



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**MAESTRÍA EN CIENCIAS BIOQUÍMICAS
CENTRO DE CIENCIAS GENÓMICAS
INSTITUTO DE BIOTECNOLOGÍA**

Análisis de la lisogenia en *Rhizobium*

TESIS

Que para optar por el grado de:
Maestro en Ciencias Bioquímicas

PRESENTA

Arun Bohindra Reverte Vera

DIRECTOR DE TESIS

Dr. Víctor Manuel González Zúñiga
Centro de Ciencias Genómicas

COMITÉ TUTOR

Dra. Susana Brom Klanner
Centro de Ciencias Genómicas
Dr. Víctor Humberto Bustamante Santillan
Instituto de Biotecnología

Centro de Ciencias Genómicas, Cuernavaca, Mor. Mayo, 2019.



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.

INDICE

RESUMEN	4
INTRODUCCIÓN	5
GENERALIDADES.....	5
Historia	
Estructura viral	
Clasificación	
Ciclos de vida	
Ciclo lítico	
Ciclo lisogénico	
BACTERIÓFAGOS (FAGOS)	12
Importancia de los bacteriófagos en los ecosistemas	
Importancia de los bacteriófagos en la evolución bacteriana	
Taxonomía y diversidad de fagos	
Estructura genómica de fagos	
Resistencia bacteriana a fagos	
PROFAGOS	17
Clasificación de profagos	
Inducción de profagos	
Conversión lisogénica	
BACTERIÓFAGOS EN <i>Rhizobium</i>	20
JUSTIFICACIÓN	23
OBJETIVOS	24
GENERAL.....	24
PARTICULARES	24
METODOLOGÍA	25
CULTIVOS	25
Cultivo líquido	
Cultivo sólido	
Agar suave	
AISLAMIENTOS	25
Aislamiento de DNA bacteriano	
Aislamiento de partículas virales	
Aislamiento de DNA viral	
Concentraciones de ADN	
Electroforesis para ADN	
FAGOS	29
Inducción	
Inducción con mitomicina C	
Inducción con pH y temperatura	
Método de placa lítica para la detección de fagos	
Cuento de unidades formadoras de placa (UFP) para fagos obtenidos por inducción con mitomicina C	
PROFAGOS	32
Predicción de profagos con PHAST	
Clasificación de profagos por identidades nucleotídicas y diseño de marcadores moleculares	
Condiciones de PCR para el marcador de primer grupo	

RESULTADOS Y DISCUSIÓN	35
INDUCCIÓN DEL PROFAGO RHEph10 A PARTIR DE LA CEPA <i>Rhizobium etli</i> CE3 LISOGENIZADA CON EL FAGO RHEph10.....	35
INDUCCIÓN DE PROFAGOS EN CEPAS DE <i>Rhizobium</i>	43
PREDICCIÓN DE PROFAGOS EN GENOMAS DE <i>Rhizobium</i>	47
AGRUPAMIENTO, POR SIMILITUD NUCLEOTIDICA, DE PROFAGOS PREDICHOS POR PHAST	56
IDENTIFICACIÓN DE SECUENCIAS DE PROFAGOS EN CEPAS DE <i>Rhizobium</i> NO SECUENCIADAS .	59
CONCLUSIÓN	64
PERSPECTIVAS	65
BIBLIOGRAFÍA	66
ANEXO	72

RESUMEN

En este trabajo se revisan características lisogénicas de aproximadamente 120 cepas bacterianas pertenecientes al género *Rhizobium*. Los análisis se enfocan en tres vertientes principales: 1) Inducción lisogénica mediante varios inductores, principalmente mitomicina C, disminución de pH y aumento de temperatura; 2) Búsqueda bioinformática de profagos en aproximadamente 50 genomas de *Rhizobium* y; 3) Generación (y utilización) de marcadores moleculares de ADN basados en las secuencias de prófagos encontrados bioinformáticamente, cuya utilidad radicó en la búsqueda de secuencias profágicas en cepas de *Rhizobium* no secuenciadas. Los resultados son negativos en cuanto a la inducción de cepas lisogénicas, pero positivos con respecto a la búsqueda bioinformática; en la cual se ubicaron 72 secuencias profágicas de 50 genomas analizados. Con respecto al tercer punto, se generaron 7 marcadores de ADN, de los cuales, uno fue probado en cerca de 70 cepas no secuenciadas de *Rhizobium*, dando positivo en muchas de ellas. Se concluye que la ausencia de inducción probablemente se debió a la función nula o incorrecta de genes indispensables para el ciclo lítico. Por otro lado, la gran cantidad de profagos encontrados fue suficiente para agruparlos coherentemente y generar marcadores con posibles aplicaciones futuras.

INTRODUCCIÓN

GENERALIDADES

Los virus son partículas que constan de material genético (RNA o DNA de cadena sencilla o doble) envuelto en una (o dos) cubierta(s) proteica(s). Algunos pueden tener una cubierta lipídica adicional, la cual tiene embebidas proteínas, glicolípidos y/o glicoproteínas que participan en el reconocimiento virus-hospedero. Estas partículas constituyen entidades incapaces de reproducirse por sí mismas; necesitan de la maquinaria celular para replicar su propio material genético, sintetizar sus proteínas y construirse (Flint, 2009).

A nivel popular la gente identifica a los virus como microorganismos o partículas infecciosas que pueden causar enfermedades e incluso la muerte. La mayoría sabe de virus que infectan a humanos y animales, pero pocos saben de virus que infectan plantas, y aún menos, sobre los que infectan a bacterias. La realidad es que existe una variedad inmensa de tipos de virus. Cada ser vivo es “infectado” cotidianamente por muchos virus, sin embargo, muy pocos causan enfermedades. La mayoría de ellos interactúa con sus hospederos mediante diversos mecanismos. De hecho, algunos autores han estimado que mucho menos del 3% de todos los virus son causantes de enfermedades (Dimmock *et al.*, 2007).

Historia. Los virus se encontraron poco después de que se descubrieran las bacterias patógenas. Estas podían verse a través del microscopio, cultivarse *in vitro* en medios nutritivos, ser retenidas en un filtro y causar enfermedades en huéspedes sanos. Sin embargo, algunos agentes infecciosos solo tenían la característica de causar enfermedades cuando se inoculaban en organismos sanos; no podían verse al microscopio, no podían cultivarse *in vitro* y no eran retenidas por los filtros que retenían bacterias (Cann, 2005).

A estos agentes “extraños” los describió Iwanowski en 1892, quien encontró que la causa de la enfermedad del mosaico del tabaco era un agente que solo cubría el requisito antes mencionado (causar enfermedades en huéspedes sanos); sin embargo, Iwanowski le dio poca importancia a su hallazgo. En 1898 Beijerinck repitió los experimentos de Iwanowski y se convenció de que la sustancia que causaba la enfermedad no era una bacteria, sino un “fluido vivo contagioso” (veneno), término del cual proviene la palabra “virus” (Cann, 2005).

Los hallazgos de Iwanowski y Beijerinck fueron poco aceptados debido a que la causa de la enfermedad (el contenido del “fluido vivo contagioso”) podría no ser un “fluido vivo”, sino toxinas secretadas por bacterias u otro tipo de células. Sin embargo, en el mismo año de la publicación de Beijerinck (1898), Friedrich Loeffler y Paul Frosch aislaron el “fluido vivo contagioso” que causaba la enfermedad de piel y boca, y con él realizaron el siguiente experimento: diluyeron el fluido en varios órdenes de magnitud, (10^0 - 10^{-8} , por citar un ejemplo) e inocularon las diluciones en organismos sanos, los cuales enfermaron

al poco tiempo. Este hecho sugirió fuertemente que la infección no se debía a toxinas ya que en una dilución tan alta (10^{-8}) habría muy pocas toxinas (si es que hubiera), siendo casi nula la posibilidad de producir una enfermedad. Fue así como se llegó a la conclusión de que el “fluido vivo contagioso” (virus de ahora en adelante) contenía partículas capaces de replicarse (Dimmock *et al.*, 2007).

Los hallazgos no los aceptó la comunidad científica hasta que se realizaron más investigaciones, lo que llevó al descubrimiento de varias enfermedades causadas por virus: 1908 (Ellerman y Bang) leucemia en gallinas; 1909 (Landsteiner y Popper) poliomielitis en humanos (parteaguas de la aceptación de la teoría); 1911 (Rous) generación de tumores en pollos; etcétera (Citados en Cann, 2005).

Los virus bacterianos los descubrió accidentalmente Frederick Twort en 1915. Él estaba probando si podía generar vacunas para la viruela, utilizando como medio de cultivo camas de células eucariontes, que comúnmente se contaminaban con bacterias. En diferentes zonas del cultivo se podían observar crecimientos bacterianos, dentro los cuales se formaban placas líticas (ausencia de crecimiento bacteriano), que Twort atribuyó a la lisis por parte de virus bacterianos (Flint *et al.*, 2009).

Felix d'Herelle, quien realizó investigaciones más claras y precisas- Él observó que los cultivos bacterianos de disentería desaparecían cuando se les agregaba un filtrado bacteriano obtenido a partir de bacterias de aguas negras, al cual bautizó como bacteriófago (comedor de bacterias). D'Herelle publicó sus hallazgos como “un microbio invisible, antagonista del bacilo disentería” (D'Herelle, 1917).

Estructura viral. La estructura viral consta de un ácido nucleico (RNA o DNA) de cadena sencilla o doble, el cual puede ser lineal o circular, así como entero o fragmentado. Este material genético genera proteínas necesarias para la sobrevivencia del virus, y también tienen zonas específicas donde actúan diversas proteínas (tanto virales como celulares). Para proteger este material genético se forma una cubierta proteica que, asociado al primero, forma la nucleocápside (o cápside). Algunos virus tienen una cubierta de proteínas adicional (también denominada cápside) y otros tienen una membrana lipídica que recubre la cápside (virus envueltos). A la bicapa lipídica (y en algunas ocasiones a la cápside), se adhieren glicoproteínas y/o glicolípidos relacionados con el reconocimiento virus-célula (Flint *et al.*, 2009).

La variación de los distintos tipos de ácidos nucleídos y de las proteínas de la cápside, así como sus diferentes acomodos, hace posible una amplia gama de morfologías virales, que se agrupan en tres tipos principales: helicoidales, icosaédricas y complejas. La simetría helicoidal corresponde a virus filamentosos y pleomórficos, la icosaédrica a formas esféricas, y la compleja se refiere a la mezcla de las dos anteriores. Aunque los virus se observen circulares, irregulares o pleomórficos, la estructura interna (cápside) tiene alguna de las tres estructuras principales (Lodish *et al.*, 2000).

La unidad estructural de la cápside se denomina capsómero, sin embargo cada tipo de partícula viral posee una cantidad diferente de capsómeros. Por ejemplo, cada cápside

del adenovirus se compone de 252 capsómeros. Cada capsómero se compone de pocos polipéptidos (1-3 generalmente) del mismo o de diferentes tipos. Cabe mencionar que estas unidades estructurales se asocian estrechamente al ácido nucleico, formando así una estructura estable (la cápside o nucleocápside), que en algunos casos puede durar miles de años sin descomponerse.

Como se menciona en los párrafos anteriores, existen tres tipos de partículas virales (helicoidales, icosaedricas y complejas). Dentro de cada estructura existe una gran variedad de tamaños, generalmente de 10-250 nm de diámetro. Por ejemplo, la nucleocápside helicoidal del virus *sendai*, perteneciente a la familia paramixoviridae, tienen una longitud de 1000 nm y un grosor de 17 nm. (Gelderblom, 1996). En la figura I1 se muestra una variedad de partículas virales, se observan los tamaños, tipos de estructura, presencia o ausencia de envoltura (bicapa lipídica) y tipo de material genético.

