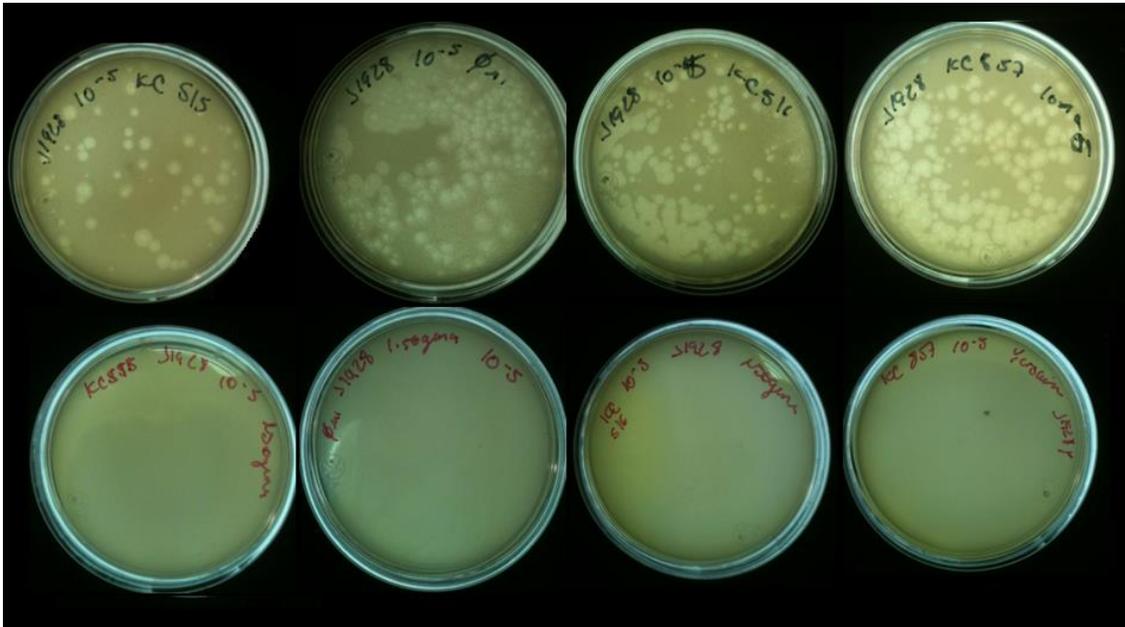


Figura 6. Ensayo de formación de placas sobre cepas lisógenas de ϕ M1. Arriba; formación de placas por los fagos KC515, ϕ M1, KC516, KC857 en la cepa no lisogénica J1928 de *S. coelicolor*. Abajo, ausencia de placas de KC515, ϕ M1, KC516, KC857 en la cepa lisógena J1928(ϕ M1).

Se obtuvieron cepas lisogénicas de *S. coelicolor* J1928 con los fagos ϕ M1 y con los derivados de ϕ C31 (KC515, KC516, KC857), según se indica en Materiales y métodos. En la figura 6 se observa que ninguno de los fagos probados fue capaz de formar placas sobre la cepa lisogénica J1928(ϕ M1), y en la figura 7 se muestra que el fago Φ M1 fue incapaz de formar placas en la cepa lisógena J1928(KC857). Adicionalmente se obtuvieron derivadas de J1928 lisogénicas para los fagos KC515 y KC516 y se repitió la misma observación: ninguno de los derivados de ϕ C31 ni el Φ M1 fueron capaz de formar placas en las lisógenas. Este resultado indica que el fago ϕ M1 en estado lisogénico da inmunidad a la infección por ϕ C31 y sus derivados. y viceversa. También se obtuvieron lisógenas de los diferentes fagos en *S. lividans* 1326 y se obtuvieron los mismos resultados que en *S. coelicolor* J1928. Esto sugiere fuertemente una relación estrecha entre ambos fagos, al menos en lo que respecta al represor.



Figura 7. Ensayo de formación de placas del fago ϕ M1 en las cepas J1928 y J1928(KC857).

Con el fin de determinar el fondo genético de las cepas lisógenas y así confirmar el resultado de la incapacidad de formar placas debida a la estrecha relación entre los fagos ϕ M1 y ϕ C31, se creció a la cepa lisógena de J1928(ϕ M1) en medio MS con el antibiótico estreptomycinina [30 μ g/mL], figura 8. La cepa J1928 contiene un gen de resistencia a estreptomycinina, sin embargo, la cepa M145 es sensible a estreptomycinina. Como se puede observar en la figura 7 el fenotipo observado (no formación de placas en las lisógenas) se debe a la incapacidad de una superinfección en la cepa J1928(ϕ M1) y ocurre muy probablemente por acción del represor, y no por ser una cepa contaminante derivada del fondo genético de la cepa M145, que contiene un sistema de restricción Pgl funcional y en la que los fagos derivados del ϕ C31 no forman placas.

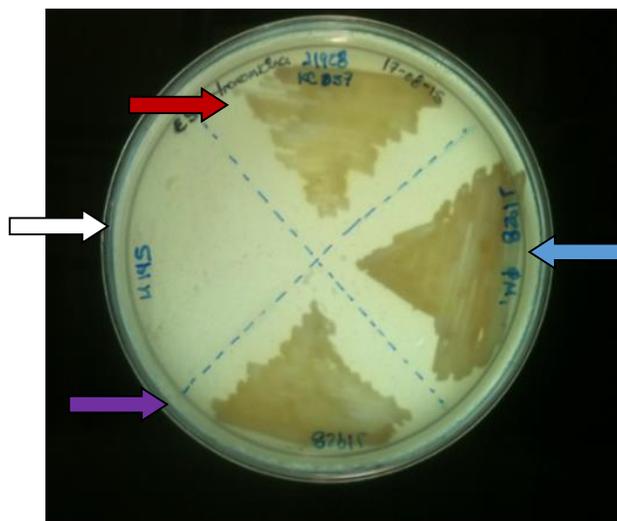
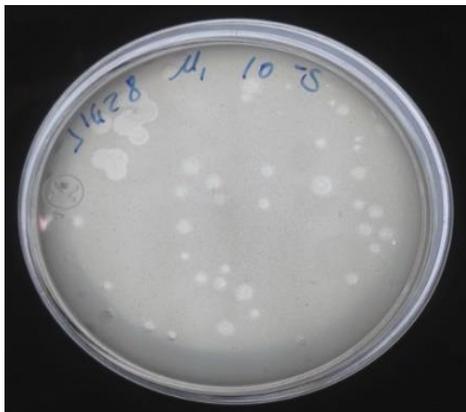


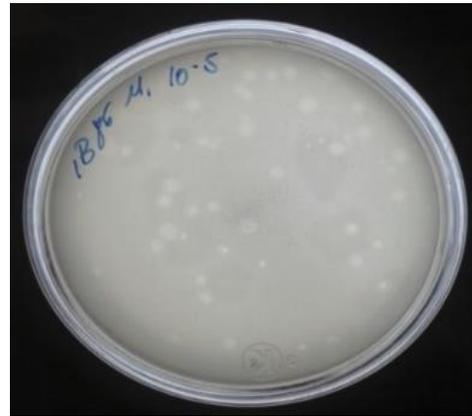
Figura 8. Confirmación del fondo genético de las cepas lisógenas derivadas de J1928. J1928 (rojo), J1928(ϕ M1) (azul), J1928(KC857) (Morado), M145(Blanco).

Eficiencia de formación de placas por el fago ϕ M1 en la cepa IB86.

El gen sco3902 codifica para una manosil-transferasa que aparentemente está involucrada en la adición de residuos de oligomanosa a las proteínas secretadas que ya han sido manosiladas en residuos de serina y treonina por la manosil-transferasa Pmt (Arroyo, 2017). La eficiencia de formación de placas del fago Φ C31 se ve afectada en la cepa IB86 de *S. coelicolor*, la cual contiene una mutación nula en el gen sco3902. En las proteínas glicosiladas de la cepa IB86 sólo existe una manosa enlazada covalentemente en cada sitio de glicosilación, ya que es mutante en las enzimas que colocan las posteriores manosas en las proteínas glicosiladas. Dado que el receptor del fago ϕ C31 es una glicoproteína, esto afecta la eficiencia de formación de placas.



Formación de placas de ϕ M1 en J1928



Formación de placas de ϕ M1 en IB86



Formación de placas de ϕ C31 en J1928

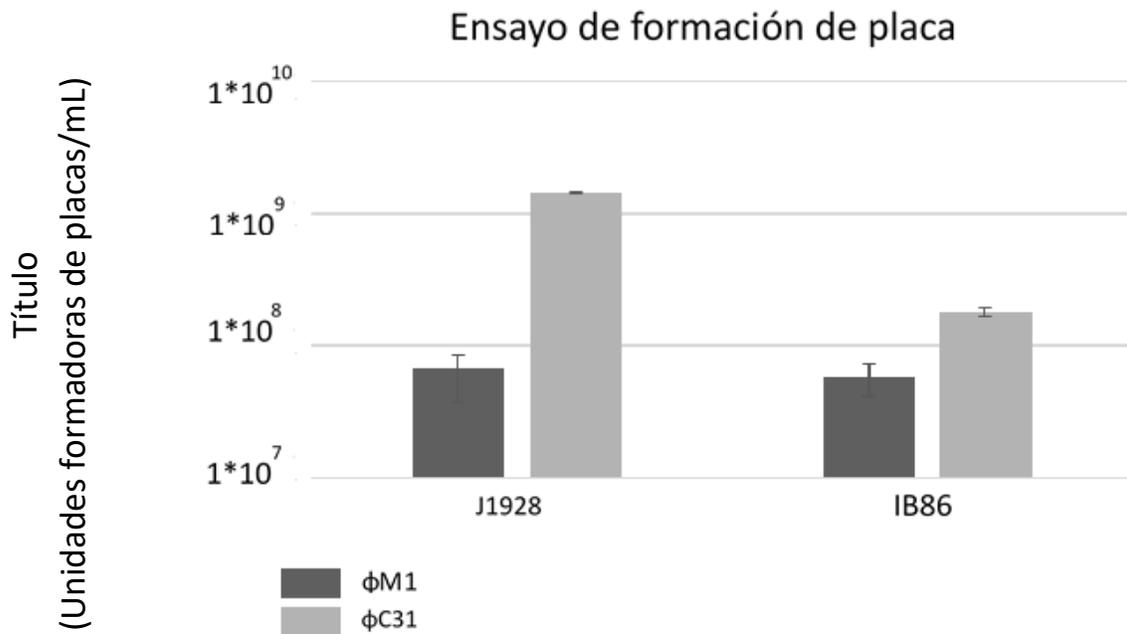


Formación de placas de ϕ C31 en IB86

Fig 9. Ensayos de formación de placa de los fagos en las cepas IB86 (sco3902 nulo) y J1928.