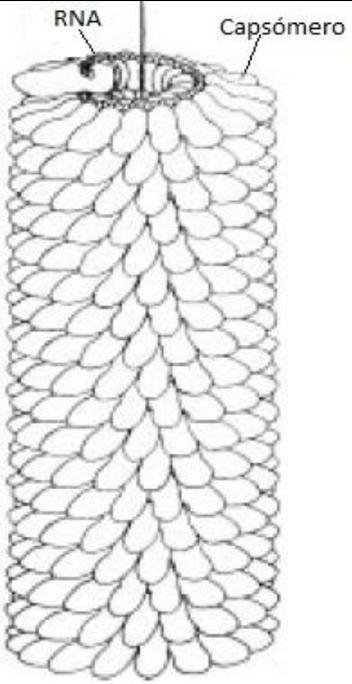
	Tamaño (nm)	Tipo de estructura	Envoltura	Material genético	Referencia
 <p><i>Virus de mosaico del tabaco</i></p>	16x300	Helicoidal	No	RNA (Monocatenario positivo)	Franklin, 1956.

Figura I1. Estructura y algunas características de virus representativos.

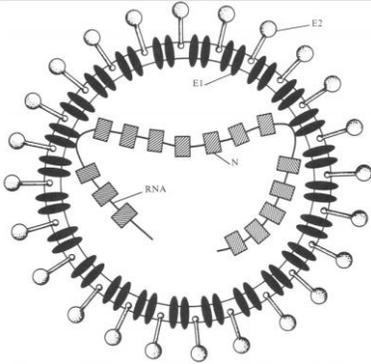
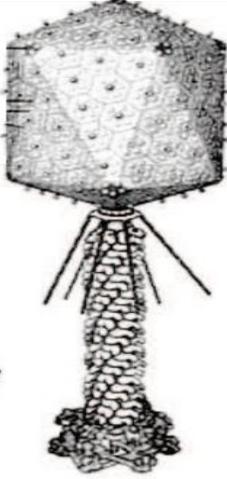
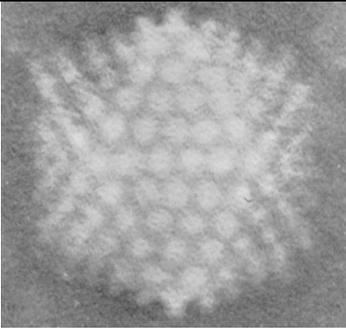
 <p><i>Coronavirus (p. ej. Hepatitis)</i></p>	60-220 de diámetro	Icosaédrica	Si	RNA (Monocatenario positivo)	Siddelle <i>et al.</i> , 1983.
 <p><i>Fago T4</i></p>	200x86	Compleja	No	DNA (Doble cadena)	Mesyanzhinov <i>et al.</i> , 2004.
 <p><i>Adenovirus A</i></p>	8 de diámetro	Icosaédrica	No	DNA (Doble cadena)	Velentine y Pereira, 1965.

Figura 11. Continuación.

Clasificación. Dependiendo del área o enfoque de la investigación, existen diversas formas de clasificación de los virus: por estructura, rango de hospedero, tipo de material genético, secuencia genómica, etc. Por ejemplo, en la comunidad médica se acostumbra clasificar a los virus por las enfermedades que producen o las zonas del cuerpo que afectan. Sin embargo, si se busca una comprensión global de todos los virus conocidos, una clasificación basada en las principales características biológicas y morfológicas sería más conveniente, tal como lo hizo Baltimore, que tomo como principal eje el tipo de material genético y como secundario la morfología y demás características (rango de hospederos, etc.) (Flint, 2009).

Con base en la clasificación de Baltimore, todos los virus conocidos se agrupan en órdenes, familias, subfamilias (en algunos casos), géneros y especies, escritas con nomenclatura binomial latina. (Cann, 2005). Este tipo de clasificación inició en 1962 con la reunión de tres investigadores (Lwoff, Horne y Tournier), dando nacimiento al comité internacional sobre la taxonomía de virus (ICTV) (Flint, 2009).

El ICTV publicó su primer reporte en 1971, dando a conocer las familias *Picornaviridae* (con 3 géneros) y *Papovaviridae* (con 2 géneros), además de 38 géneros no emparentados, siendo 286 el número total de especies. El noveno reporte fue publicado en el año 2011 y dio a conocer 6 órdenes, 87 familias (19 subfamilias contadas dentro de algunas familias), 349 géneros, y 2285 especies. Y aunque no es un reporte oficial, la última actualización del ICTV (en 2015) dio a conocer 7 órdenes, 111 familias, 27 subfamilias, 609 géneros y 3704 especies (ICTV oficial web site).

En años recientes han surgido clasificaciones basadas en genomas y proteomas. Por ejemplo; la comparación de regiones génicas específicas entre distintas cepas de una especie (Simmonds *et al.*, 1993); análisis filogenéticos de una especie basados en la totalidad de su genoma (Lu *et al.*, 2006); comparaciones genómicas entre todos los virus con genomas secuenciados hasta el 2008 (Hatfull, 2008); y una cantidad enorme de trabajos con diferentes enfoques clasificatorios.

Como se puede notar en los párrafos anteriores, hay muchas formas de clasificar a los virus, elegir la adecuada depende de los objetivos del investigador. Debido a que este trabajo tiene un enfoque genómico, convendría considerar las 2 últimas clasificaciones, en especial la de Hatfull, 2008.

Ciclos de vida. Los ciclos de vida están bien estudiados en algunos bacteriófagos, por ejemplo en el fago λ y el T4. Dichos ciclos se separan en lítico y lisogénico: el ciclo lítico es la fase en la que se da una producción masiva de partículas virales; el ciclo lisogénico es la fase de latencia en la que el material genético de origen viral se encuentra en un estado “durmiente” dentro de la célula hospedera, replicándose independientemente (episoma) o junto con el ADN del hospedero (provirus), figura I2. Además, existen algunos subtipos o procesos alternos, por ejemplo, la fase crónica está caracterizada por la producción y liberación constante y lenta de partículas virales, de tal forma que el hospedero nunca se lisa (Abedon, 2008).

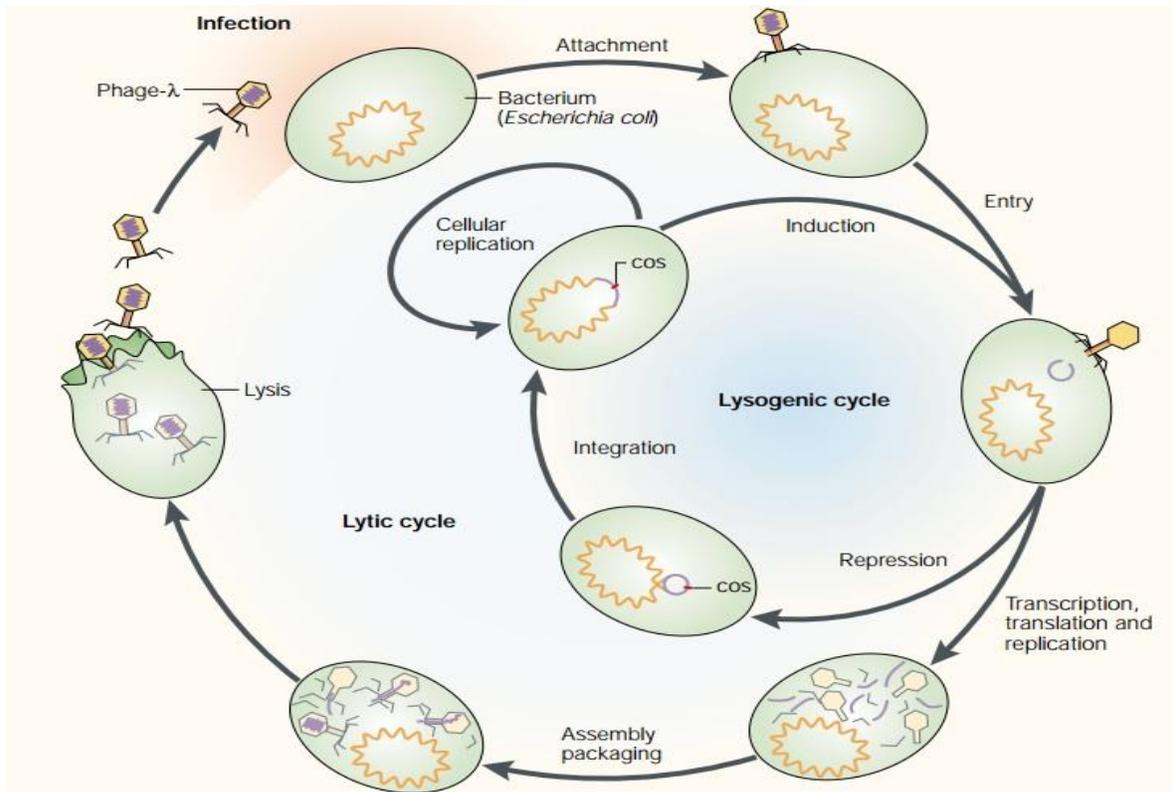


Figura 12. Representación general de los ciclos lítico y lisogénico (Campbell, 2003).

Ciclo lítico. El ciclo lítico, también conocido como productivo o infeccioso, se caracteriza por la generación masiva de partículas virales, y que finaliza con la muerte de la célula hospedera debido a que la salida de una cantidad considerable de partículas virales (generalmente de 100-1000) necesita de la lisis celular para liberarse. El ciclo se divide en periodo temprano (antes de la replicación del material genético viral) y periodo tardío (después de la replicación del material genético viral). El punto es lograr la producción de partículas virales completas e infecciosas (viriones) a partir de la expresión secuencial de genes presentes en el genoma viral (Lewin, 2008).

Cada familia, género, e incluso especie y subespecie de virus tiene ciclos líticos con diferentes grados de discrepancia. Los mecanismos de entrada del virión son variados, al igual que los factores necesarios para la expresión de sus genes (Flint *et al.*, 2009). Como ejemplo del ciclo lítico, se resumirá el ciclo de vida del fago lambda.

En el periodo temprano se transcriben genes de enzimas implicadas en la síntesis de material genético viral, las cuales trabajan junto a enzimas celulares (transcriptasas y algunos factores de transcripción). El trabajo en conjunto finaliza con un banco de genomas virales que además de servir como moldes para sucesivas replications y transcripciones, serán posteriormente encapsidadas (Lewin, 2008).

En el periodo tardío, las proteínas de la cápside, de la cola, de la placa, de ensamble, y las relacionadas con la lisis, serán sintetizadas mediante la maquinaria celular. A partir de estas proteínas y de los genomas virales previamente sintetizados se construyen las

partículas virales, autoensamblándose siguiendo la entropía máxima de ese subsistema. El resultado es una rica progenie de viriones caudados (simetría compleja) que serán expulsados mediante la lisis celular causada por lisozimas de origen viral, repitiéndose el ciclo con cada virión que infecte otra célula y que entre en ciclo lítico (Friedman y Gottesman, 1983).

Ciclo lisogénico. Esta estrategia de sobrevivencia está ampliamente distribuida en todo el mundo viral y se ha estudiado ampliamente en los fagos. Es el proceso en el que el material genético de origen viral (ya sea de ADN o ARN) se almacena en forma de ADN dentro de la célula hospedera. Este ADN de origen viral (almacenado como episoma o integrado al cromosoma) se replica junto con el material genético de la célula, permaneciendo allí por “n” generaciones hasta que algún estímulo externo (luz UV, temperatura, alguna droga, etc.) hace que el “virus” entre en fase lítica, e inicie la generación de partículas virales infecciosas (Bertani, 1953). Para ejemplificar este tipo de ciclo utilizaré el modelo del fago λ .

El fago lambda tiene un conjunto de genes que se dividen en tres paquetes; tempranos inmediatos, tempranos intermedios y tardíos. Los tempranos inmediatos incluyen a los genes N y *cro*, que se expresan cuando el fago ingresa a la célula huésped. N promueve la transcripción de 3 genes tempranos intermedios; CII y CIII (implicados en el ciclo lisogénico), y Q (implicado en el ciclo lítico) (Lewin, 2008).