En la gráfica 1 se observa que el número de placas formadas en la cepa IB86 por ϕ C31 es aproximadamente un orden de magnitud menor que el número de placas formadas en la cepa cepa J1928. También se puede observar que el número de unidades formadoras de placas del fago ϕ M1 en la cepa IB86 y J1928 se encuentran en el mismo orden de magnitud. Lo cual significa una diferencia entre los fagos ϕ C31 y ϕ M1.

Se ha reportado que los genes putativos para proteínas de la cola de ϕ C31 son g44-g49 (Smith, 1999), y que una mutación sin sentido, en la proteína gp44, puede cambiar el rango de hospederos en fagos derivados de ϕ C31 y revertir la incapacidad del ϕ C31 de infectar a una mutante de *S. coelicolor* carente de Pmt (Cowlshaw, 2001). Las secuencias de residuos de aminoácidos de estas proteínas en ϕ M1 no son 100% idénticas a las de ϕ C31, lo que podría explicar la diferencia en el reconocimiento del receptor (ver más abajo). Por ejemplo, la proteína gp49 de ϕ M1, cuenta con el 86.83 % de identidad con la proteína homóloga en ϕ C31 (gp19), porcentaje más bajo de identidad entre las proteínas de la cola con respecto a ϕ C31. La proteína gp49 codifica para la fibra de la cola, la cual en muchos fagos realiza el primer contacto con el hospedero. También puede deberse a las modificaciones en la secuencia de aminoácidos de la proteína gp14, que también está involucrada con el reconocimiento del receptor.



Grafica 1. Eficiencia de en la formación de placas de los fagos ϕ C31 y ϕ M1 en las cepas J1928 e IB86 (ANOVA, $p < 0.05$).

Ensamblaje del genoma de ϕ M1

Los genomas de Φ M1 (purificado de *S. albus* J1074) y de la cepa lisógena J1928(ϕ M1) fueron secuenciados en la Unidad Universitaria de Secuenciación Masiva de DNA de la UNAM.

Las secuencias cortas de nucleótidos (lecturas) de 150 pb, se ensamblaron en el programa Abbyss (Simpson, 2009). En el caso del ensamblaje de la cepa lisógena se realizó una búsqueda de homología mediante BLAST utilizando la secuencia de ϕ C31 con el objetivo de determinar el contig en el cual se encontraba el genoma del fago ϕ M1. El resultado de la búsqueda determinó que un *contig* con tamaño de 520,669pb contenía una secuencia homóloga a la del ϕ C31. Se utilizó el genoma de ϕ C31 como secuencia de búsqueda, porque se sabía que era un genoma cercano al del fago ϕ M1.

Posteriormente esta secuencia obtenida (*contig_520,669pb*) se utilizó como la secuencia de referencia, y se alinearon las lecturas individuales utilizando el programa Bowtie, en el cual se pueden comparar grandes conjuntos de secuencias cortas a una secuencia de referencia (Langmead, 2010), para así determinar variaciones o cambios en la secuencia de referencia.

Del mismo modo se ensamblaron las lecturas del genoma del fago ϕ M1, con los programas Abbyss y Velvet, obteniendo un solo *contig*, el cual se analizó y comparó con los resultados obtenidos en la secuencia de la lisógena de ϕ M1. De esta forma se obtuvo la secuencia definitiva del genoma del fago.

Características del genoma de ϕ M1

El tamaño del genoma del fago Φ M1 es de 41,538 pb. Presenta un porcentaje de G + C de 63.6%, menor al de *Streptomyces coelicolor* (72.3%; Bentley et al., 2002). El genoma se anotó en el servidor RAST (Rapid Annotation using Subsystem Technology). Se predijeron 53 genes codificantes de proteínas y un gen de tRNA, utilizando Argón y tRNAscan-SE. Las secuencias de aminoácidos de los ORFs fueron comparadas para encontrar similitud con secuencias presentes en las bases de datos disponibles (figura 10). Además, se les asignaron funciones putativas a las proteínas predichas, con base en la similitud de la secuencia o la presencia de dominios.

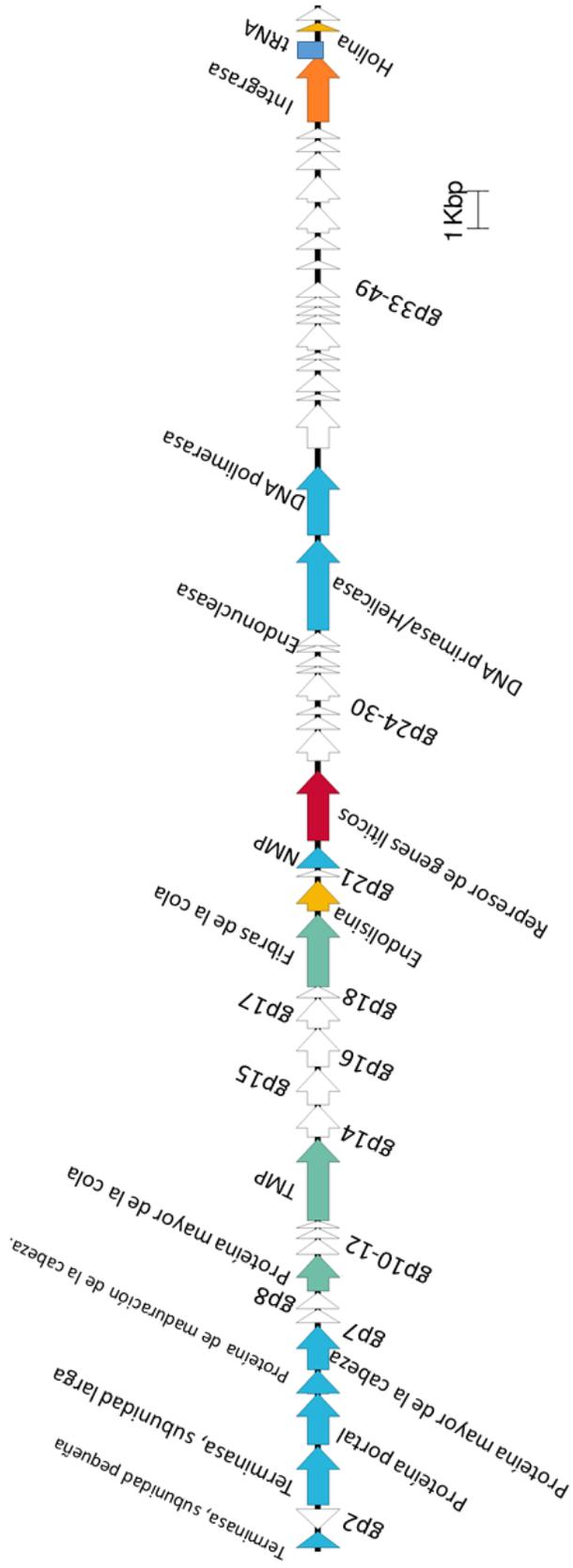


Figura 10. Organización del genoma de ϕ M1. Las funciones putativas fueron marcadas con el nombre. Los colores de las flechas indican el módulo funcional de los genes dentro del genoma, Azul claro: empaquetamiento/ cápside. Verde: Cola/ fibras de la cola. Amarillo: Lisis. Rojo; Regulación. Azul marino: DNA/ temprano. Naranja: Integrasa. Rectángulo azul: tRNA.

Componentes del genoma de Φ M1.

Para asignar funciones a los productos génicos que surgen de los ORFs predichos, se interrogaron bases de datos para genes homólogos a los supuestos ORFs del fago ϕ M1. Las búsquedas se realizaron a nivel de nucleótidos, así como a nivel de aminoácidos.

El genoma de ϕ M1 se puede dividir en cinco grupos provisionales de genes con funciones relacionadas. Los grupos definidos se discutirán con más detalle.

Lisogenia y regulación. El orf23 posiblemente está involucrado en el establecimiento de la lisogenia, ya que codifica para una proteína 80% idéntica a la codificada por el gen "c" de ϕ C31. La región altamente conservada se encuentra en el extremo C-terminal, el cual es un motivo de unión a DNA HTH e interacciona con las secuencias invertidas repetidas (CIRs) localizadas en 16 sitios dentro del genoma (Ingham,1994; Wilson,1995). Esta región también se encuentra conservada en los fagos ϕ BT1 y TG1, siendo el extremo N-terminal más variable.

El orf51 se propone como el gen codificante de la integrasa. A diferencia del fago λ (*E. coli*) y Mu1/6 (*S. aureofaciens*), se encuentra del lado derecho del mapa génico, orientación común en la mayoría de los fagos. Muestra una identidad del 99% con la proteína codificada por ϕ C31, 51% con ϕ TG1 y 25% respecto a ϕ BT1. No presenta identidad con otra integrasa de otros fagos de *Streptomyces*. La secuencia de la proteína gp51 revela similitud a la familia de integrasas de serina, con un mecanismo similar al que ocurre en la familia invertasa, resolvasa. Estudios *in vitro* de la integrasa de ϕ C31 muestran que la recombinasa provoca recombinación integrativa con el *attP* y *attB*, pero no recombinación escisiva. (Kuhstoss,1991).

A pesar de que existe una alta similitud de secuencia en el dominio catalítico N-terminal de la integrasa de TG1, ésta reconoce sitios distintos *attP* y *attB* de la integrasa de attachment (Kentaro, 2009). Además, también, los sitios de reconocimiento *attP* y *attB* del fago ϕ BT1 son distintos a los reconocidos por la integrasa de ϕ C31.

La secuencia del genoma de la cepa J1928 lisógena de ϕ M1, permitió determinar los sitios de unión o "*attachment*" del cromosoma (*attB*) y del fago (*attP*). La secuencia reportada del sitio *attB* en el cromosoma de *Streptomyces coelicolor* se buscó en la secuencia de la cepa lisógena de Φ M1, y se alineó con el *attB* de *S. coelicolor* (Fig. 9.D). El sitio *attP* se buscó en la cercanía del gen *Int* (el cual es una recombinasa de sitio específico) del genoma de ϕ M1 (figura 11.A) y se comparó con la secuencia de la cepa lisógena, y se determinó el *attP* de ϕ M1 (figura 11.B). El sitio de entrecruzamiento se identificó como un dinucleótido 5'-TT en el centro de una secuencia repetida inversa imperfecta (figura 12). El sitio de "*attachment*" es idéntico en secuencia al del fago ϕ C31.