La entrada al ciclo lítico o lisogénico se basa en una “lucha” entre la concentración de las proteínas *cro* y CII. La que se encuentre en mayor cantidad determinará la entrada al ciclo lítico o lisogénico. Cuando *cro* “gana” se reprime la síntesis de las proteínas CII y CIII, que estimulan la producción y permanencia de CI (represor del ciclo lítico), además, *cro* permite la síntesis de la proteína Q, que activa la transcripción de los genes tardíos (cola, cápside y lizosima). Cuando CII “gana” se promueve la síntesis de su potenciador CIII, estimulándose la producción y mantenimiento de CI, que inhibe el ciclo lítico mediante la represión de los genes tempranos inmediatos, así como de algunos tempranos intermedios (incluyendo a Q), además de estimular la síntesis y mantenimiento de CII y CIII, lo que genera un circuito autorregulado donde CII y CIII activan a CI, y a su vez, CI activa a CII y CIII, lo que mantiene el genoma viral en estado latente, es decir, como provirus ó virus en estado lisogénico (Lewin, 2008).

Analizar la teoría del ciclo lisogénico desde un enfoque ecológico-evolutivo es sumamente interesante. La gente no especializada piensa que la única función de los virus radica en infectar células para reproducirse y dispersarse, matando a su hospedero en el proceso. Sin embargo, al analizar ese comportamiento desde un enfoque ecológico-evolutivo, resulta demasiado ineficiente, incluso absurdo, ya que indubitablemente llegará el momento en el que los viriones eliminen por completo a sus hospederos, provocando su propio exterminio (Abedon, 2008).

Muchas investigaciones han llegado a la conclusión de que la mayor parte de las partículas virales (si no es que todas) “necesitan” vivir como provirus para tener mayor éxito y permanecer a lo largo del tiempo. Además se les considera como grandes

contribuyentes de la evolución biológica, ya que actúan como vectores de información genética, activadores de rearrreglos genómicos y moduladores poblacionales dentro de comunidades biológicas. Se podría deducir que el impacto viral sobre la vida es mucho más profundo si se mira desde un enfoque ecológico-evolutivo, cuestión que se ha comenzado a explorar profundamente, principalmente mediante el estudio de los bacteriófagos (Abedon, 2008).

BACTERIÓFAGOS (FAGOS)

Se piensa que todas las bacterias pueden ser infectadas por bacteriófagos, que a pesar de su abundancia, poco se conoce de su diversidad, evolución y estructura genómica. La mayor parte de lo que se sabe proviene de estudios moleculares en sistemas modelo como λ , Mu, tipo T (T4, T7) y tipo P (P1, P2, P22), que infectan células de *Escherichia coli* y entran en estado lítico o lisogénico.

Se han identificado algunos factores que promueven la lisogenia. Por ejemplo, la probabilidad de encontrar profagos (genomas virales integrados a genomas bacterianos) en bacterias con genomas pequeños (menores a 2.5 Mb) es pequeña, aumentando en genomas mayores a 4 Mb (Touchon *et al.*, 2017). También se ha observado que las bacterias lisogénicas aparecen con mayor frecuencia en condiciones de baja densidad celular, bajos nutrientes y baja temperatura, es decir, en condiciones de crecimiento pobres. Por otro lado, la lisogenia se ha asociado a la presencia de algunos genes virales específicos como integrasas, represores y antirepresores, sistemas de restricción-modificación, y sistemas CRISP-Cas.

Los profagos son capaces de conceder fenotipos únicos al hospedero que infectan, lo que puede contribuir a la adaptación del mismo. Este fenómeno (conocido como conversión lisogénica) puede proporcionar inmunidad a la superinfección de fagos homólogos, expresar toxinas que pueden participar en la patogenia de especies, entre otras cosas (Canchaya, 2003; Abedon, 2008).

Por otro lado, los bacteriófagos más estudiados (tanto a nivel lítico como a nivel lisogénico) son los que tienen que ver con una aplicación más directa a la industria, por ejemplo, las bacterias lácticas (*Lactobacillus*, *Lactococcus*, etc.), y patógenas (*Staphylococcus*, *Mycobacterium*, *Streptococcus*, etc.). Por lo tanto, es normal que los genomas completos registrados en bases de datos como el GenBank muestren una mayor cantidad de estos tipos bacterianos. De hecho de los aproximadamente 2,000 genomas de fagos depositados en el GenBank, la mayoría se describieron en géneros bacterianos tales como *Mycobacterium*, *Synechococcus*, *Pseudomonas*, *Staphylococcus*, *Bacillus*, *Lactococcus*, *Vibrio*, *Propionibacterium*, *Escherichia*, *Salmonella*, *Listeria*, *Streptococcus*, *Enterobacter* y *Lactobacillus* (Hatfull y Hendrix, 2011).

Los párrafos anteriores explican las generalidades de los bacteriófagos, sin embargo, más adelante se detallarán conceptos que son relevantes para contextualizar el presente

trabajo. Entre ellos se hablará del impacto de los profagos en los ecosistemas y en la evolución de los seres vivos, además de características propias de fagos y profagos, tales como estructura, función clasificación, dinámica, entre otras.

Importancia de los bacteriófagos en los ecosistemas. En el mundo existen alrededor 10^{31} partículas virales que habitan todos los linajes celulares. Estas partículas viven en promedio 48 horas y se generan a una tasa aproximada de 10^{27} por minuto, hablando a nivel global. Al ver estos números se puede vislumbrar el impacto que provocan los virus en el mundo vivo; por ejemplo, se estima que 100 millones de toneladas cubicas de microbios mueren cada minuto a causa de lisis virales. Esto pone a los virus como importantes depredadores microbianos. Tan es así que tienen la capacidad de influenciar los ciclos biogeoquímicos globales, dirigiendo la ecología y evolución microbiana, sin embargo, su impacto se subestima frecuentemente. La mayor parte de este conocimiento viene de los estudios de comunidades microbianas marinas (Rhower *et al.*, 2009).

En los océanos marinos, tanto los organismos fotosintéticos (productores primarios) como los organismos heterótrofos son infectados por virus, que en muchos casos matan a sus hospederos. Las células lisadas se convierten en materia orgánica disuelta que puede ser usada por otras bacterias heterótrofas. Esto significa que la mortalidad provocada por los virus incrementa la respiración neta, liberando CO_2 y reciclando materia orgánica. Esto resalta el gran impacto que tienen los virus sobre los ciclos biogeoquímicos de los océanos (Rhower *et al.*, 2009).

Rastelli *et al.*, en 2017 demostraron que la inducción lisogénica de bacterias hidrotermales causa una amplia mortalidad celular y reduce significativamente la emisión de CO_2 . Lo que lleva a la generación de nuevos organismos fotosintéticos y a un aumento en la nutrición y metabolismo de los organismos heterótrofos. Esto sugiere la amplia participación de los virus inducidos en los procesos de producción primaria y secundaria de organismos marinos, principalmente de los que se encuentran cerca de las ventilas hidrotermales.

Sin embargo, no es muy conveniente que los virus maten a sus hospederos, lo ideal es coevolucionar con ellos hasta llegar a un tipo de simbiosis donde ambos “con-vivan” a través del tiempo, ya que de otra forma los virus terminarían por exterminar a sus propios hospederos. (Rohwer *et al.*, 2009). De hecho, se ha visto que la interacción lisis-lisogenia en las comunidades bacterianas es indispensable para la función correcta de los ecosistemas microbianos y, por ende, de todo el ecosistema; además de ser fundamental en los mecanismos de intercambios genéticos que dirigen la evolución (Abedon, 2008).

Impacto de los bacteriófagos en la evolución bacteriana. Pensando en la gran interacción genética mediada por vectores virales, vale la pena mencionar la teoría propuesta por Koonin *et al.*, en el 2006. Ellos sugieren que los agentes autorreplicantes y los virus se generaron antes de que surgieran las células. Hablan de un pool genético primitivo, donde existían todo tipo de genes que se replicaban y recombinaban constantemente, viajando y protegiéndose con biomoléculas. Dicho intercambio masivo

dio origen a los linajes primarios de agentes autorreplicantes y virus, primero de RNA y después de DNA, seguidos por el primer grupo de células primitivas.

Dichos virus, agentes autorreplicantes y células, comenzaron una coevolución; seleccionándose los virus y agentes autorreplicantes menos agresivos, ya que su mejor forma de sobrevivir era protegiéndose en el interior de una célula. Así se originó un segundo pool genético, donde participaron todos los agentes autorreplicantes y virus poco agresivos, además de las células. Allí se llevó a cabo una gran diversidad de mecanismos de intercambio, separación y fusión de material genético, así como diversos tipos de endosimbiosis. A partir de este segundo gran intercambio surgieron los principales linajes celulares (incluyendo al Dominio Eucaria), virales y de agentes autorreplicantes, que después fueron (y continúan) coevolucionando de una forma mucho más lenta (Koonin *et al.*, 2006).

Se piensa que la coevolución célula-virus se debe, en gran parte, al inmenso intercambio de material genético que se da a través de la transducción, ya que aunque existen otros mecanismos de transferencia horizontal y de evolución, la transducción ha demostrado ser (debido a su alta frecuencia) uno de los principales motores de la evolución. Prueba de ello es que en la mayoría de genomas secuenciados se encuentran secuencias homólogas a las de virus (Nasir *et al.*, 2014).

Esta gran capacidad viral de transferencia horizontal, sumada a que probablemente los virus sean el reservorio más grande de información genética de la biosfera, hace pensar en los virus como uno de los principales motores de la evolución. Además, estas entidades se encuentran en cualquier lugar donde existan células: aguas sulfurosas, ventilas hidrotermales, rizosfera, intestino, sangre, etc. Se piensa que dirigen la evolución de los seres vivos mediante la selección de organismos resistentes a los virus que los atacan y mediante transferencia horizontal de material genético (HGT), realizando esta última incluso entre especies filogenéticamente alejadas (Rohwer *et al.*, 2009).

Por ejemplo, arriba de 10^{14} eventos de transducción pueden ocurrir por año en el estuario de Tampa Bay (Jiang y Paul 1998), y si a ello le sumamos las partículas virales que sirven como agentes de transferencia de genes (GTAs), que no son propiamente virus, esta tasa podría aumentar hasta un millón de veces (McDaniel *et al.*, 2010). Otro ejemplo se encuentra en las ventilas hidrotermales, donde la dinámica, diversidad y densidad de microorganismos, así como su flujo constante, subyacen un frecuente intercambio de genes. Elsaied *et al.*, 2007, así como Brazelton y Baross (2009) han demostrado que los genes responsables para este proceso, incluyendo transposasas e integrasas, se encuentran abundantemente (en comparación con otros ecosistemas) en muestras metagenómicas celulares de ventilas hidrotermales.

Otros análisis genómicos han concluido que entre procariontes termófilos ocurre una transducción mucho más frecuente que entre organismos mesófilos o psicrófilos (Koonin, *et al.* 2001 y Beiko *et al.*, 2005), además han demostrado que la HGT ocurre entre organismos de diferentes dominios. La prevalencia de la HGT en las ventilas hidrotermales también puede ampliar el fenotipo de algunas especies bacterianas,

umentando su pangenoma y dándoles acceso a diferentes nichos ecológicos, lo que les podría proveer de una gran ventaja evolutiva (Anderson *et al.*, 2014).

Por otro lado, experimentos de inducción han sugerido que los profagos se inducen ampliamente en los genomas de procariontes de las ventilas hidrotermales. Sin embargo se ha visto que estos organismos juegan un papel mucho más importante que en el de la dinámica lisis-lisogenia. Se trata de su endogenización “total”, es decir, su integración al genoma del hospedero, lo que se debe a que el profago pierde o muta alguna(s) secuencias genómicas importantes, haciéndole casi imposible su entrada al ciclo lítico (o sea, su inducción), lo que lo convierte en parte del genoma celular, haciéndose participe de la evolución del mismo, ya que puede provocar cambios en los patrones de expresión génica, dirigir rearrreglos genómicos, etcétera (Arkhipova *et al.*, 2012). Estas actividades pueden llevar a la generación de nuevos genes, así como a modificaciones en la regulación de los ya existentes (Forterre, 2013).