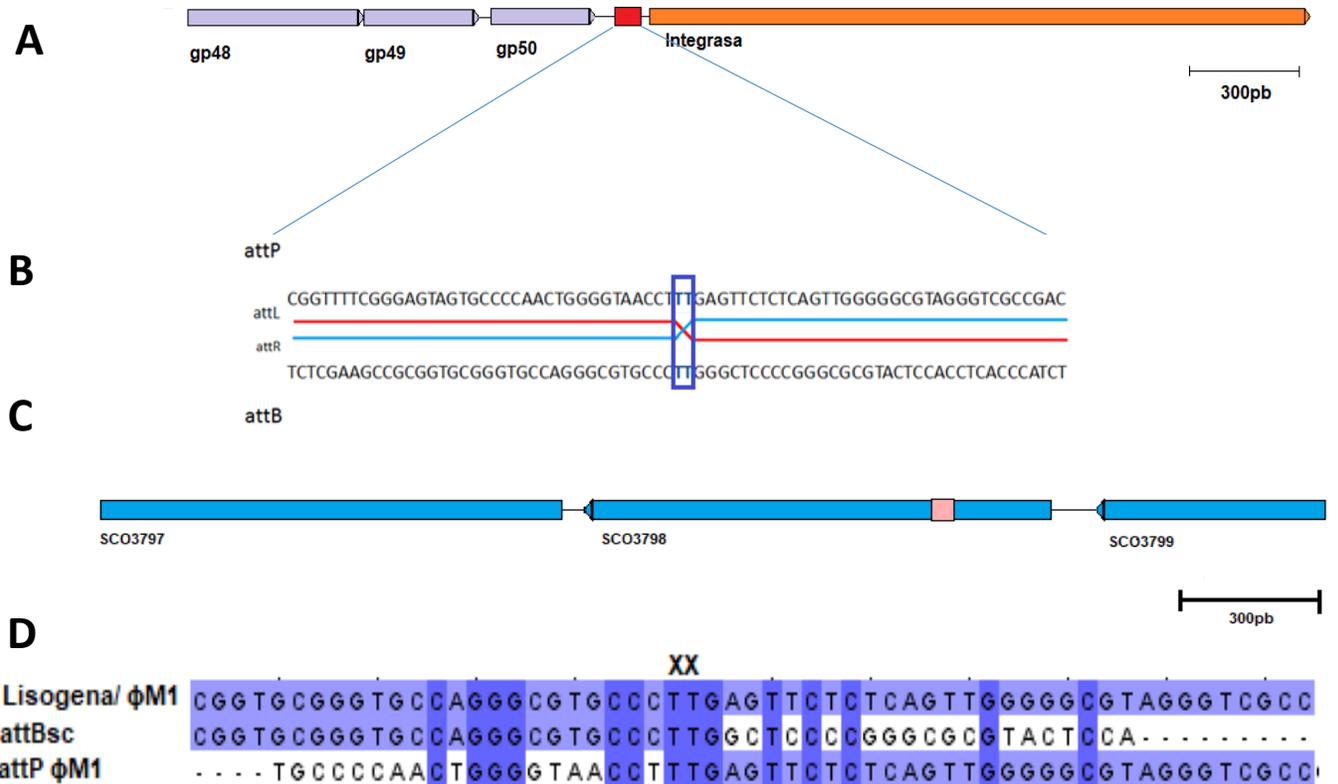


Figura 11. Representación de los sitios *attB* de *Streptomyces coelicolor* y *attP* de ϕ M1. A) Localización en el genoma de ϕ M1 el sitio *attP* (recuadro rojo), se encuentra “Upstream” del gel *int*. B) La primera secuencia de nucleótidos representa el sitio *attP* del fago ϕ M1 y la última secuencia representa el sitio *attB* del fago ϕ M1. La segunda y tercera línea muestran esquemáticamente las secuencias recombinadas, *attR* y *attL*, que corresponden a mitades del sitio *attP* ϕ M1 (*attP*-L y *attP*-R) mostrado por una línea naranja, y el sitio *attB* ϕ M1 (*attB*-L y *attB*-R) mostrado por una línea azul. La secuencia de cruce es un dinucleótido TT el cual está representado por un rectángulo morado. C) Se representa la localización en el genoma del sitio *attB* ϕ M1, el cual se encuentra en el gen *sco3798*. D) Alineamiento de las secuencias *attB* ϕ M1, *attP* ϕ M1 y la secuencia de la cepa lisogénica de ϕ M1. La secuencia de entrecruzamiento se representa con una doble X.



Figura 12. Diagrama del sitio *attP*. Se muestran los dinucleótidos centrales (morado), y las secuencias repetidas invertidas imperfectas (flechas negras).

Se determinó que el sitio de integración al cromosoma en *S. coelicolor* del fago ϕ M1 es el mismo reconocido por ϕ C31. El sitio de integración en el genoma ocurre dentro del gen *sco3798* (Fig. 11.C), el cual codifica para una proteína de la familia de las pirinas, las cuales están conservadas tanto en procariontes como en eucariontes y que no es esencial para el crecimiento de *S. coelicolor* (Combes, 2002).

Replicación. Entre las proteínas codificadas con homología reconocida entre los genomas de fagos de *Mycobacterium* y *Streptomyces*, se encuentran las asociadas a la replicación del DNA (primasas y DNA polimerasas).

Los productos de los orfs 32 y 33 putativamente se encuentran involucrados en la replicación. ϕ M1 codificaría una DNA polimerasa I (orf33), enzima catalizadora de la replicación de DNA. La identidad entre las proteínas de los fagos de *Streptomyces* se encuentra entre el 55% de ϕ SASD1 hasta 84% respecto a ϕ C31. Sin embargo, existen fagos de *Streptomyces* que sólo codifican la subunidad épsilon de la DNA polimerasa como el fago VWB (*Streptomyces venezuelae*).

La proteína gp34 muestra un dominio de región 2 del factor sigma 70: Esta proteína se encuentra conservada entre los fagos de *Streptomyces*, y no presenta homología con las RNA polimerasas mejor estudiadas en fagos, como las del fago T7 (*E. coli*) o ϕ KMV (*Pseudomonas aeruginosa*).

El producto del orf32 tiene homología con dominios de helicasas de la superfamilia 3, DNA primasa, y dedos de Zinc. No muestra dominios de unión a ATP. El dominio de DNA primasa se encuentra conservado en fagos de *Streptomyces* tales como: ϕ C31, 83%; TG1, 79%; ϕ BT1, 77%; ϕ SASD1, 51%; y algunos fagos de *Mycobacterium*. En el dominio de helicasa además de la homología con actinofagos se puede observar relación con fagos de *Bacillus*, fago BM5, 36%; *Lactobacillus* fago A2, 39%; *Xylella* fago Xfas53, 34%.

Un aspecto a considerar es que los actinofagos poseen los dos dominios (helicasa y primasa) en una sola proteína, a diferencia de otros fagos como T4 de *E. coli* donde están en proteínas separadas. El fago T7 produce una proteína con ambas actividades (helicasa-primasa), y en algunas ocasiones puede presentar una proteína con el extremo N-terminal trunco, en la que solo presenta actividad de helicasa (Zhang, 2012). Sin embargo, fagos como P4 (*E. coli*), presentan ambas actividades en una proteína multifuncional, ésta proteína muestra homología con proteínas tipo D5 al igual que la primasa de los fagos de *Streptomyces*. Esto demuestra las diferentes estrategias de replicación de los bacteriófagos.

El orf39 se predice como una timidilato sintasa (ThyX). Las timidilato sintasas pueden ser de dos tipos: ThyA o ThyX, la cual representa una clase alternativa. Ambas enzimas catalizan la conversión de dUMP a dTMP, pero el mecanismo reductor es diferente (Bhattacharya, 2008). Aproximadamente el 30% de las especies microbianas (basado en secuencias de genomas completos), dependen de ThyX, mientras que la formación canónica de dTMP, vía ThyA está presente en 70% de los microorganismos. A diferencia de *Mycobacterium* que codifica ThyA y ThyX, *Streptomyces* solo presenta la enzima ThyX, al igual que sus fagos relacionados. Solo los actinofagos ϕ C31 y ϕ TG1 codifican para la enzima ThyX.

La proteína gp43 tiene homología con la enzima DCMP desaminasa. Esta proteína está involucrada en la transformación de dCMP a dUMP. Estos genes se distribuyen esporádicamente entre los genomas de los fagos, y sus ubicaciones no muestran una relación consistente con los otros genes en los genomas en los que aparecen (Smith, 2013). Sin embargo, sí se conserva la sintenia entre ϕ M1 y ϕ C31. Además, todos los fagos del clúster- ϕ C31 (ϕ C31, ϕ M1, ϕ BT1, TG1) y clúster-R4 (R4, EIB20, lika, sujidade, ϕ Hau3) presentan este gen.

Morfogénesis. Los productos de los ORFs *gp1* y *gp3* poseen homología con proteínas estructurales en fagos, y particularmente codifican la subunidad menor y mayor de la terminasa (respectivamente). En los fagos de *Streptomyces* es común la localización del sitio *cos*, cerca de los genes de la terminasa.

La subunidad menor se une específicamente al DNA viral y la subunidad mayor tiene actividad de endonucleasa para la escisión de DNA y actividad de ATPasa que mejora el empaquetamiento del DNA (Bukosvska, 2006). La secuencia de residuos de aminoácidos muestra gran identidad con las mismas proteínas de fagos homoinmunes de ϕ C31.

El producto del ORF gp4 fue anotado como la proteína portal la cual es una proteína importante involucrada en el transporte de DNA durante el empaquetamiento. Además, los genes de la terminasa y portal son aparentemente adyacentes en fagos de diferentes orígenes, tales como ϕ C31, muchos mycobacteriofagos, algunos fagos de *Streptococcus thermophilus* y fagos lambdaoides tales como HK022 y HK97 (Fokine, 2016)

Muchos virus grandes, incluyendo los bacteriófagos de dsDNA con cola y herpesvirus, ensamblan su cápside por medio de la formación de precursores, llamados procápsides. La procápside normalmente contiene una proteasa que es activa durante la maduración de la cápside, y destruye el centro y libera espacio para el material genómico (Fokine, 2016). El producto del ORF5 fue identificado como una proteinasa procápside putativa. Solo los fagos del clúster ϕ C31 codifican para esta proteína. Otra proteína de esta región (gp6) muestra similitud a la proteína mayor de la cápside, la cual está involucrada en la estabilización de la condensación del DNA en la cápside de los fagos.

En la cercanía del gen de la integrasa de ϕ M1, se encuentra el gen para una proteína con motivos HNH de endonucleasa (gp53). Este cuenta con los

aminoácidos conservados entre proteínas HNH asociadas a actividad de terminasa. Entre éstas se encuentra la del fago HK96 de *E. coli*, cuya proteína es necesaria para la función de la enzima terminasa. El rol en la función de la terminasa está extendido en los fagos con cola (Kala, 2014).

Ensamblaje de la cola. La región de gp13-gp19 está probablemente involucrada en esta función. El orf13 es el segundo con mayor número de nucleótidos. Codifica a una proteína putativa TMP (*tape measure protein*), la cual está involucrada en la determinación del tamaño de la cola (Farkasovska, 2007). Katusta (1984), propuso una relación entre la longitud de la secuencia de residuos de aminoácidos y el tamaño en nanómetros de la TMP, donde un aminoácido equivale a 0.15 nm. La secuencia de aminoácidos cuenta con 730 residuos, lo que significaría un tamaño de 109nm, cercano al establecido en ϕ C31 por microscopía electrónica, de 100 nm (Suarez, 1985). La predicción de la estructura secundaria muestra que 657 aa se asignaron como alfa hélices por análisis de HNN; esta conformación se encuentra en varias TMP de fagos. Estas proteínas a menudo tienen un número variable de repeticiones en tándem con aminoácidos altamente conservados de triptófano (W) y fenilalanina (F) en posiciones fijas que se usan como anclajes por proteínas auxiliares pequeñas las cuales se asocian al ensamblaje de la cola (Belcaid, 2011). La proteína gp11 se predice como una proteína chaperona de la cola y posiblemente se encuentre involucrada en la función antes mencionada). Aunque la secuencia de aminoácidos entre la proteína TMP de ϕ M1 y ϕ C31 presenta una identidad del 96% de identidad, las repeticiones en *tándem* no son idénticas, ni se localizan en la misma posición de la secuencia (Se utilizó el programa MEME para observar repeticiones).

La proteína gp19 cuenta con motivos repetidos de colágeno, y está relacionado con las fibras de la cola. Los orfs 7-13 probablemente estén involucrados en la formación y conexión de la estructura de la cabeza y cola. Ya que en muchos fagos los genes localizados entre la proteína mayor de la cabeza y la MTP cumplen esa función (Brosend, 2001).

En fagos de DNA de doble cadena, el orden de los genes relacionados a la cola se encuentra altamente conservado, notablemente la proteína MTP (*major tail protein*), normalmente se encuentra upstream del gen codificante de la TMP. Además, entre estos dos genes, existen típicamente 2 ORFs sobrelapados, los cuales están relacionados por un “frameshift programado” o corrimiento del marco de lectura programado (Xu, 2004).

Este corrimiento de lectura programado es un fenómeno asociado a la traducción, en donde un ribosoma detenido usualmente se desliza +1 ó -1 y continúa la traducción del mensaje resultando en una proteína con un extremo C-terminal afectado (Harger,2002). El *frameshift* programado -1, posee una secuencia “slippery” o “resbaladiza” heptámera característica XXXYYYZ (Y= A o T), aunque en ϕ M1, ϕ C31, L5, y Mu esta secuencia no es canónica. La secuencia “slippery” (GGGAAG) en ϕ M1, se encuentra en los ORFs11 y12, los cuales muestran

homología con chaperonas de la cola. Estos ORF están entre la proteína gp10 (*major tail*) y g13 (*tape measure protein*).

Genes de lisis. Algunos fagos de dsDNA realizan la lisis del hospedero a través de la acción sinérgica de un sistema de lisis con dos componentes: una endolisina que degrada la pared celular y una holina, que permeabiliza la membrana. El resultado de este proceso es la degradación de la pared celular y una subsecuente lisis (Wang, 2010). A pesar de la heterogeneidad de las secuencias, la mayoría de las holinas comparten características comunes; los genes de las holinas se localizan inmediatamente *upstream* de los genes de endolisinas, las proteínas tienen una alta hidrofobicidad dos o tres dominios transmembranales y un doble motivo de inicio traduccional (Bläsi, 1996)

La proteína gp52 se asignó como una holina; es una proteína pequeña de 78 aa, conservada en muchos fagos de *Streptomyces*, presenta dos dominios transmembranales. Sin embargo, el gen de la endolisina no se encuentra adyacente, lo cual tampoco ocurre en los fagos de *Streptomyces*, pero esta organización de los genes relacionados a la lisis, sí se presenta en fagos de *Mycobacterium*.

El orf20 fue predicho como una endolisina putativa, la cual se encuentra conservada en el clúster ϕ C31, sin embargo, presenta homología con otros fagos como R4 y ϕ SASD1.

Comparación de la estructura genómica de los fagos ϕ C31 y ϕ M1.

El análisis de los genomas de ϕ C31 y ϕ M1, reveló 54 y 53 ORFs putativos, respectivamente. Los ORFs en ϕ C31 son numerados consecutivamente excepto el gen del represor "C" (31-53, C,1-30) (Smith *et al.*, 1999.) Los genes de ϕ M1 fueron numerados del 1-53. En ambos genomas se encuentra un tRNA. Un total de 50 ORFs del genoma de ϕ M1 muestran niveles altos de similitud con ORFs conocidos, pero únicamente a 22 ORFs se les asignó una función putativa. Ambos genomas se encuentran organizados modularmente, con los genes en el orden: empaquetamiento, cola y proteínas de la cola, regulación, síntesis y replicación de DNA, regulación, región hipervariable e integrasa (Fig. 13.B).

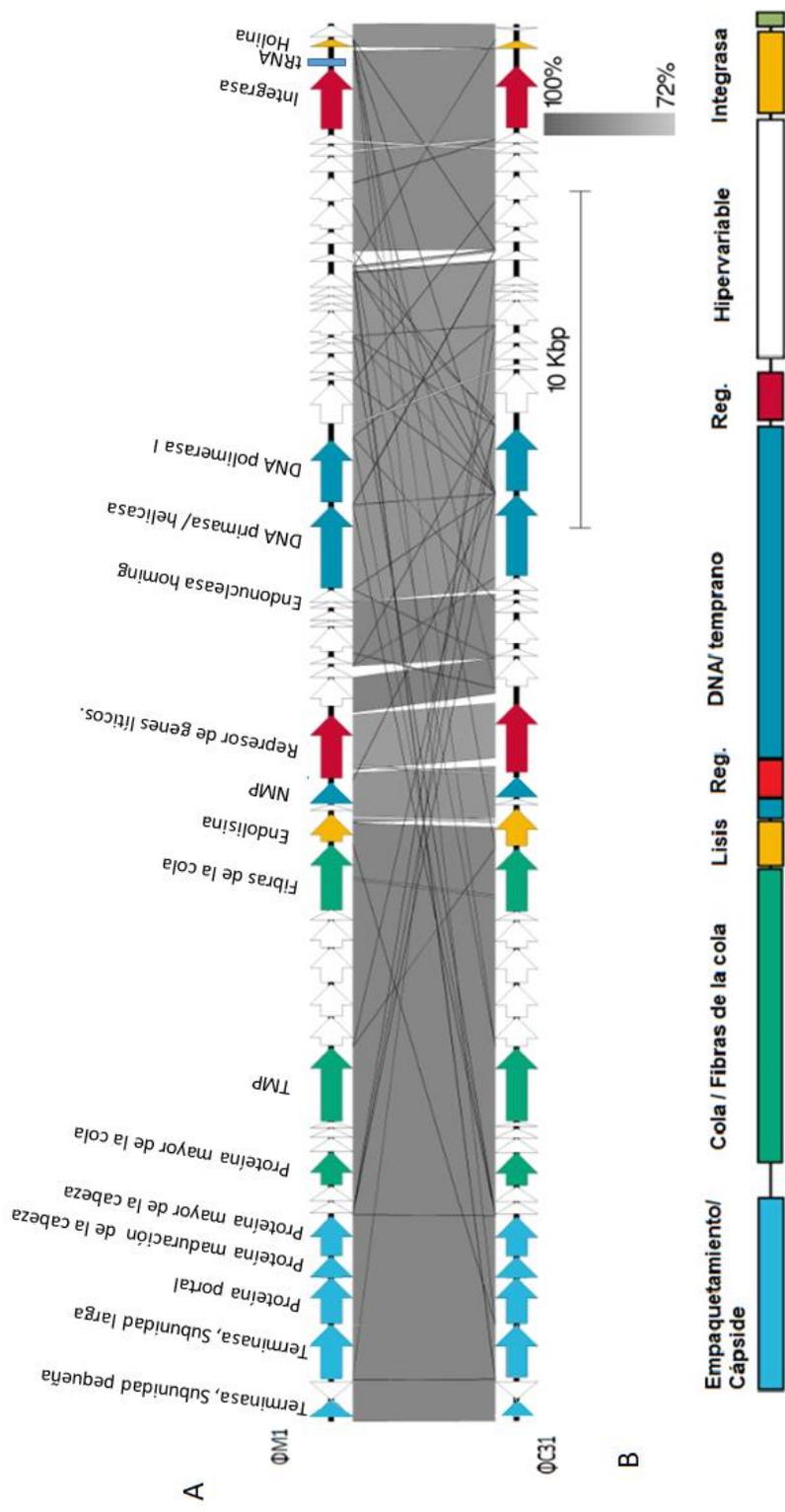


Fig. 13 Fagos ϕ M1 y ϕ C31. A) Las funciones putativas fueron marcadas con el nombre. B) Organización de los genomas del clúster ϕ C31 (Modificada de Smith, 2013), los recuadros y sus colores indican el módulo funcional de los genes dentro del genoma de los fagos. Reg=regulación.

El alineamiento de los genomas de ϕ C31 y ϕ M1 revela un alto nivel de similitud, con regiones que van de una alta identidad del 100% hasta regiones variables con menor identidad (68%). Los genomas de ϕ C31 y ϕ M1 tienen una identidad promedio de 89.42% a nivel de nucleótidos. De los 53 ORFs predichos en el genoma de ϕ M1, 50 mostraron similitud con proteínas predichas del genoma del fago ϕ C31, con la excepción de los que codifican para las proteínas gp25, gp38 y gp44.

Uno de los ORFs no presentes en ϕ C31, que codifica, gp25, sí se encuentra en los fagos ϕ BT1 (82% de identidad) y ϕ SASD1 (52% de identidad).

La gp38 sólo muestra identidad con una proteína del fago ϕ TG1 (51%) donde está anotada como “*overlapping terminator*” ya que el inicio del gen para la gp38 se encuentra sobrelapado con el codón de término de la gp37. Esto no ocurre en ϕ C31, ni en ϕ BT1. La proteína gp37 de ϕ M1 presenta un dominio C-terminal Csp2_I-U (los miembros de esta familia de proteínas son del tipo Cas o proteínas asociadas a CRISPR). Esta proteína se encuentra conservada en el clúster- ϕ C31.

Por último, la proteína gp44 de ϕ M1 no muestra homología con ninguna secuencia en las bases de datos. Por otra parte, ϕ C31 contiene en su genoma, dos genes que no tienen homólogos en ϕ M1. La proteína gp40 y gp49 de ϕ C31 no presenta homología con ninguna proteína reportada en las bases de datos.

Se ha propuesto por Desiere et al. 1998, que la falta de similitud entre los ORFs de los fagos y la base de datos a nivel de DNA, y la identidad generalmente modesta a nivel de residuos de aminoácidos, se explica por: 1) la falta de genes homólogos secuenciados en la base de datos y 2) la acumulación de deleciones/ inserciones y mutaciones puntuales dentro de un gen después del evento de recombinación, lo que conduciría a una diversidad de genes. Con lo cual se explicaría la falta de homología de muchas de las proteínas de los fagos.