Taxonomía y diversidad de fagos. Según la ICTV, los virus se clasifican de acuerdo con el rango de hospedero, ciclo de replicación, la estructura del virión (morfología), tipo de material genético (RNA o DNA de cadena sencilla o doble) y secuencia genómica (o de proteínas). En este sistema, los virus se han clasificado en 7 órdenes, 112 familias, 610 géneros, y 3,704 especies (Adams *et al.*, 2017), en donde la especie es definida como un “grupo monofilético cuyas propiedades pueden distinguirse de otras especies por múltiples criterios”. Estos criterios son sumamente laxos y se nota la ausencia de un elemento unificador y cuantitativo para establecer el parentesco entre distintos aislamientos de virus.

La clasificación tradicional de los bacteriófagos se basa en los factores antes mencionados, además de algunas estructuras auxiliares (colas, envolturas, etcétera). La mayor parte de los fagos tienen DNA de doble cadena (dsDNA), aunque existen algunos con DNA de cadena sencilla, RNA de doble cadena y RNA de cadena sencilla (ssDNA, dsRNA y ssRNA respectivamente). Así mismo, la mayoría corresponden al orden de los caudovirales (cápsides con cola), es decir, fagos con una estructura compleja. Los restantes tienen estructuras filamentosas, pleomórficas y poliédricas (Orlova, 2012).

Las estructuras virales se han determinado mediante su visualización con el microscopio electrónico. Para el 2011 se habían analizado 6000 fagos aproximadamente, de los cuales alrededor de 5500 son complejos (caudados) y el resto son de estructuras filamentosas, poliédricas y pleomórficas (cúbicas), teniendo las dos últimas estructuras icosaédricas internas (Ackerman, 2011).

Los fagos se clasifican en 1 orden (con tres familias) y 7 familias extras (sin orden). El único orden corresponde a los caudovirales (fagos con cola), ver figura I1, que tienen dsDNA como material genético y que se dividen en las familias Myoviridae (colas largas contráctiles), Siphoviridae (colas largas no contráctiles) y Podoviridae (colas cortas), ver figura I3 (B-D). Tienen una estructura icosaédrica y, entre las tres, suman el 96% de todos los fagos (Ackerman, 2011).

Las 7 familias restantes tienen diferencias muy pronunciadas, tanto en tipo de material genético como en estructura morfológica, por lo que no cuentan con problemas de clasificación. Para describirlos se mencionarán algunas de sus características: Microviridae, ssDNA, 27 nm de diámetro y capsómeros con 12 dominios tipo knob (estructuras de nudo que funcionan como ligandos en receptores celulares); Corticoviridae, dsDNA, 63 nm de diámetro y lípidos asociados a cápsidas complejas; Tectiviridae, dsDNA, 60 nm de diámetro, pseudocola y vesículas lipídicas internas; Leviviridae, ssRNA, 23 nm de diámetro y una estructura parecida al virus de la polio (simetría icosaédrica); Cystoviridae, dsRNA segmentado, 70-80 nm de diámetro y envoltura lipídica; Inoviridae (figura I3-A), ssDNA y 85-1950 nm por 7 nm (estructura filamentosa); y Plasmaviridae, dsDNA, 80 nm de diámetro y no cuentan con cápsida sino con cubierta lipídica (Ackerman, 2011).

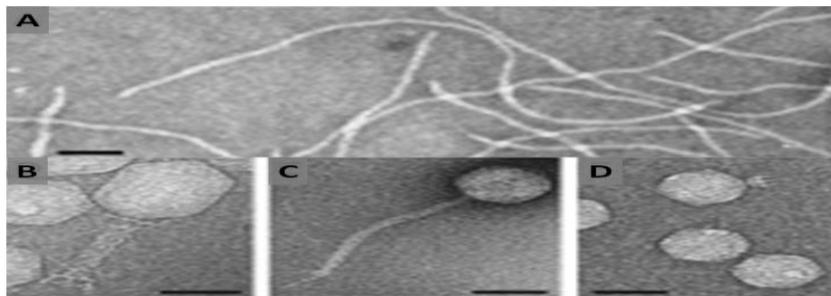


Figura I3. A: Fago filamentoso B5, perteneciente a la familia Inoviridae. B: Bacteriophage T4, perteneciente a la familia Myoviridae. C: Fago SPP1, perteneciente a la familia Siphoviridae. D: Fago P22, perteneciente a la familia Podoviridae. Las barras negras representan 50 nm de longitud (Modificado de Orlova, 2012).

El sistema tradicional del ICTV ha sido alcanzado por la revolución en los métodos de secuenciación masiva, que han permitido examinar con más profundidad la amplitud de la virósfera mediante metagenómica, aun con la falta de información de las propiedades biológicas y estructurales de las especies de virus (Simmonds *et al.*, 2017). Las secuencias genómicas aportan gran cantidad de información que puede usarse para construir relaciones evolutivas entre diferentes especies de virus, lo que puede realizarse mediante la comparación de su organización genética, uso de codones, contenido de GC, etcétera. Además, los métodos de agrupamiento (clusterización) como el promedio de identidad nucleotídica (similar al ANIm aplicado a bacterias) o la distancia filogenética basada en el Genome Based Distance Phylogeny (GBDP), (Meier-Kolthoff and Goker, 2017) comienzan a proporcionar una visión mucho más coherente de la clasificación viral.

Estructura genómica de fagos. Los estudios actuales muestran que la estructura del genoma de los fagos comprende un “cromosoma” en mosaico, con elementos genéticos de diferentes orígenes que están sujetos a frecuencias altas de ganancia y pérdida, mediadas por transferencia horizontal y recombinación no homologa (Pedulla *et al.*, 2003; Hendrix *et al.*, 1999). Sin embargo, los estudios no han sido realizados homogéneamente debido a que la colección de estos genomas no es muy amplia ni homogénea, por ejemplo: no se han encontrado colecciones de fagos pertenecientes a un sitio de aislamiento (simpátricos) y; hay pocos estudios relacionados con secuencias pertenecientes a fagos que infectan un solo huésped.

El estudio de Hendrix *et al.* (2003) es más completo. Consistió en la formación de 26 agrupaciones (*clusters*) a partir de 1,401 genomas de fagos que infectan a *Mycobacterium smegmatis*. Cada *cluster* se eligió a partir de identidades de ADN genómico viral, contenido de genes y clase de genes, lo que generó unidades (*clusters*) discretas. Sin embargo, en una comparación más detallada a partir de solo 627 genomas de demostró que aún entre *clusters*, se comparte un número sustancial de genes divergentes, es decir, se concluyó que los genes de diferentes *clusters* son variantes de un ancestro común, y no “*clusters*” discretos.

La última observación contrasta con un estudio de fagos que infectan una sola especie de *Synechococcus*, donde se observó claramente que cada fago representa un *cluster* discreto, aislado genéticamente de los demás (Deng *et al.*, 2015). Lo mismo se ha reportado para fagos que infectan a la familia Enterobacteriaceae, ya que la mayoría de ellos carecen de homología genética, por lo que solo representan grupos individuales (Grosse y Casjens, 2014).

Estos conceptos en contradicción podrían resolverse cuando se disponga de un mayor número de secuencias genómicas de fagos que infecten al mismo grupo bacteriano. Además, es necesario el desarrollo de métodos y criterios para: primero, diseñar muestreos homogéneos con el propósito de comparar la estructura y diversidad genómica entre comunidades; segundo, identificar las relaciones de homología entre los genes de diferentes fagos; y por último, hacer una evaluación de la tasa de recombinación homóloga contra la tasa de recombinación no-homóloga; entre otros enfoques.

Resistencia bacteriana a los fagos. Hasta la fecha se han descubierto diversos mecanismos bacterianos de resistencia a fagos, entre ellos encontramos los siguientes: modificación de la producción o estructura de los lipopolisacáridos que forman el antígeno O de la membrana celular; modificación de los receptores fágicos de membrana para impedir la penetración del material genético viral; sistema de restricción modificación CRISP-Cas; entre otros. Dichos mecanismos se han comprobado en diversos estudios, por lo que su conocimiento es bastante sólido.

Hasta este punto se han revisado los conceptos generales de virus y de bacteriófagos. A continuación se profundizará más acerca de los profagos y posteriormente se mencionarán una serie de estudios que hablan un poco de los fagos relacionados con las bacterias del genero *Rhizobium*.

PROFAGOS

La palabra profago deriva del vocablo pro (antes) y fago (virus que come bacterias), por lo que un profago es una entidad que precede a los fagos, la que esta antes de que se produzcan las partículas virales llamadas fagos. Se encuentra en forma de ADN y pueden estar integrados al cromosoma o permanecer como episomas. Los profagos completos (no degradados o mutados) pertenecen a la fase lisogénica de un bacteriófago que, por

tener la característica de no dañar a su hospedero, también se les conoce como fagos aatemperados (Canchaya *et al.*, 2004).

La integración y escisión del un profago usualmente se da mediante una tirosina o serina recombinasa codificada por el mismo profago, aunque en algunos casos la proteína puede ser codificada por otro profago o alguna secuencia móvil (Nunes-Duby *et al.*, 1998; Smith and Thorpe 2002). Otros profagos (Mu por ejemplo) se integran y se transponen al azar mediante transposasas DDE (Mizuuchi 1992). Algunos, llamados satélite, tienen las secuencias que sintetizan proteínas para formar partículas virales completas, sin embargo, las producen para otro virus, por ejemplo, P4 produce viriones de P2 (Six y Klug, 1973). Como último ejemplo se citará a la familia Inoviridae (ssDNA), cuya integración la media una recombinasa del hospedero (Huber y Waldor, 2002), (Citados en Louis-Marie *et al.*, 2012).

Clasificación de profagos. Con respecto a la clasificación de profagos, aun no existe un consenso claro entre la comunidad científica, a pesar de los avances en genómica tanto en secuencias disponibles como en herramientas informáticas para analizarlas. Sin embargo el ICTV clasifica los profagos basándose en las características de los siguientes profagos tipo: Profagos tipo-lambda, tipo-P22, tipo-P2, tipo-Epsilon15, tipo-PhiC31, tipo-Mu, tipo-P4 e Inovirus. Su clasificación se basa en características genómicas (sintenia) el tipo de material genético y su morfología (Louis-Marie *et al.*, 2012).

Por otro lado, Casjens, en 2003, dividió a los profagos en 5 grupos: 1) Profagos funcionales, es decir, profagos que pueden pasar a fase lítica; 2) Profagos defectuosos, también llamados profagos crípticos, que se encuentran en decaimiento mutacional pero que pueden contener genes funcionales; 3) Profagos satélite, que producen proteínas de cápside correspondientes a otros profagos, es decir, ayudan a formar viriones no propios; 4) Bacteriocinas, secuencias homologas a secuencias fágicas conocidas (cola de lambda por ejemplo) pero con la capacidad de matar a otras bacterias; 5) Agentes de transferencia genética (GTAs), que son profagos con la capacidad de formar partículas virales, pero no con el material genético del profago, sino con cualquier parte del genoma bacteriano (en un tipo de empaquetamiento al azar) que llega completo a otra bacteria de la misma especie (o muy similar) y que se intercambia por recombinación homologa con la secuencia propia del segundo hospedero, replazándola.

Por otro lado, Brussow y Hendrix (2002) hacen notar que la clasificación de profagos basada en comparaciones genómicas podría ser más adecuada. Ellos mencionan algunos ejemplos: Los profagos del Podovirus P22 (fago de cola corta) y de Siphovirus lambda (fagos de cola larga) normalmente se clasifican en grupos diferentes, sin embargo tienen una sintenia lo suficientemente grande para clasificarlos en el mismo grupo; Algunos profagos de tipo arqueovirus tienen una estructura genética y morfología muy similar a la de los fagos caudados, lo que debería ser suficiente para clasificarlos dentro de un mismo grupo, lo que no sucede; Los profagos de bacterias ácido lácticas, micobacterias y los fagos tipo lambda tienen grandes sintenias, por lo que todos podrían clasificarse en un mismo grupo.

Un trabajo más reciente (Louis-Marie, 2012) da cuenta de un análisis profundo de aproximadamente 500 profagos. Primero ubicó las secuencias profágicas mediante programas bioinformáticos (Page finder, PHAST y profinder). Después los agrupó mediante una matriz que comparó secuencias génicas mediante el algoritmo de Needleman-Wunsch. También comparó proteínas y realizó análisis sintéticos, además de tomar en cuenta las clasificaciones del ICTV. Los resultados que arroja son muy diversos, por lo que si se quiere ahondar se recomienda revisar la referencia.

La clasificación de profagos es un tema joven que se ha desarrollado lentamente, en conjunto con la evolución de las ciencias genómicas. Los trabajos más recientes buscan ordenar estas secuencias basados en el mayor número de factores, tomando en cuenta el pangenoma, panproteoma, fenotipos, movilidad, etcétera.

Inducción de profagos. El movimiento de los profagos dentro de los genomas y dentro los ecosistemas es de suma importancia para la biotecnología ya que con dichos estudios se contribuye al conocimiento concerniente a la manipulación de genomas. Comprendiendo los mecanismos que subyacen la movilidad de las diferentes secuencias genómicas se pueden generar diferentes herramientas que sean muy valiosas para procesos médicos e industriales. Por otro lado, debido a la economía en tiempo y dinero, además de todo el avance en el campo, trabajar con fagos y bacterias es una opción muy viable para investigar los mecanismos de movilidad de secuencias genómicas.

A la activación de un profago para que pase del estado lisogénico al lítico se le conoce como inducción lisogénica, fenómeno que ocurre cuando las bacterias lisogénicas se inducen, ya sea por mecanismos desconocidos (“espontáneamente”) o conocidos (luz UV, mitomicina C, etcétera). En secciones como metodología y resultados se mencionarán estudios de inducción con diferentes inductores.

Conversión lisogénica. Los profagos completos se encuentran en estado latente hasta que un estímulo los hace entrar en fase lítica, como se menciona en los párrafos de arriba, sin embargo, muchos mutan o pierden secuencias importantes, haciéndoles casi imposible entrar en estado lítico, estos profagos se denominan “profagos incompletos”. Ambos tipos de profagos se encuentran en la mayor parte de las bacterias secuenciadas y pueden cubrir porcentajes importantes dentro del genoma del hospedero (Abedon 2008). Su presencia puede tener diversos efectos en las células, por ejemplo: pueden contener genes con factores patogénicos que hacen que la bacteria hospedera pueda colonizar organismos que antes no podía; otros portan genes de Bacteriocinas que tienen la capacidad de atacar otras bacterias para ganar espacio y/o alimento (Louis-Marie *et al.*, 2012).

Dichos efectos se conocen como conversión lisogénica. Se trata de modificaciones en el metabolismo de la bacteria que porta cierto profago. Estas modificaciones pueden ser tanto positivas como negativas y son uno de los principales motores de la evolución.

Otros ejemplos de conversión lisogénica son los siguientes: *E. coli* O157:H7 adquiere la toxina shiga del genoma de su profago Sp5 o Sp15 (Ohnishi *et al.*, 2001); *Vibrio cholerae*

adquiere la toxina del cólera de su profago CTX ϕ (Waldor y Mekalanos, 1996); Los profagos L54a y phi13 de *Staphylococcus aureus* causan la inactivación del gen de lipasa y β -toxina respectivamente (Coleman *et al.*, 1991).

Es importante resaltar que en la literatura científica actualmente existe una gran cantidad de trabajos que ejemplifican la conversión lisogénica, sin embargo los anteriores son suficientes para explicar este proceso. Además en estos últimos años (5 aproximadamente) han crecido las investigaciones acerca de coevolución fago-bacteria en sistemas sólidos.

BACTERIOFAGOS EN *Rhizobium*.

Con respecto a la investigación de lisógenos en bacterias rizobiales, los estudios se han enfocado principalmente en la dinámica entre fagos líticos y lisogénicos compitiendo por un huésped en un nicho (suelo), así como la identificación de fagos lisogénicos en cepas rizobiales de importancia agronómica (Kimura *et al.*, 2008).

Otros estudios de importancia se enfocan en: la identificación de fagos que se convierten en lisógenos de cepas específicas; el mecanismo de inserción del genoma fágico en el cromosoma del hospedero; en la identificación de las diferentes propiedades bioquímicas de cepas lisogénicas en comparación con cepas no lisogénicas y; la repercusión (a nivel de producción agrícola, por ejemplo) que tienen los fagos atemperados sobre sus hospederos. Algunas referencias se encuentran a lo largo de este texto, o se pueden buscar en la bibliografía.

Por ejemplo, la inoculación en la rizosfera de trébol, con dos cepas bacterianas y un fago de *Rhizobium trifolii*, provocó que la población de la cepa susceptible se redujera en un orden de magnitud, y que la formación de nódulos se redujera unas 4 veces aproximadamente (Evans *et al.*, 1979). Esto sugiere una gran capacidad de los fagos rizobiales para la selección de cepas.

En un estudio realizado en 1989 por Uchiumi *et al.*, se encontró un fago (denominado fago U-Mole) que lisogeniza a la cepa de *Rhizobium leguminosarum* biovar trifolii UK. Adicionalmente se observó que el fago también lisogeniza a la cepa 4S de la misma especie, produciendo en ella la pérdida de un plásmido de 315 Kbs y su capacidad para nodular.

La cepa de *Rhizobium trifolii* SU297, lisogenizada con el fago 7 perdió la habilidad de absorber tanto a su fago hospedero (fago 7) como a los fagos 7cr y 8. Además se comprobó (mediante la comparación con una cepa no lisogenizada) un cambio sustancial en su superficie antigénica (receptores membranales conocidos) (Barnet y Vincent, 1970). Esto sugiere que la incapacidad que tiene la cepa lisógena de absorber a los fagos citados se debe a un cambio (o ausencia) de algunos receptores en su superficie membranal, los cuales podrían funcionar como receptores específicos para los fagos.

Otro estudio mostró que de una colección de 54 cepas (de las especies *R. leguminosarum*, *R. trifolii* y *R. phaseoli*), 7 se encontraban lisogenizadas. Dos de estas (de *R. leguminosarum*) produjeron agentes letales parecidos a bacteriocinas. La manera en que comprobaron la existencia de estos agentes se basa en la siguiente técnica: se hacen cultivos bicapa en placas de agar, cada placa contiene una cepa diferente, la de abajo contiene una cepa lisógena y la de arriba una no lisogena; cuando se induce la liberación del fago lisógeno (primera capa) con rayos X se produce una placa lítica en la segunda capa, sugiriendo que el fago es lítico para la segunda cepa, posteriormente se reproduce el fago mediante el método de enriquecimiento (Barnety Vincent, 1970), el filtrado obtenido (aparentemente con el fago) se comprueba para lisis en la segunda cepa. Cuando se observó que no existían placas líticas en la infección a la segunda cepa, se sugirió que la lisis en el experimento bicapa la provocó un agente (con actividad de bacteriocina) sintetizado por la primera cepa (Shwinghamer y Reinhardt, 1963).

Con respecto a la integración genómica del ADN fágico en su huésped, Semsey *et al.*, en 1999, encontraron dos genes importantes en el fago lisogénico 16-3 de *Rhizobium melliloti* 41. Tales genes participan en la integración y liberación del DNA fágico dentro del cromosoma de su huésped; los genes encontrados fueron *int* (integración del DNA fágico en un sitio especificado por secuencia) y *xis* (escisión del profago).

En un estudio reciente se probaron 21 cepas de *Rhizobium leguminosarum* para producción nodular, fijación de nitrógeno y peso seco de su planta huésped (la lenteja), cuatro de ellas (HUL-57, L41-47, K-75 y PL-4/DPL-6) mostraron una mayor cantidad en los tres parámetros. Lo interesante del trabajo fue que esas cuatro cepas eran sensibles a fagos, en comparación con otras cepas inmunes a fagos, cuyos resultados en los tres parámetros mencionados fueron bajos. Esto muestra que los fagos pueden contribuir con la selección de cepas rizobiales para su uso en la producción de cultivos exitosos (Jaiswal y Dhar, 2011).

Bacteriófagos infectivos y lisogénicos de *Rhizobium* se conocen desde los años cuarenta y se han aislado paulatinamente para casi todas las especies de *Rhizobium*, *Sinorhizobium*, *Bradyrhizobium* y *Mesorhizobium*. Pocos de ellos se han caracterizados genómicamente. Al 2014, se encontraron 17 secuencias genómicas completas de fagos de estos géneros depositados en el Genbank. De éstos, 9 pertenecen a *R. etli* y el resto a *Sinorhizobium*, *R. leguminosarum* y *R. gallicum*. Estos fagos son líticos en su mayoría, y algunos (Por ejemplo, el fago 16-3 de *Sinorhizobium*) son capaces de transducir material genético del huésped. La comparación genómica entre los genomas de estos fagos no revela gran conservación sino una estructura modular con bloques de genes que codifican las proteínas de lisis, morfogénesis, síntesis de ADN, y una gran proporción (70%) de genes hipotéticos y huérfanos sin función asignada.

Santamaría *et al.*, en 2013 realizaron un trabajo para la caracterización morfológica como genómica de los fagos (principalmente líticos) de 48 cepas de *Rhizobium etli* de diferentes regiones. Aislaron 14 fagos en total, de los cuales secuenciaron el genoma de 9. Algunos resultados sobresalientes con respecto a las características de los fagos son los siguientes: los fagos tienen un rango de hospederos estrecho y variable; el tamaño del

genoma varió de 45 a 50 kbs (en un fago fue de 115 kbs); 30 a 35% de su DNA codifica para proteínas de morfogénesis fagica y procesamiento y síntesis de DNA; los fagos encontrados son los primeros reportados para *Rhizobium etli* y se clasificaron como RHph01 al RHph14; un grupo de fagos (definidos como P1-like) mostró fagos lisógenos con secuencias génicas parecidas a genes bacterianos (asociados al metabolismo, regulación y transporte).

JUSTIFICACIÓN

Como se ha mencionado, los fagos en su estado lisogénico, o fagos atemperados, han mostrado ser sumamente importantes en la estructuración de ecosistemas y en la evolución de bacterias. Además, se ha visto que los estudios de fagos y bacterias han aportado conocimientos fundamentales a nivel básico y aplicado, como por ejemplo, el mecanismo de inmunidad CRISPR-Cas (Barrangou *et al.*, 2007), cuyo impacto en la ciencia aplicada está dando resultados importantes.

En el caso de los ecosistemas del suelo, y en específico, en los que se encuentran bacterias fijadoras de nitrógeno, sería sumamente enriquecedor entender los mecanismos que subyacen a la modulación de sus ecosistemas, incluyendo la abundancia de ciertas cepas bacterianas. En este sentido, se ha encontrado que gran parte de la modulación de dicha abundancia está determinada por un mecanismo denominado predador-presa, mediante el cual las densidades poblacionales están controladas por la alternancia entre el ciclo lítico y lisogénico de los fagos que infectan a dichas poblaciones (Chibani-Chennoufi *et al.*, 2004).

El descubrimiento de este mecanismo (y otros), además de generar conocimiento *per se*, podría ser aplicado a la agricultura. Por ejemplo, al añadir algún o algunos tipos específicos de fagos a campos de cultivo, se podrían controlar las densidades poblacionales de bacterias que participan en el crecimiento (incluso en condiciones hostiles) de vegetales usados para la alimentación de humanos o animales.