La comparación entre los genomas que se presenta en la figura 13 se construyó contando las sustituciones e inserciones y deleciones (“indels”) en ventanas consecutivas de 1,000 pb en un alineamiento de ϕ C31 y ϕ M1. Aunque se observan diferencias, principalmente sustituciones, las mayores similitudes ocurren en los primeros 17,000 nucleótidos, y posteriormente existe un mayor número de cambios (indels). En ϕ C31 la región de g15-g28 se considera una región innecesaria, pues deleciones de esta región no afectan la formación de placas (Smith, 1999). Es esa región en donde se presentan las mayores cantidades de diferencias entre los genomas. Lo mismo se observó en una comparación entre los genomas de los fagos R4 y ELB20 (Smiteh et al., 2013). Si bien existen sustituciones e indels, la estructura genética (figura 14), así como las señales regulatorias no se ven modificadas (figura 16).

También se puede observar que existen genes sin homología cercanos a genes con alta identidad, por ejemplo, el orf25 no está presente en ϕ C31, sin embargo, el orf24 y orf26 presentan identidades de 93% y 85% respectivamente.

Los resultados sugieren que los genomas de estos fagos son mosaicos uno con respecto al otro. Esta organización se ha reportado en diversos grupos de fagos, como los mycofagos D29 y L5 (Pelluda,2003)

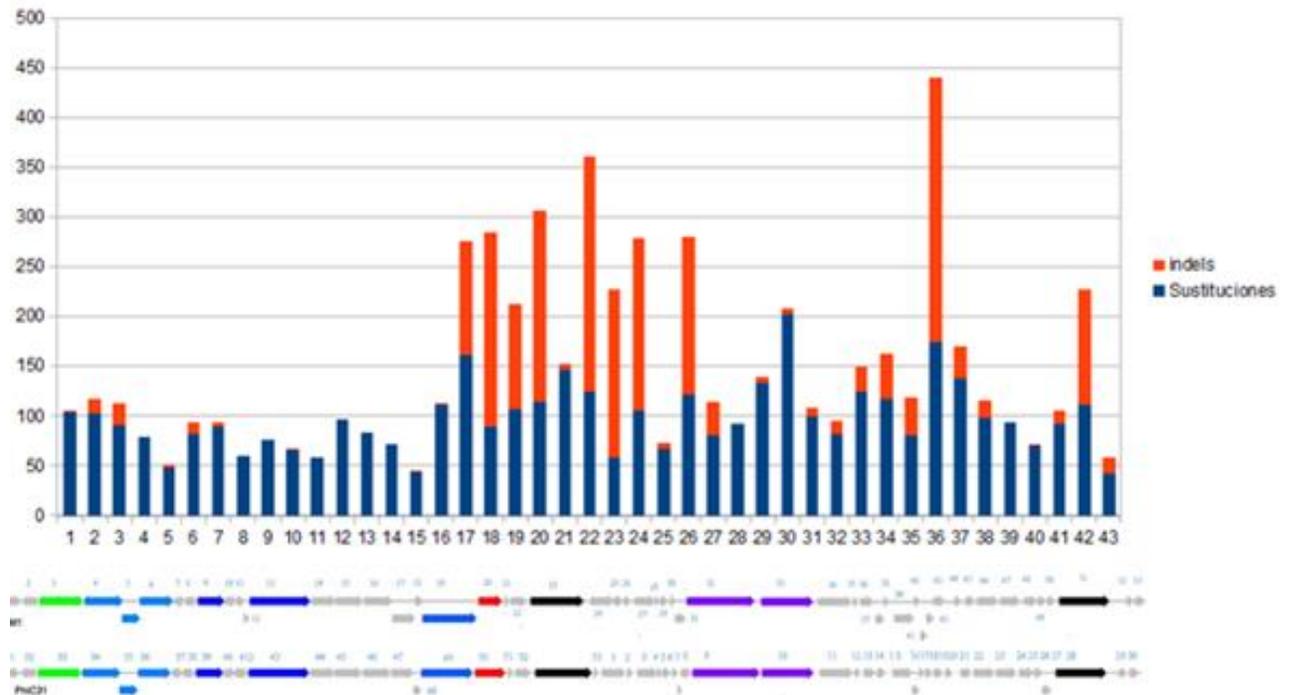


Figura 14. Número de sustituciones e “Indels” en ventanas consecutivas de 1,000 pb en un alineamiento de DNA, entre ϕ M1 y ϕ C31.

Los genomas de ϕ C31 y ϕ M1 tienen una identidad de 89.42% a nivel de nucleótidos. La gráfica (figura 14) se construyó contando las sustituciones e indels en ventanas consecutivas de 1,000 pb en un alineamiento entre Φ c31 y ϕ M1.

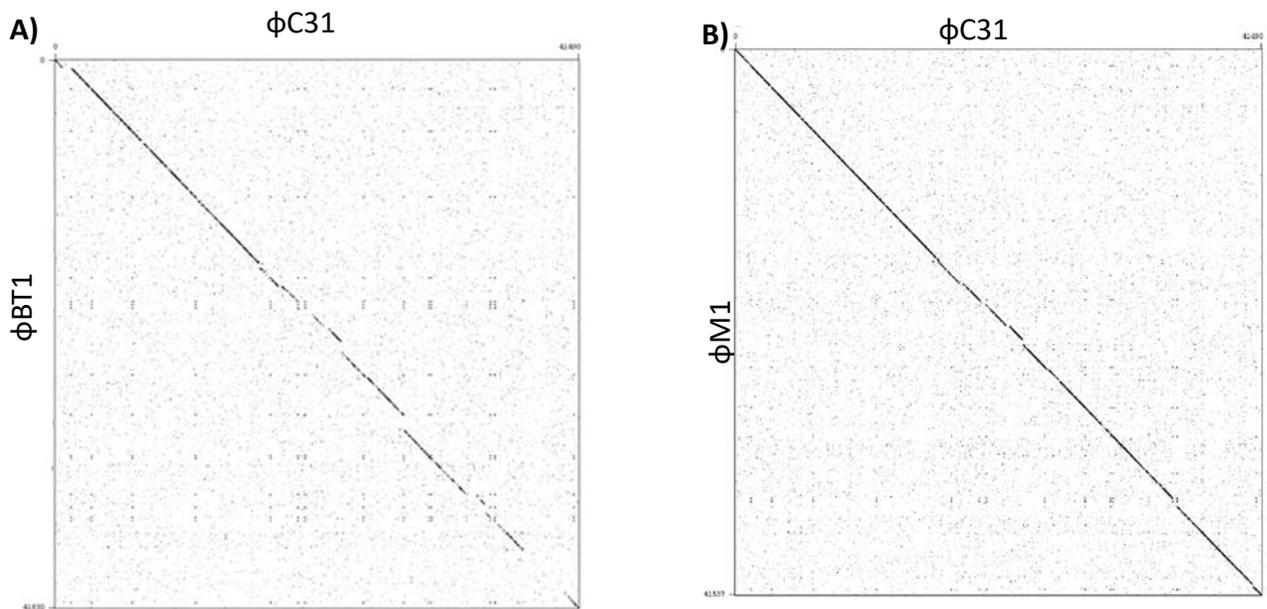


Figura 15. A) Dotplot entre secuencia de nucleótidos de ϕ BT1 y ϕ C31. B) Dotplot entre secuencia de nucleótidos de ϕ M1 y ϕ C31.

El análisis de matrices de puntos (dotmatrix) es un método rápido de comparación de secuencias de aminoácidos o nucleótidos. Los dotplots son el resultado gráfico de las matrices y son utilizados para el análisis de relaciones evolutivas como dominios conservados (Krumstiek, 2007). En la figura 15 se observan las comparaciones entre las secuencias de nucleótidos de ϕ BT1 y ϕ C31 y entre ϕ M1 y ϕ C31. Se observan módulos conservados, sin embargo, como se mostraba en la figura 14 (indels y sus), las mayores diferencias ocurren en las regiones centrales del genoma (figura 15.B). En cuanto al gráfico de ϕ C31 y ϕ BT1 (figura 15.A) se observan módulos conservados, sin embargo, tanto en la región central como en la terminal, existen secuencias sin homología. Es posible observar que la similitud entre ϕ M1 y ϕ C31 es mayor la similitud entre ϕ BT1 y ϕ C31. La Tabla 2 muestra cómo los fagos homoinmunes de ϕ C31 muestran diferentes porcentajes de identidad a nivel de DNA; también mediante este análisis es posible observar que entre todos estos fagos los más parecidos entre sí son ϕ M1 y ϕ C31.

Fago	ϕ M1	ϕ C31	ϕ BT1	TG1
ϕ M1	100			
ϕ C31	89.42	100		
ϕ BT1	77.24	77.70	100	
TG1	70.23	70.67	68.08	100

Tabla 2. Porcentaje de identidad a nivel de nucleótidos entre los fagos homoinmunes de ϕ C31.

Secuencias regulatorias conservadas en los fagos ϕ C31 y ϕ M1.

Los sitios de unión del represor (CIRs), los promotores fago específicos y los terminadores, se encuentran conservados en los fagos ϕ C31, ϕ BT1 y TG1 (Smith, 2013). Como se puede observar en la figura 16, que estas secuencias también se encuentran conservadas en el fago ϕ M1. Aunque existen deleciones o inserciones en el genoma, la posición de las señales regulatorias se encuentra conservada.