Sin embargo, los estudios de microorganismos del suelo son muy escasos, y aún más los relacionados a fagos que se asocian a cepas de *Rhizobium*. Algunos de estos estudios, principalmente el de Santamaría *et al.*, en 2013, se mencionan en los párrafos de arriba. No obstante, los estudios con respecto a la interacción lisis-lisogenia entre fagos y bacterias de *Rhizobium* se comenzaron a realizar recientemente en nuestro laboratorio, y este trabajo muestra parte de algunos resultados preliminares.

OBJETIVOS

GENERAL

Analizar la lisogenia en *Rhizobium*.

PARTICULARES

- a) Buscar cepas de *Rhizobium* lisógenas mediante inducción experimental de profagos.
- b) Determinar mediante análisis bioinformáticos la presencia de profagos en algunos genomas de *Rhizobium*

METODOLOGÍA

Para contribuir a los objetivos planteados se realizaron trabajos de laboratorio y bioinformáticos. Los primeros estuvieron basados en cultivos bacterianos, técnicas clásicas para analizar fagos y técnicas de biología molecular como asilamiento de ADN y PCRs. Los objetivos bioinformáticos se lograron gracias al uso de algunas plataformas como PHAST, MEGA 5.0 y Oligo 7.0, y del lenguaje de programación Phyton. Estas dos vertientes se combinaron y acoplaron de diferentes maneras y en diferentes momentos del proyecto, todo para que poco a poco el proyecto adquiriera un sentido cada vez más coherente, cubriendo (en la medida de lo posible) los objetivos estipulados.

Es importante recordar que las siguientes técnicas están específicamente diseñadas para cepas del genero *Rhizobium*. Los métodos más generales se basan en *E. coli* y aplicarlos a otras especies requiere de modificaciones, a veces mínimas, a veces considerables.

CULTIVOS

Cultivo líquido: El medio se compone de caldo PY adicionado con 7 mM de CaCl₂, 10 mM de MgSO₄ y 20 µg/ml de ácido nalidíxico. Para hacer crecer las bacterias se deben agregar 5 mL de caldo PY y verterlo a un tubo de ensayo de vidrio, posteriormente se añaden las células, ya sea diluyendo una asada recogida de un cultivo sólido o agregando 100 µl de un cultivo líquido previamente crecido. El cultivo se incuba a 30°C y a 250 rpm. El tiempo de incubación dependerá del objetivo del experimento. Los compuestos que se adicionan al caldo PY puro son para que las cepas *Rhizobium sp.* tengan un crecimiento óptimo. Para facilitar la lectoescritura de este trabajo, a toda la mezcla se le denominará PY, caldo PY o medio PY.

Cultivo sólido: Para formar los botones de agar en las cajas de petri estándar, agregar 1.5% de agar al medio PY.

Agar suave: Caldo TB (10 g/l de triptona, 5 g/l de NaCl y 0.65% de agar).

AISLAMIENTOS

Aislamiento de DNA bacteriano

- I. Crecer toda la noche en medio líquido la cepa de interés.
- II. Pasar 1.5 mL del cultivo a un tubo eppendorf y centrifugar 2 minutos a 13000 rpm, retirar el sobrenadante con micropipeta y volver a agregar 1.5 mL del cultivo, centrifugar nuevamente 2 minutos a 13000 rpm y volver a retirar el sobrenadante con micropipeta.
- III. Resuspender la pastilla con 1ml de TE 50/20 (50 mL de Tris-HCl, 20 mM y 20 mM de EDTA a pH 8) filtrado y esterilizado.
- IV. Centrifugar 2 minutos a 13000 rpm.
- V. Retirar el sobrenadante con una micropipeta.

- VI. Resuspender la pastilla con 550 μ l de TE 50/20 y mezclar con un vortex.
- VII. Agregar 20 μ l de lisozima a 20 mg/ml, 50 μ l de pronasa (5 mg/ml en TE 50/20 pre digerida 1h a 37°C) y 50 μ l de SDS al 10% (filtrado y esterilizado). Por cada agregado, mezclar rápidamente con vortex.
- VIII. Incubar por 20 minutos a 37°C. Verificar que este viscoso y transparente, o dejarlo incubar un poco más.
- IX. Incubar por 10 minutos a 80°C.
- X. Dejar enfriar a temperatura ambiente por aproximadamente 2 minutos.
- XI. Agregar 5 μ l de RNAsa (10 mg/ml disuelta en TE 10/1 pre tratado 10 minutos a 95°C).
- XII. Mezclar inmediatamente con vortex e incubar 10 minutos a 37°C.
- XIII. Agregar 220 μ l de acetato de amonio saturado (7 M) y mezclar durante 20 segundos con vortex.
- XIV. Incubar 5 minutos en hielo.
- XV. Centrifugar 5 minutos a 13000 rpm.
- XVI. Transferir 760 μ l del sobrenadante a otro eppendorf de 1.5 mL y añadirle 570 μ l de isopropanol frío al 100%.
- XVII. Mezclar varias veces por inversión.
- XVIII. Recuperar con una micropipeta de 10 o 100 μ l las hebras de DNA que se ven flotando como pequeños residuos de algodón y pasarlas a un tubo eppendorf que contenga 1ml de etanol al 70%.
- XIX. Centrifugar 3 minutos a 13000 rpm.
- XX. Retirar el sobrenadante con mucho cuidado (la pastilla es casi imperceptible) usando una micropipeta de 1 mL.
- XXI. Resuspender con 1ml de etanol al 70% y disolver la pastilla.
- XXII. Repetir los pasos XIX y XX.
- XXIII. Dejar el tubo a temperatura ambiente (de preferencia toda la noche) para que se evapore todo el etanol y la pastilla quede seca.
- XXIV. Resuspender con 100 μ l de agua destilada o TE10/1 (ambos filtrados y esterilizados).
- XXV. Añadir 1 μ l de RNAsa (10 mg/ml disuelto en TE 10/1, filtrado, esterilizado y pre digerido 10 minutos a 95°C).
- XXVI. Mezclar suavemente y almacenar a 4°C.

Para estimar la concentración de este compuesto se utilizó el Nanodrop (A280/A260). En el cual se usan tubos eppendorf pequeños para agregarles 5 o 10 μ l de la muestra de DNA, los resultados se visualizaron a través de un software especializado.

Aislamiento de partículas virales

Primero se propagó el fago (por inducción o por reproducción de fagos líticos) dentro de un crecimiento bacteriano toda la noche. Posteriormente, el cultivo se centrifugó a 10,000 rpm durante 10 minutos y a 4°C. Por último, el sobrenadante se filtró con una malla estéril

con poros de 22 μm de diámetro. El filtrado se almacenó en tubos nuevos grado molecular a 4°C.

Aislamiento de DNA viral

A partir del filtrado de partículas virales (punto anterior) se sigue la siguiente metodología.

- I. Añadir 100 μl de buffer L1 por cada 10 mL del concentrado de fagos y mezclar por inversión en tubos Falcon de 50 mL. *El buffer L1 debe mantenerse frío. Se prepara mezclando los siguientes reactivos: 300 mM de NaCl, 100 mM de Tris-HCl a pH 7.5, 10 mM de EDTA, 0.2 mg/ml BSA, 20 mg/ml de RNAsa A y 6 mg/ml de DNAsa 1.*
- II. Incubar 1 h. a 37°C y 250 rpm. *Después de este punto las muestras deberán mantenerse a 4°C en bandejas con hielos triturado.*
- III. Añadir 2 mL de buffer L2 por cada 10 mL de muestra y mezclar suavemente por inversión. *El buffer L2 debe mantenerse a 4°C. Consiste de 30% de polietilenglicol (PEG8000) y 3 M de NaCl.*
- IV. Incubar en hielo por 2 h.
- V. Centrifugar por 20 minutos a 10000 rpm y 4°C.
- VI. Decantar el sobrenadante y dejar escurriendo el resto por 1 minuto aproximadamente.
- VII. Resuspender la pastilla con 1 mL de buffer L3 por cada 10ml del concentrado inicial de fagos. *El buffer L3 se compone de 100 mM de Tris-HCl a pH8 y 25 mM de EDTA.*
- VIII. Pasar la solución resuspendida a tubos de ensayo de vidrio de 3 mL.
- IX. Añadir 1 mL de buffer L4 por cada 10 mL del concentrado inicial de fagos y mezclar suavemente por inversión. *El buffer L4 consiste de una solución de SDS al 4%.*
- X. Añadir 5 μl de proteinasa K por cada 10 mL del concentrado inicial de fagos y mezclar 4-6 veces por inversión.
- XI. Incubar 1 h. a 50°C.
- XII. Incubar 10 minutos a 70°C.
- XIII. Poner en hielo por 2 minutos.
- XIV. Pasar a tubo Falcon de 50 mL.
- XV. Añadir 1 mL de buffer L5 por cada 10 mL del concentrado inicial de fagos y mezclar suavemente por inversión de 4-6 veces. *El buffer L5 consta de una solución 3 M de acetato de potasio a pH 5.5.*
- XVI. Centrifugar por 30 minutos a 12000 rpm y 4°C.
- XVII. Recuperar el sobrenadante en tubos Falcon de 50 mL.
- XVIII. Centrifugar por 10 minutos a 12000 rpm y 4°C.
- XIX. Recuperar el sobrenadante en tubos eppendorf de 1.5 mL. Agregar 1 mL de sobrenadante a cada tubo.
- XX. Añadir 0.7 mL de isopropanol (100%) a cada tubo eppendorf y mezclar por inversión.
- XXI. Centrifugar por 30 minutos a 12000 rpm y 4°C.

- XXII. Decantar cuidadosamente ya que la pastilla es casi imperceptible.
- XXIII. Resuspender pastilla con 1 mL de etanol y mezclar con vórtex.
- XXIV. Centrifugar 10 minutos a 12000 rpm y 4°C.
- XXV. Decantar cuidadosamente.
- XXVI. Repetir los pasos w-y.
- XXVII. Secar 5 minutos en savant.
- XXVIII. Resuspender pastilla (cada eppendorf) con 35 µl de agua destilada.
- XXIX. Refrigerar toda la noche a 4°C.
- XXX. Las muestras de todos los tubos se puede juntar y almacenar a -20 °C.

Concentración de ADN

Se estimó la concentración de ADN con un espectrofotómetro Nanodrop A280/A260

Electroforesis para ADN

Como es sabido, para comprobar la amplificación por PCR de los segmentos de ADN en cuestión se necesita la electroforesis. Las siguientes condiciones electroforéticas fueron las mismas para las amplificaciones correspondientes al ADN bacteriano y viral.

- I. Se utiliza una cámara electroforética Horizon 58 de LifeTechnologies Model 200.
- II. El gel para la placa de corrida consiste en una solución de Tris-Acetato EDTA 1X con 1% de agarosa.
- III. El buffer de corrida consiste de Tris-acetato EDTA 1X. Este se agrega al interior de toda la cámara, de tal forma que cubra completamente el gel de corrida.
- IV. La corriente de corrida que se utiliza es de 100 V.
- V. En cada carril se agregan 10 µl de la siguiente solución: 5 µl de la solución sometida a PCR (“secuencia amplificada” o marcador molecular), 3 µl de H₂O grado HPLC, y 2 µl de azul de bromofenol.
- VI. En el carril 1 se coloca el marcador de peso molecular GeneRuler 1 kb plus DNA leader de Thermoscientific.
- VII. En los carriles posteriores se colocan las muestras.
- VIII. Aunque las muestras (incluyendo el marcador de peso) tienen el color azul como indicador de migración, el tiempo de corrida debe oscilar entre 40 minutos a 2 horas.
- IX. Después de la corrida, las muestras se revelan de la siguiente forma:
 - X. Se retira el gel y se pone a remojar de 5-10 minutos en bromuro de etidio (es importante usar guantes de latex debido a la toxicidad de este compuesto).
 - XI. Se escurre el gel y se revela en un fotodocumentador Gbox (Syngene).
 - XII.** Se toma una fotografía (el Gbox tiene la opción para hacerlo) y se observan las secuencias amplificadas, cuyo tamaño (comparado con el marcador) debe ser igual al de la secuencia que se busca amplificar.