Secuencias CIR ϕ M1		Localización		1		18		Localización		1		18		Secuencias CIR ϕ C31	
	1223	CCTACTC	CACG	TGAGT	AGG				1223	CCTACTC	CACGT	GAGT	AGG		
	34054	CCTACTC	CACG	TGAGT	AGG				2848	CCTACTC	CACGT	GAGT	AGG		
	42000	CCTACTC	CACG	TGAGT	AGG				11089	CCTACTC	CACGT	GAGT	AGG		
	11134	CCTACTC	GCGG	TGAGT	AGG				17009	CCTACTC	CACGT	GAGT	AGG		
	35718	CCTACTC	GCGG	TGAGT	AGG				29547	CCTACTC	CACGT	GAGT	AGG		
	19357	CCTACTC	CACG	TGAGT	AGC				34431	CCTACTC	CACGT	GAGT	AGG		
	30231	CCTACTC	CACG	TGAGT	AGT				34781	CCTACTC	CACGT	GAGT	AGG		
	2848	CCTACTC	CACG	CGAGT	AGG				34820	CCTACTC	CACGT	GAGT	AGG		
	17174	CCTACTC	CACG	CGAGT	AGG				41031	CCTACTC	CACGT	GAGT	AGG		
	19936	CCTACTC	CACG	CGAGT	AGT				19200	CCTACTC	CACGT	GAGT	AGC		
	30108	CCTACTC	GCGG	TGAGT	AGG				27542	CCTACTC	CACGT	GAGT	AGC		
	6097	CGTACTC	CACG	TGAGT	AGG				19772	CCTACTC	CACGT	GAGT	AGT		
	24879	CGTACTC	CACG	TGAGT	AGG				29741	CCTACTC	CACGT	GAGT	AGT		
	28094	CGTACTC	CACG	TGAGT	ATG				6097	CGTACTC	CACGT	GAGT	AGG		
Consenso		* *****		*****		****			Consenso		* *****				

Promotores ϕ M1		Localización		1		18		Localización		1		18		Promotores ϕ C31	
	23175	CGGGTTG	CCG	ACTCC	CCTT				27583	CGGGTTG	CCCG	ACTCC	CCTT		
	23544	CGGGTTG	CCG	ACTCC	CCTT				36222	CGGGTTG	CCCG	ACTCC	CCTT		
	25570	CGGGTTG	CCG	ACTCC	CCTT				25053	CGGGTTG	CCCG	ACTCC	CCTT		
	37136	CGGGTTG	CCG	ACTCC	CCTT				22666	CGGGTTG	CCCG	ACTCC	CCTT		
	13025	AGGGTTG	CCG	ACTCC	CCTT				12976	AGGGTTG	CCCG	ACTCC	CCTT		
	18634	TAGGGTTG	CCG	ACTCC	CCTT				29906	CGGGTTG	CCCG	ACTCC	CCTT		
	21930	CTAGGGTTG	CCG	ACTCC	CCTT				31189	CGGGTTG	CCCG	ACTCC	CCTT		
	28135	CAGGGTTG	CCG	ACTTC	CCTT				23031	CGGGTTG	CCCG	ACTCC	CCTT		
	30467	CAGGGTTG	CCG	ACTTC	CCTT				21759	CGGGTTG	CCCG	ACTCC	CCTT		
	41491	CAGGGTTG	CCG	ACAAC	CCTT				18487	CAGGGTTG	CCCG	ACTCC	CCTT		
	23545	CGGGTTG	CCG	ACTCC	CCTT				40566	CAGGGTTG	CCCG	ACAAC	CCTT		
Consenso		*****		**		****			Consenso		*****		**		****

Figura 16. Secuencias CIR y promotores de ϕ C31 y ϕ M1. En la parte superior se encuentran las secuencias de unión al represor, del lado izquierdo se localizan las secuencias del fago ϕ M1 y del lado derecho las secuencias del fago ϕ C31, se puede observar que estas secuencias, se encuentran conservadas, con algunas sustituciones en algunas posiciones. En la parte inferior se encuentran las secuencias promotoras de los fagos. Se observa que también conservadas las secuencias entre estos fagos.

Sitio probable de reconocimiento por el sistema Pgl.

Se ha propuesto que las secuencias invertidas repetidas conservadas (CIR) (Fig. 16), que están involucradas en la unión del represor, serían también los sitios de reconocimiento de las proteínas del sistema Pgl, debido a la presencia de estas secuencias y la sensibilidad al sistema Pgl en los fagos homoinmunes de ϕ C31 (Laity, 1993, Ighman, 1994). El hecho de que no haya sido posible obtener mutantes

del fago resistentes al sistema Pgl, ha sugerido que la secuencia blanco para dicho sistema está repetida a lo largo del genoma, o bien que es una región esencial (Sumbly, 2002).

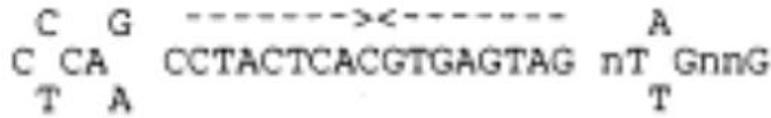
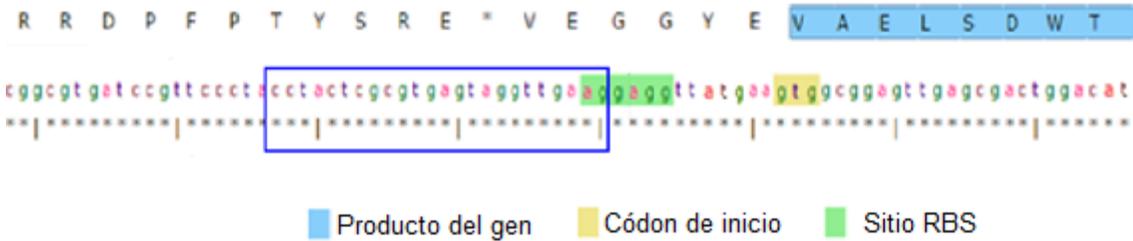


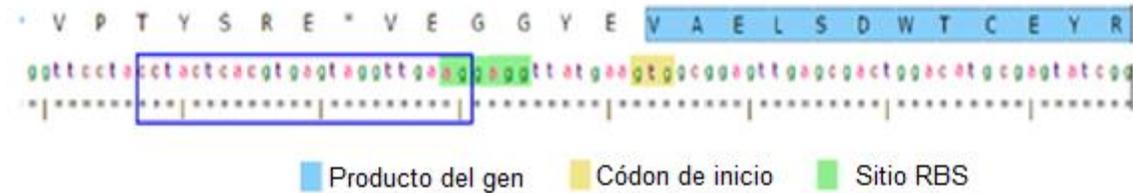
Figura 17. La secuencia consenso de las CIR encontradas en ϕ C31 (Ighman, 1994).

Realizando una búsqueda de las secuencias CIR en los genomas de ϕ M1, ϕ C31 y ϕ BT1, se localizó una cercana a una proteína gp14. (Fig.16)

A. ϕ M1 gp14



B. ϕ C31 gp44



C. BT1 gp12

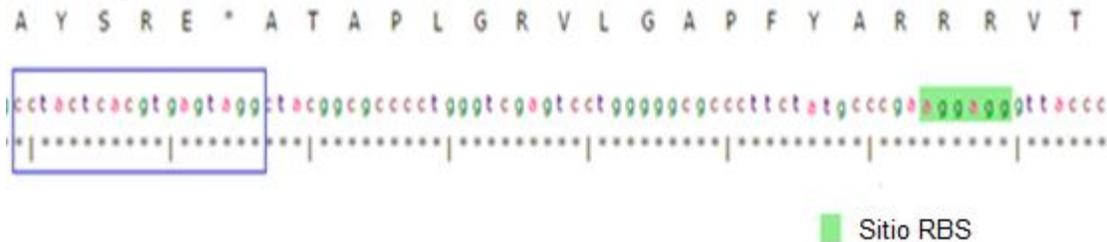


Fig. 18. Localización de las secuencias CIRs en la región cercana al sitio RBS de la proteína gp14.

El recuadro morado de la figura 18, muestra la secuencia CIR localizada en la región intergénica cercana a una proteína homóloga (gp14) entre los fagos; la identidad de residuos de aminoácidos entre la proteína gp44 y gp14 es alta. En ϕ C31 y ϕ M1, la secuencia CIR está empalmada sobre el sitio de unión a ribosoma. Esta secuencia es la única encontrada en los genomas que se empalma en el sitio

RBS. Las demás se encuentran en la región intergénica. Esta región podría estar involucrada en el reconocimiento del sistema Pgl.

Diferencias entre ϕ C31 y ϕ M1 relacionadas con el receptor glicosilado

En adición, un fago derivado de ϕ C31cD25, ϕ DT4002, puede crecer con gran eficiencia de plaqueo en la cepa DT1017 (mutante en glicosilación), mientras ϕ C31 Δ D25 tiene una deficiente formación de placas. Con la región secuenciada de los genes tardíos, el gen *gp44* de ϕ DT4002, comparado con ϕ C31 Δ D25, contiene una mutación sin sentido, implicando a *gp44* en la interacción con el receptor del hospedero (Smith, 2006.).

En un alineamiento de las secuencias de aminoácidos de las proteínas homólogas *gp44* y *gp 14* de ϕ C31 y ϕ M1, respectivamente (Fig. 17), se observan sustituciones en la secuencia de aminoácidos y como se mencionó anteriormente, una mutación sin sentido en una proteína de la cola podría modificar el rango de hospederos de los fagos derivados de ϕ C31. Por lo cual, alguna de estas sustituciones podría ser la responsable de la diferencia en la eficiencia del plaqueo de ambos fagos en la cepa afectada en la adición de residuos de oligomanosa (IB86), ya que se conoce que la o las proteína(s) receptoras están glicosiladas.

<i>gp44/PhiC31</i>	VAELSDWTCEYRGLVMGLPDSAISIVGVDGLLTPDVRSSDLTILVQRNGLWAGRDYLNGR	60
<i>gp14/phiM1</i>	VAELSDWTCEFRGLVMGLPDSAISIVQVDGLLTPDVRSSDLTILVQRHGLWAGDDYMNGR	60
	*****:***** *****:***** **:*	
<i>gp44/PhiC31</i>	TVTLTLEVYGRDRAEFTEALNALQAAFMPGVDESPFRFRFPGAASDRTAFVMARARKRSA	120
<i>gp14/phiM1</i>	TVTLTLEVYGRDRAEFTDALNALQAAFMPGVEESPFRFRFPGAASDRTAFVMARARKRSA	120
	*****:*****:*****:*****:*****	
<i>gp44/PhiC31</i>	PLDLNFAYLTCNMSVELYATSPYIVGDAARTVTVRSYKRDKVPTGLVLPVAVFPWQIEGQG	180
<i>gp14/phiM1</i>	PLDLNFAYLTCNMVVELFATSPYIIGDAARTATVRSYKRDRVPTGLVLPVAVFPWQIEGQG	180
	***** **:******:*****:*****:*****:*****	
<i>gp44/PhiC31</i>	PAPDDPVS RFTQYGSVAARPSIVITDAASPWLVDVDTGAFFAIDYDGTVVIDSAAETVTN	240
<i>gp14/phiM1</i>	PAPDDPVS RFTQYGSVAARPVVTIIDGASPWLMDVDTGAFFGIDYDGTVVIDSAAETVTD	240
	***** :.***.*****:*****.*****:*****:*****:	
<i>gp44/PhiC31</i>	AEGSDIRGLIADGSTWPEYGP GDHRLRLRSRDEYTAASASLTWSDRWV	288
<i>gp14/phiM1</i>	AQGNDIRGLITDGSTWPEYGP GDHRLRLRSRDEFTAASASLTWSDRWV	288
	:.*****:*****:*****:*****:*****	

Fig 17. Alineamiento de la secuencia de residuos de aminoácidos de las proteínas *gp44* y *gp14* de los fagos PhiC31 y PhiM1, respectivamente, la identidad es del 92%.

Relación filogenética de ϕ M1 con otros actinofagos.