La fotografía que da la Gbox se puede editar y posteriormente guardar como imagen, la cual puede ser utilizada para los fines que convengan.

FAGOS

Inducción

- I. A un cultivo bacteriano líquido se le agrega una cantidad establecida del inductor, lo que provocará que los profagos entren en fase lítica y se comiencen a producir partículas virales.
- II. El cultivo (ya con el inductor añadido) se incuba con agitación orbital, temperatura y tiempos establecidos (según el tipo de bacteria y profago que se busque inducir).
- III. Posterior a la incubación (generalmente de 2-6 horas después) se centrifuga la muestra para retirar las células (pastilla) y recuperar el sobrenadante con partículas virales.

Inducción con mitomicina C

- I. Se toman 100 μ l del stock (cepa al 1 M, almacenada a -20°C en una solución al 40% de glicerol) de alguna cepa lisogénica del género *Rhizobium* y se añaden a un tubo de ensayo con 5 mL de caldo PY.
- II. El tubo se incuba toda la noche a 30°C con agitación orbital de 250 rpm, obteniéndose cultivos saturados de la cepa.
- III. Se toman 100 μ l del cultivo saturado y se pasan a un tubo con 5 mL de medio fresco. Se agita en un vórtex y se incuba 40 minutos.
- IV. El tubo se retira del horno de cultivo y se le agrega mitomicina C para llegar a una concentración final de 2.5 $\mu\text{g/ml}$. El proceso se realiza en condiciones de oscuridad debido a la fotosensibilidad de la mitomicina C. *La concentración efectiva para inducir profagos en Rhizobium va de 0.5-2.5 $\mu\text{g/ml}$.*
- V. El tubo se incuba durante 4 h (proceso realizado en oscuridad). *El tiempo efectivo para la inducción de profagos de Rhizobium va de 4-6 horas.*
- VI. El cultivo se retira del horno y se pasa a tubos eppendorf de 1.5 mL para ser centrifugado a 12000 rpm durante 10 minutos y a 4°C .
- VII. Se recupera el sobrenadante, que probablemente contenga partículas virales. *Este puede ser usado directamente para otros experimentos o ser almacenado a corto plazo (4°C) o largo plazo (glicerol al 40% y a -20°C).*
- VIII. Para hacer una curva de crecimiento se preparan varios tubos y se someten a las mismas condiciones para después ser retirados a diferentes tiempos y medirles la densidad óptica (DO) a una longitud de onda de 620 nm. Dichas DOs serán graficadas para observar cómo crece cierto cultivo bacteriano.

Inducción con pH y temperatura.

El pH y la temperatura son factores universales en el medio ambiente. Al cambiar modifican comportamientos en diversas comunidades, pudiendo ser la inducción de profagos parte de esos comportamientos modificados.

Para estudiar la capacidad inductiva en respuesta a una disminución del pH y un aumento de temperatura se utilizaron los organismos modelo CE3 y CE3f10. Aunque las mismas

también se aplicaron en organismos que probablemente contenían profagos intactos. Todo eso se detalla en los resultados.

En los siguientes puntos se especifica la metodología empleada para inducir profagos mediante una disminución del pH:

- I. Preparar 24 tubos con caldo fresco.
- II. A cada uno de 12 tubos con caldo fresco añadirle 100 μ l de cultivo líquido de la cepa *Rhizobium etli* CE3 crecida previamente toda la noche.
- III. A cada uno de 12 tubos con caldo fresco añadirle 100 μ l de cultivo líquido de la cepa *Rhizobium etli* CE3f10 crecida previamente toda la noche.
- IV. Incubar los 24 tubos por 40 min.
- V. Tomar 6 tubos con cultivos de CE3 y 6 con cultivos de CE3f10 y agregarle a cada uno (en condiciones de oscuridad) 3.6 μ l de HCl 37%. *Esto hará que los tubos disminuyan su pH de 6.8 a 5.*
- VI. Incubar todos los tubos (24 en total) en condiciones de oscuridad.
- VII. Retirar los tubos a diferentes tiempos (0, 1, 2, 3, 4 y 5 h.) y medir la DO de cada uno a 620 nm. *Por cada tiempo se retiran 4 tubos: 1 de CE3 y 1 de CE3f10 sin reducción de pH; y 1 de CE3 y 1 de CE3f10 con reducción de pH.*
- VIII. Cada tubo de los tratados con HCl al 37% (tubos con tratamiento) se centrifuga en tubos eppendorf de 1.5 mL a 12000 rpm por 10 min. a 4°C. *La centrifugación se realiza inmediatamente después de medir las DOs.*
- IX. Se recupera el sobrenadante de los tubos centrifugados y se almacena a 4°C.
- X. El sobrenadante recuperado se somete a la prueba de placa lítica para detección de fagos. *En este caso, el control negativo (sobrenadante sin el profago RHEph10) correspondería a los sobrenadantes de la cepa CE3, es inútil que se le haga la prueba a estos sobrenadantes. El control positivo (sobrenadante que probablemente contenga al fago 10) corresponde a los sobrenadantes de la cepa CE3f10.*

Con respecto a la temperatura, se buscó inducir a las lisógenas mediante un aumento de 7°C con respecto a la temperatura normal de crecimiento de las células del género *Rhizobium* (30°C). Dicho aumento es tolerable por las células de éste género.

La metodología es muy similar a la anterior y se explica en los siguientes puntos:

- I. Preparar 24 tubos con caldo fresco
- II. A cada uno de 12 tubos con caldo fresco añadirle 100 μ l de cultivo líquido de la cepa *Rhizobium etli* CE3 crecida previamente toda la noche.
- III. A cada uno de 12 tubos con caldo fresco añadirle 100 μ l de cultivo líquido de la cepa *Rhizobium etli* CE3f10 crecida previamente toda la noche.
- IV. Incubar los 24 tubos por 40 min.
- V. Tomar 6 tubos con cultivos de CE3 y 6 con cultivos de CE3f10 y pasarlos a una incubadora de 37°C. Los 12 tubos restantes continuarán su crecimiento a la

temperatura normal (30°C). Todos los tubos deben permanecer en condiciones de oscuridad.

- VI. Retirar los tubos a diferentes tiempos (0, 1, 2, 3, 4 y 5 h.) y medir la DO de cada uno a 620 nm. *Por cada tiempo se retiran 4 tubos: 1 de CE3 y 1 de CE3f10 sin aumento de temperatura y; 1 de CE3 y 1 de CE3f10 con aumento de temperatura.*
- VII. De los tubos con tratamiento (crecidos a 37°C), cada uno se pasa a tubos eppendorf de 1.5 mL y se centrifuga a 12000 rpm por 10 min. a 4°C. *La centrifugación se realiza inmediatamente después de medir las DOs.*
- VIII. Se recupera el sobrenadante de los tubos centrifugados y se almacena a 4°C.
- IX. El sobrenadante recuperado se somete a la prueba clásica para detección de fagos. *En este caso el control negativo (sobrenadante sin el profago RHEph10) correspondería a los sobrenadantes de la cepa CE3 y el positivo (sobrenadante que probablemente contenga al fago 10) a los sobrenadantes de la cepa CE3f10.*

Como en las inducciones con mitomicina C, para las pruebas con pH y temperatura también se realizaron curvas de crecimiento bacteriano con el objetivo de observar si existió inducción en cepas probablemente inducidas. Así mismo, se les aplicaron los métodos de detección de fagos por placa lítica.

Método de placa lítica para la detección de fagos

- I. Se prepara una caja Petri con agar PY para obtener un botón sólido de agar.
- II. Se prepara medio de agar suave y se mantiene a 44°C.
- III. Se toman 200 µl de un cultivo de *Rhizobium* susceptible a ser lisado por el fago en cuestión y se mezclan con 3 mL de agar suave a 44°C. Se agita suavemente para que las células se difundan por todo el medio sin producir burbujas, y se añade al botón de agar (caja petri con medio sólido).
- IV. Se esperan 15 minutos para que seque el agar suave.
- V. Sobre la placa de agar suave se gotean 10 µl de una solución que contenga fagos.
- VI. Se esperan 5 minutos a que se absorba la gota.
- VII. Se incuba toda la noche.
- VIII. Al día siguiente se observan placas líticas claras sobre el césped bacteriano, donde se goteó la solución de fagos.

Es importante recordar que solo se forma placa lítica si la cepa bacteriana crecida en agar suave es lisada por el fago inducido. Además, la ausencia de placa lítica no necesariamente exime la presencia de partículas virales, ya que pueden estar presentes pero no tener la capacidad de lisar a las bacterias del césped.

Conteo de unidades formadoras de placa (UFP) para fagos obtenidos por inducción con mitomicina C

- I. Preparar 36 cajas Petri con medio de cultivo sólido.
- II. Preparar 120 mL de medio suave. Es importante mantener este medio a 44°C.
- III. Incubar toda la noche: 1 tubo con la cepa *R. etli* CE3 lisogenizada con el fago RHEph10 (CE3f10) y 4 tubos con la cepa *R. etli* CE3 (CE3).

- IV. Retirar los tubos y almacenarlos a 4°C para posteriores usos.
- V. Agregar 100 µl/tubo de cultivo CE3f10 a cada uno de 4 tubos con caldo fresco.
- VI. Vortexear en incubador por 40 min.
- VII. Sacar los tubos de la incubadora y, en condiciones de oscuridad, agregar a cada uno mitomicina C para llegar a una concentración final de 0.5 µg/ml.
- VIII. Incubar los tubos en condiciones de oscuridad e irlos sacando a sus respectivos tiempos.
- IX. Cada tubo se sacará a diferentes tiempos (1, 2, 4 y 23 h.). Los siguientes pasos se aplicarán para cada tubo (4 en total) cada vez que se retire de la incubadora.
- X. Pasar el cultivo a tubos eppendorf de 1.5 mL para después centrifugarlo a 12000 rpm por 10 min y a 4°C.
- XI. Recuperar el sobrenadante (concentrado de fagos) en tubos eppendorf.
- XII. Del sobrenadante recuperado, hacer las siguientes diluciones: 10⁻⁰, 10⁻¹, 10⁻², 10⁻³, 10⁻⁴, 10⁻⁵, 10⁻⁶, 10⁻⁷, 10⁻⁸. Es suficiente obtener 1 mL por cada dilución.
- XIII. Por cada dilución se hace la siguiente mezcla: 100 µl de la dilución mas 200 µl de cultivo toda la noche de la cepa CE3 mas 3 mL de agar suave a 44°C.
- XIV. Se agita la mezcla suavemente para que las células se difundan por todo el medio sin producir burbujas y posteriormente se añade al botón de agar (caja petri con medio solido).
- XV. Esperar 15 minutos para que seque el agar suave.
- XVI. Incubar toda la noche.

Al día siguiente se observan cajas Petri con gran cantidad de placas líticas, cada una representa una UFP, es decir, un virión único que se replicó en fase lítica.

PROFAGOS

Predicción de profagos con PHAST

Para la búsqueda de profagos en cepas secuenciadas se utilizó el programa PHAST de la siguiente manera.

- I. Abrir la página de internet “phast.wishartlab.com”
- II. En la página principal esta la opción para subir la secuencia, ya sea mediante un archivo .fasta (directamente de la computadora del usuario), un numero de acceso del Genbank o directamente copiando y pegando la secuencia nucleotídica.
- III. Posteriormente se oprime el botón “submit” y automáticamente el programa comenzará a operar, tardando de 1-5 minutos para arrojar los resultados
- IV. El tipo de resultados que arroja el programa, así como su organización y formato se detallan en el apartado de resultados de este trabajo. Donde se pone el ejemplo de los resultados de un genoma de *Rhizobium*, el cual cuenta con los 3 tipos de profagos determinados por PHAST (explicados también en resultados).