Con el fin de tener una visión actualizada entre los diferentes actinofagos descritos a la fecha, se alinearon los genomas completos de dichos fagos de *Streptomyces* en el servidor MAFFT (Kato, 2013) y posteriormente se construyó un árbol filogenético (figura 18). Los grupos identificados como clúster se basan en la organización de su genoma propuesta por Smith, 2013. El fago ϕ M1 se encuentra en el clúster de ϕ C31, ϕ BT1 y TG1.

El clúster BG, se caracteriza por estar formado por fagos que solo infectan a *Streptomyces griseus* y cuyo ciclo de vida es únicamente lítico.

El cluster-R4 presenta la mayor cantidad de miembros (10), en él se localizan fagos de aislados de *S. coelicolor* (ϕ Hau3, R4 por ejemplo), *S. venezuelae* (ϕ CAM, Verse), estos fagos presenta homología con un clúster de micofagos.

Curiosamente los fagos R4 y ELB20 son casi idénticos en secuencia, aunque fueron aislados aproximadamente con 35 años de diferencia de localidades distintas (Smith, 2013), lo cual también ocurrió con el aislamiento del fago ϕ M1 y ϕ C31 el cual fue descrito por primera vez en 1972 (Lomovskaya, 1972).

Los fagos μ 1/6, VWB, SV1 y ϕ SASD1 no muestran una similitud global de secuencias respecto a otros fagos para formar un clúster.

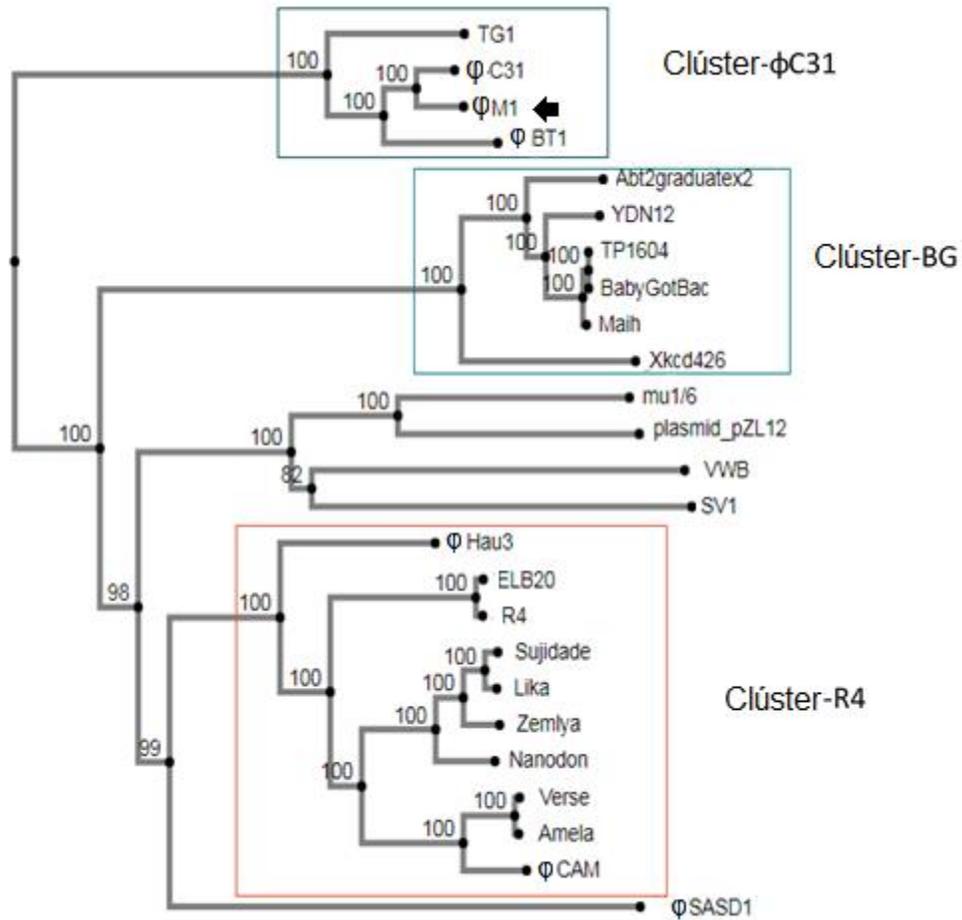


Figura 18. Análisis filogenético de los fagos secuenciados de *Streptomyces* (25) basado en su genoma completo. El algoritmo utilizado fue Neighbor-Joining (NJ). Los valores de “bootstrap” indican el número de veces que un nodo fue soportado en 1000 réplicas de las secuencias. Se identifican 3 clúster, en verde: clúster- φC31, se señala con una flecha el fago φM1. En azul: el clúster-BG y en rojo: clúster-R4.

Conclusiones

- De 70 fagos aislados, sólo uno, denominado ϕ M1, fue incapaz de infectar a una mutante deficiente de glicosilación, indicando que este fago requiere proteínas glicosiladas como receptor. Este fago fue también el único incapaz de infectar a la cepa M145, que es Pgl⁺, pero sí fue capaz de formar placas turbias en la mutante Pgl⁻ J1928, indicando que es sensible al sistema Pgl. Lo anterior demuestra que ninguno de los dos fenotipos es frecuente en la naturaleza.
- El aislamiento de ϕ M1 muestra alguna asociación entre el fenotipo Pgl y el requerimiento de glicosilación de proteínas para infectar.
- Los genomas de los fagos ϕ M1 y ϕ C31 presentan una identidad global a nivel de nucleótidos del 90%, y pertenecen al mismo "clúster" filogenético junto con los fagos ϕ BT1 y TG1.
- El genoma del fago ϕ M1, consiste de 41,538 pb, y contiene 53 ORFs, de los cuales a 22 se les asignó una función putativa.

Perspectivas

- Aislamiento de más actinofagos con el objetivo de hacer genómica comparativa de fagos, con la finalidad de enriquecer las bases de datos para poder identificar las regiones de DNA involucradas en el reconocimiento por el sistema de restricción Pgl de *S. coelicolor*, así como determinar qué secuencias están involucradas en el requerimiento de receptores glicosilados.

Referencias

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." *J. Mol. Biol.* 215:403-410.
- Arroyo, E. (2017). Identificación de una manosiltransferasa involucrada en la glicosilación de proteínas en *Streptomyces coelicolor*. (tesis de Licenciatura). UNAM, México.
- Belcaid, M., Bergeron, A., & Poisson, G. (2011). The evolution of the tape measure protein: units, duplications and losses. *BMC Bioinformatics*. 12(9).
- Bentley, S., et al. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*. 41,141-147
- Bhattacharya, B. et al. (2008). Cloning, characterization and expression analysis of nucleotide metabolism-related genes of mycobacteriophage L5. *FEMS Microbiol Lett.* 280(1):64-72.
- Bläsi, U. (1996). Two beginnings for a single purpose: the dual-start holins in the regulation of phage lysis. *Mol Microbiol.* 21(4):675-82.
- Bruton, C. et al. (1991). Phage Vectors that Allow Monitoring of Transcription of Secondary Metabolism Genes in *Streptomyces*. *Bio/Technology*. (9), 652–656
- Bukovska, G. et al. (2006). Complete nucleotide sequence and genome analysis of bacteriophage BFK20 — A lytic phage of the industrial producer *Brevibacterium flavum*. *Virology*. 348. 57-71.
- Chandra, G., & Chater, K. F. (2014). Developmental biology of *Streptomyces* from the perspective of 100 actinobacterial genome sequences. *Fems Microbiol Rev*, 38(3), 345–379.
- Chater K. (1986). *Streptomyces* phages and their application to *Streptomyces* genetics En Queener N & Day, L (Eds.), The Bacteria. Press, Orlando. 119-158.
- Combes, P., Till, R., Bee, S., & Smith, M. C. M. (2002). The *Streptomyces* Genome Contains Multiple Pseudo-*attB* Sites for the ϕ C31-Encoded Site-Specific Recombination System. *J. Bacteriol.*, 184(20). 5746–5752.
- Córdova-Dávalos LE. et al. (2014) Lipoprotein N-acyl transferase (Lnt1) is dispensable for protein O-mannosylation by *Streptomyces coelicolor*. *FEMS Microbiol Lett.* 350(1):72-82
- Cowlshaw D & Smith, M. (2001). Glycosylation of a *Streptomyces coelicolor* A3(2) cell envelope protein is required for infection by bacteriophage ϕ C31. *Mol Microbiol.* 41(3):601-10.
- Cox, K. L., & Baltz, R. H. (1984). Restriction of bacteriophage plaque formation in *Streptomyces* spp. *J. Bacteriol.* 159(2), 499–504.
- F. Desiere, et al. (1998). Evolution of *Streptococcus thermophilus* bacteriophage genomes by modular exchanges followed by point mutations and small deletions and insertions. *Virology*. 241;345-356
- Donegan-Quick, R. et al. (2017). Genome Sequences of Five *Streptomyces* Bacteriophages Forming Cluster BG. *Genome Announc.*
- Erill, I., & Caruso, S. M. (2016). Complete Genome Sequence of the *Streptomyces* Phage Nanodon. *Genome Announc*, 4(5).

- Espitia, C. et al (2010). New insights into protein O-mannosylation in actinomycetes. *Mol Biosyst.* 6(5):775-81.
- Espitia, C., and R. Mancilla. 1989. Identification, isolation and partial characterization of *Mycobacterium tuberculosis* glycoprotein antigens. *Clin. Exp. Immunol.* 77:378-383.
- Farkasovska, J Godany, A. Vlcek, C. (2003) Identification and characterization of an endolysin encoded by the *Streptomyces aureofaciens* phage mu1/6. *Folia Microbiol.*48 (6), 737-744
- Fokine, A., & Rossmann, M. G. (2016). Common Evolutionary Origin of Procapsid Proteases, Phage Tail Tubes, and Tubes of Bacterial Type VI Secretion Systems. *Structure* . 24(11), 1928–1935.
- González-Cerón Gabriela, Osvaldo Josué Miranda-Olivares, Luis Servín-González; Characterization of the methyl-specific restriction system of *Streptomyces coelicolor* A3(2) and of the role played by laterally acquired nucleases. *FEMS Microbiol Lett.*, 301 (1), 35–43.
- Gregory, M. A., Till, R., & Smith, M. C. M. (2003). Integration Site for *Streptomyces* Phage ϕ BT1 and Development of Site-Specific Integrating Vectors. *J. Bacteriol*, 185(17), 5320–5323.
- Gregory, M. A., Till, R., & Smith, M. C. M. (2003). Integration Site for *Streptomyces* Phage ϕ BT1 and Development of Site-Specific Integrating Vectors. *Journal of Bacteriology*, 185(17), 5320–5323.
- Harger J. Meskauskas, A & Dinman JD. (2002). An "integrated model" of programmed ribosomal frameshifting. *Trends Biochem Sci.* 27(9):448-54.
- Hopwood, D. (2006). Soil to genomics: the *Streptomyces* chromosome. *Annu Rev Genet.* 40, 1-23.
- Hoskisson, P. A., Sumbly, P., & Smith, M. C. M. (2015). The phage growth limitation system in *Streptomyces coelicolor* A(3)2 is a toxin/antitoxin system, comprising enzymes with DNA methyltransferase, protein kinase and ATPase activity. *Virology*, 477, 100–109.
- Huang, X., & Madan, A. (1999). CAP3: A DNA Sequence Assembly Program. *Genome Res*, 9(9), 868–877.
- Ikeda, H, Shin-ya, k & Omura Satoshi. (2014). Genome mining of the *Streptomyces avermitilis* genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. *J Ind Microbiol Biotechnol.* 41, 233-250
- Ingham CJ. et al. (1993). Multiple novel promoters from the early region in the *Streptomyces* temperate phage ϕ C31 are activated during lytic development. *Mol Microbiol.* 9(6):1267-74.
- Jayapal, K. et al. (2007). Comparative genomic hybridizations reveal absence of large *Streptomyces coelicolor* genomic islands in *Streptomyces lividans*. *BMC Genomics*, 8, 229.
- Kala, S., Cumby, N., Sadowski, P. D., Hyder, B. Z., Kanelis, V., Davidson, A. R., & Maxwell, K. L. (2014). HNH proteins are a widespread component of phage DNA packaging machines. *PNAS*, 111(16), 6022–6027.
- Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol*, 30(4), 772–780.
- Katsura, I, Hendrix, R. (1984) Length determination in bacteriophage lambda tails. *Mol Microbiol.* 41(3):601-10.

- Krumsiek J, Arnold R, & Rattei T. (2007). Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics*. 15;23(8):1026-8.
- Kuhstoss, S & Rao, RN. (1991). Analysis of the integration function of the streptomycete bacteriophage ϕ C31. *J Mol Biol*. 20;222(4):897-908.
- Kwak, J., McCue, L. A., Trczianka, K., & Kendrick, K. E. (2001). Identification and Characterization of a Developmentally Regulated Protein, EshA, Required for Sporogenic Hyphal Branches in *Streptomyces griseus*. *J. Bacteriol*, 183(10):3004–3015.
- Laity C., Chater K.F., Lewis C.G., Buttner M.J. (1993). Genetic analysis of the ϕ c31-specific phage growth limitation (Pgl) system of *Streptomyces coelicolor* A3(2) *Mol. Microbiol.* 7:329–336.
- Langmead, B. (2010). Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics* .11, 11.7
- Laslett, D., & Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.*, 32(1), 11–16.
- Layton, S. et al. (2016). Genome Sequences of *Streptomyces* Phages Amela and Verse. *Genome Announc.*
- Loenen, W. A., and Raleigh, E. A. (2014) The other face of restriction: modification-dependent enzymes, *Nucleic Acids Res.*, 42, 56–69.
- Lommel, M., & Strahl, S. (2009) Protein O-mannosylation: Conserved from bacteria to humans, *Glycobiology*, 9(5), 816–828.
- Lomovskaya, N. Mkrtumian, N, Gostimskaya, L. (1970). Isolation and characteristics of *Streptomyces coelicolor* actinophage. *Genetika* 6(12): 135-136.
- Lomovskaya, N. D., Chater, K. F., & Mkrtumian, N. M. (1980). Genetics and molecular biology of *Streptomyces* bacteriophages. *Microbiol Rev*, 44(2), 206–229.
- Lomovskaya, N. et al. (1972). Characterization of temperate actinophage ϕ C31 isolated from *Streptomyces coelicolor* A3(2). *J. Virol.*, 9, 258-262.
- Morita, K. et al. (2009), The site-specific recombination system of actinophage TG1. *FEMS Microbiol Lett*, 297: 234-240.
- Myronovskiy, M., et al (2014). Genome rearrangements of *Streptomyces albus* J1074 lead to the carotenoid gene cluster activation. *Appl. Microbiol. Biotechnol.* 98, 795–806.
- Ong, E. et al. (1994). *Streptomyces lividans* Glycosylates the Linker Region of a B-3-1,4-Glycanase from *Cellulomonas fimi*. *J. Bacteriol*. 176: 999-1008
- Pablo Cruz-Morales et. al. (2013). The Genome Sequence of *Streptomyces lividans* 66 Reveals a Novel tRNA-Dependent Peptide Biosynthetic System within a Metal-Related Genomic Island, *Genome Biol Evol.*, 5, 1165–1175.
- Pingoud, A., Wilson, G. G., & Wende, W. (2014) Type II restriction endonucleases— a historical perspective and more, *Nucleic Acids Res.*, 42, 7489–7527.
- Roberts, R.J. et al. (2003) A nomenclature for restriction enzymes, DNAmethyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.* 31: 1805–1812.
- Rutjerford, K. et al. (2000). Artemis: sequence visualization and annotation. *Bioinformatics*. 16(10):944-5
- Schmidt MA1, Riley LW, Benz I. (2003). Sweet new world: glycoproteins in bacterial pathogens. *Trends Microbiol*. 11(12):554-61.

- Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J. M., & Birol, İ. (2009). ABySS: A parallel assembler for short read sequence data. *Genome Res*, 19(6), 1117–1123.
- Smith, G. T., Sweredoski, M. J., & Hess, S. (2014). O-linked glycosylation sites profiling in *Mycobacterium tuberculosis* culture filtrate proteins. *J Proteomics*, 97, 296–306.
- Smith, M, et al (2013). Evolutionary Relationships among Actinophages and a Putative Adaptation for Growth in *Streptomyces* spp. *J. Bacteriol.*, 195(21), 4924–4935.
- Smith, M. C. et al. (2013). Evolutionary Relationships among Actinophages and a Putative Adaptation for Growth in *Streptomyces* spp. *J. Bacteriol* , 195(21), 4924–4935.
- Smith, M. C., Burns, R. N., Wilson, S. E., & Gregory, M. A. (1999). The complete genome sequence of the *Streptomyces* temperate phage straight ϕ C31: evolutionary relationships to other viruses. *Nucleic Acids Res.* 27(10), 2145–2155.
- Smith, M. (2006). Molecular Genetics of *Streptomyces* Phages. en Calendar, R (Ed). The Bacteriophages. Oxford University Press.
- Suárez, J. E., Caso, J., Rodriguez, A. and Hardisson, C. (1984), Structural characteristics of the *Streptomyces* bacteriophage ϕ C31. *FEMS Microbiol Lett.* 22: 113-117.
- Sumbly P., Smith M.C.M. (2002) Genetics of the phage growth limitation (Pgl) system of *Streptomyces coelicolor* A3(2). *Mol. Microbiol.* 44:489–500.
- Sumbly P., Smith M.C.M. (2003). Phase variation in the phage growth limitation system of *Streptomyces coelicolor* A3(2) *J. Bacteriol.* 185:4558–4563.
- Tock, M.R., and Dryden, D.T.F. (2005) The biology of restriction and anti-restriction. *Curr Opin Microbiol.*8: 466–472.
- VanderVen BC et al. (2005) Export-mediated assembly of mycobacterial glycoproteins parallels eukaryotic pathways. *Science.* 309: 941–943.
- Wang, S. et al. (2010). Complete genomic sequence analysis of the temperate bacteriophage ϕ SASD1 of *Streptomyces avermitilis*. *Virology.* 403, 78-94.
- Wehmeier S, et al. (2009). Glycosylation of the phosphate binding protein, PstS, in *Streptomyces coelicolor* by a pathway that resembles protein O-mannosylation in eukaryotes. *Mol Microbiol.*71:421–433
- Wilson S. et al. (2009). Control of lytic development in the *Streptomyces* temperate phage ϕ C31. *Mol Microbiol.* 16(1):131-43.
- Xu, J. Hendrix, R & Duda, R. (2004). Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. *Mol. Cell*, 16, 11-21.
- Yagüe, P., Lopez-Garcia, M. T., Rioseras, B., Sanchez, J., & Manteca, A. (2012). New insights on the development of *Streptomyces* and their relationships with secondary metabolite production. *Curr Trends Microbiol.*, 8, 65–73.
- Zaburanyi, N., Rabyk, M., Ostash, B., Fedorenko, V., & Luzhetskyy, A. (2014). Insights into naturally minimised *Streptomyces albus* J1074 genome. *BMC Genomics*, 15, 97.
- Zerbino, R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18(5), 821–829.
- Zhang, H., Lee, S.-J., Kulczyk, A., Zhu, B., & Richardson, C. (2012). Heterohexamer of 56- and 63-kDa Gene 4 Helicase-Primase of Bacteriophage T7 in DNA Replication. *J Biol Chem.* , 287(41), 34273–34287.

Protocolos

Protocolo 1: Aislamiento de fagos nuevos de *Streptomyces*

1. En un matraz con 20 mL de DNB enriquecido [(glucosa 0.5%, $\text{Ca}(\text{NO}_3)_2$], mezclar 10g de suelo y 100 μL de una preparación densa de esporas.
2. Incubar toda la noche a 30°C con agitación.
3. Pasar el contenido del matraz a tubos de centrifuga de 35 mL.
4. Centrifugar a 10,000 rpm durante 10 min.
5. Transferir el sobrenadante a otro tubo de 35 mL y filtrarlo a través de un filtro estéril de 0.45 μm , usando una jeringa de 10 mL.
6. Hacer diluciones seriales del filtrado de 10^{-1} hasta 10^{-7} en buffer SM.
7. Plaquear 0.1 mL de cada dilución con 20 μL de esporas en 2.5 mL de medio SNA sobre cajas de medio DNA con MgSO_4 10 mM y $\text{Ca}(\text{NO}_3)_2$ 25 mM.
8. Incubar toda la noche.

Protocolo 2: Aislamiento de las placas de lisis

1. Succionar con una pipeta Pasteur estéril una placa aislada.
2. Transferirla a 1 mL de buffer SM y dejar difundiendo los fagos durante 2 horas a temperatura ambiente.
3. Realizar diluciones seriales y proceder al ensayo de formación de placas.

Protocolo 3: Obtención de stocks con un título alto

1. Hacer diluciones seriales de una suspensión de fagos hasta 10^{-7} en buffer SM.
2. A 2.5 mL de SNA añadir 100 μL de las diluciones de fagos y 20 μL de esporas concentradas.
3. Verter esta mezcla en cajas Petri con medio DNA enriquecido y remover hasta que cubra uniformemente la superficie.
4. Esperar a que seque completamente el medio suave.
5. Incubar toda la noche a 30°C.
6. Añadir 2.5 mL de buffer SM a las cajas que contengan placas casi confluentes.
7. Dejar reposar 2 horas a temperatura ambiente.
8. Extraer con pipeta el buffer de las cajas y filtrarlo a través de un filtro estéril de 0.45 μm , usando una jeringa de 10 mL.
9. Guardar a 4°C.