Clasificación de profagos en grupos por identidades nucleotídicas y diseño de marcadores moleculares

Para agrupar los profagos encontrados con PHAST, junto con algunas secuencias de fagos del laboratorio, se siguió la siguiente metodología.

- I. Se realizó un alineamiento múltiple con 83 secuencias genómicas, 11 fagos de la colección del laboratorio y 72 profagos encontrados con el programa PHAST (los ubicados en este trabajo). El alineamiento nucleotídico se realizó con el programa MAUVE (MAUVE , on line), obteniéndose una matriz con los bloques que se aparean, donde también se señala el porcentaje de identidad nucleotídica y tamaño (pbs) de cada bloque.
- II. Basándose en dicha matriz, y con ayuda del programa MAUVE, se obtuvo un árbol guía que reflejó de una manera más amable los mismos resultados.
- III. Con base en el árbol del punto anterior se eligieron los fagos y profagos más cercanos para generar bloques. Cuyas secuencias genómicas fueron alineadas con un script en python (Phyton, on line), generándose secuencias comunes entre los genomas de cada bloque.
- IV. De cada bloque (14 en total) salió una secuencia en común, que representó a los grupos de fagos y profagos anteriormente mencionados. *Es importante recordar que en cada grupo existen fagos y/o profagos provenientes de distintas cepas bacterianas, algunas muy distantes entre sí en lo que respecta a su origen geográfico.*
- V. De los 14 grupos se eligieron los que contenían fagos y profagos más heterogéneos en cuanto a la geografía de las cepas bacterianas que les dan origen, eligiéndose 7.
- VI. Teniendo las secuencias de los 7 grupos se procedió al análisis (mediante visualización por colores con el programa MEGA (MEGA, on line)) de las secuencias de cada grupo, eligiéndose las zonas casi idénticas de los fagos y profagos de cada grupo, cuyo tamaño osciló entre 500 y 1000 pbs.

Después de obtener los 7 secuencias (marcadores moleculares de ADN) de los 7 grupos principales se procedió al diseño de oligos de cada uno, para ello se siguió la siguiente metodología.

- I. Utilizar el software OLIGO 7.0 (OLIGO 7.0, on line).
- II. Ya elegido el marcador de 500-1000 pbs se procede a elegir un segmento de 20 pbs aproximadamente para cada extremo de cada segmento.
- III. Las secuencias de los extremos de ida y regreso (forward y reward) deberán ir en dirección 5'-3', de tal forma que al hibridarse con su contrario (en el PCR) sean capaces de cebar el crecimiento de la nueva cadena.
- IV. Cada segmento de 20 pbs aprox. se introduce al programa OLIGO 7.0 y este arroja una "Tm".
- V. Las Tms deben ser muy cercanas entre cada par de oligos, ya que ambos serán corridos en el mismo PCR para amplificar la misma secuencia. Es decir, cada

marcador deberá contar con un par de oligos (los que flanqueen la secuencia) con Tms similares.

Ya seleccionados los pares de oligos, se mandan a sintetizar en la unidad de biosíntesis del IBT (UNAM) para ser corridos en las condiciones adecuadas.

Condiciones de PCR para el marcador del primer grupo

Como se explicará en los resultados, de los 7 marcadores moleculares diseñados solo se ha probado el correspondiente al grupo 1, el cual representa 23 profagos, la mayoría de Tepoztlán. Este marcador (de 880 pbs aproximadamente) fue probado en cepas no secuenciadas de *Rhizobium* (ver la última parte de los resultados) para determinar cuáles de ellas podrían tener profagos. Las condiciones para hacer los PCRs en los DNAs totales de cepas de *Rhizobium* no secuenciadas fueron las siguientes.

- I. Mezclar lo siguiente (en μl):
 - 16 de H_2O
 - 2.5 de buffer 10X
 - 1.5 de DNTP'S al 0.4 mM
 - 1.75 de MgCl_2 al 50 mM
 - 1 de oligo forward
 - 1 de oligo reverse
 - 0.25 de Taq DNA Polimerase (Altaenzymes)
 - 1 de DNA muestra (aislado de la bacteria) a 15 ng/ml
- II. La suma de todo (25 μl) se mete al termociclador, que se programa para 35 ciclos de la siguiente forma:
 - 94°C por 4 minutos
 - 94°C por 1 minuto
 - 66°C por 1 minuto
 - 72°C por 1.5 minutos
 - 72°C por 5 minutos
 - 4°C en almacenaje hasta que sea retirado
- III. Posterior a la PCR se realiza una electroforesis con las especificaciones señaladas en el apartado "Electroforesis" de esta metodología.
- IV. Las muestras que salgan positivas serán candidatas a contener profagos similares a los ubicados en el grupo 1. Las cepas, fotos, etc., se detallan en el apartado correspondiente de resultados y discusión.

RESULTADOS Y DISCUSIÓN

Este trabajo consistió tanto de un enfoque experimental como de uno bioinformático. De acuerdo a los antecedentes se asumió que los bacteriófagos de *Rhizobium* pueden encontrarse como profagos y, por lo tanto, pueden inducirse bajo los efectos de agentes externos causantes de estrés celular. Por ejemplo, la mitomicina C se ha usado ampliamente para inducir profagos en varias especies bacterianas, provocándoles daño en el ADN y encendiendo su sistema de reparación, lo que activa la fase lítica de algunos profagos. Bajo estas premisas, la inducción fue probada en diferentes cepas lisogénicas de *Rhizobium*.

Con respecto al enfoque bioinformático, aprovechamos la disponibilidad de secuencias genómicas completas de un conjunto de cepas de *Rhizobium* para inferir la presencia de profagos mediante programas bioinformáticos. Dichos profagos se agruparon de acuerdo a sus identidades nucleotídicas y dichos grupos se utilizaron para construir marcadores moleculares (pequeñas secuencias de ADN), de los cuales, uno se utilizó para rastrear profagos en cepas de *Rhizobium* no secuenciadas.

INDUCCIÓN DEL PROFAGO RHEph10 A PARTIR DE LA CEPA *Rhizobium etli* CE3 LISOGENIZADA CON EL FAGO RHEph10

En trabajos previos, el laboratorio obtuvo evidencia de lisogenicidad del fago 10 (RHEph10) en varias cepas de *R. etli*. El fago RHEph10 se aisló de la cepa *R. etli* CE3t4, que se obtuvo de un campo donde usualmente se cultiva frijol, en Ahuatenco, Estado de México. El genoma de este fago se reportó y depositó en el Genbank (RHEph10 ó JX483881). Entre sus genes se encuentran algunos de origen bacteriano como los *repAB* y *repC* (cuyos productos proteicos se requieren para la replicación y partición de plásmidos de *Rhizobium*), los genes *cob* para la síntesis de cobalamina, genes clásicos de bacteriófagos (lisozima, proteínas de la cápside, terminasas, represores, etc.) y una gran cantidad de genes de función desconocida (genes hipotéticos) (Santamaría *et al.*, 2014).

La presencia de posibles reguladores transcripcionales (represores) y de genes de procedencia bacteriana originó la idea de que este fago podría ser atemperado (lisogénico) y capaz de realizar transducción. Experimentos posteriores lograron la lisogenización de la cepa de *R. etli* CFN42 (CE3) con el fago RHEph10, que en lo sucesivo será abreviada como CE3f10. Además, análisis ulteriores demostraron que este profago, a diferencia de otros, se establece como una molécula replicativa autónoma (plásmido) en sus hospederos. Figura R1 (Santamaría *et al.*, no publicado).

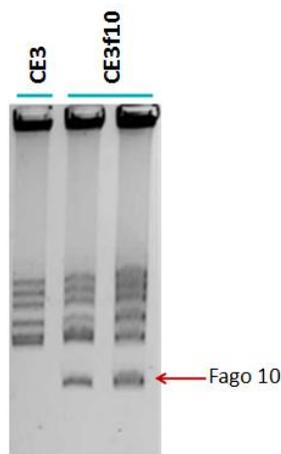


Figura R1. Electroforesis de DNA bacteriano tratado la técnica de Eckhardt, utilizada para visualizar los plásmidos de las cepas CE3 y CE3f10. En las muestras provenientes de la cepa CE3f10, en la parte de abajo se observa el plásmido correspondiente al fago 10. *Modificado de Santamaría et. al., no publicado.*

Ya que la presencia del profago del fago 10 se comprobó varias veces en la cepa CE3f10 se decidió tomar dicha cepa para estandarizar los experimentos de inducción de este trabajo. El objetivo fue estandarizar condiciones de inducción óptimas en cepas de *Rhizobium* mediante la búsqueda de los siguientes factores: inductor indicado (físico o químico); concentraciones o características específicas del mismo; tiempo con el inductor que necesita la cepa bacteriana para que este actúe eficazmente; etcétera.

En la literatura del área se reportan diversos agentes químicos y condiciones de crecimiento particulares para la inducción de profagos en distintas especies de bacterianas. Como ejemplos se mencionan los siguientes: 3 mM de H₂O₂, 1 µg/ml de mitomicina C, 0.2 µg/ml de norflaxacina y 200 mM de NaCl (Los *et al.*, 2009); el compuesto antimalaria fansidar (30 mg/placa), daraprim (20 mg/placa), nivaquine (11 mg/placa) y camoquine (11 mg/placa) (Fasunon *et al.*, 1989); pH 5, 0.1 mM de Cr (IV), 0.15 µM de KCN y un aumento de temperatura de 7°C comparado con la temperatura optima a la que crece cierta cepa bacteriana (Choi *et al.*, 2010); 1.5 µM de ácido nalidixico, 1.49 µM de ácido oxolinico, 2.69 µM de adriamien, 126.86 µM de elipticina, 8.88 µM de H₂O₂, 33.51 µM de paraquat y 6353.4 µM m-AMSA (DeMarini 1992).

Con respecto a la mitomicina C, que es una de las principales drogas usadas en la inducción de bacterias lisógenas, Werkin *et al.*, en 1988 utilizaron concentraciones variables (0.5, 1.0, 5.0 y 10 µg/ml) para encontrar la cantidad que indujera más eficazmente a 33 profagos de diferentes cepas de la especie *Rhizobium meliloti*. En otro trabajo, diez años después, Faruque *et al.*, en 1998 indujeron a la lisógena *Vibrio cholerae* toxigénica (que contiene al profago que produce la toxina del cólera) utilizando concentraciones de 20 ng/ml de esta droga. Ambos trabajos tuvieron éxito con respecto a la inducción, que fue comprobada mediante el método clásico de placa lítica, además, de ellos se deduce que las concentraciones necesarias de Mitomicina C para la inducción de lisógenas son variables de acuerdo con la especie bacteriana tratada.

Para la evaluación de la inducción del profago 10 en la cepa CEf10 se realizaron curvas de crecimiento con concentraciones de mitomicina C a 0.5, 2.5 y 5.0 $\mu\text{g/ml}$. En la figura R2, se muestra las curvas de crecimiento (DO_{620}) de las cepas CE3 y CE3f10 después de agregarles mitomicina C a dichas concentraciones. Se observa que la cepa CE3f10 disminuye su crecimiento después de 240 minutos de ser inducida, fenómeno que se aprecia claramente en las cepas lisogénicas inducidas a 2.5 y 5 $\mu\text{g/ml}$. El efecto más evidente lo observamos a la concentración de 5 $\mu\text{g/ml}$.

Las cepas inducidas con 0.5 $\mu\text{g/ml}$ muestran una diferencia prácticamente imperceptible, que puede apreciarse mejor en la tabla R1. Allí se observa que la cepa CE3f10 (con respecto a la CE3) tiene mayores densidades bacterianas hasta los 240 minutos, las cuales disminuyen ligeramente a los 300 minutos. Este comportamiento podría deberse a que la baja concentración del inductor no tuvo la capacidad de penetrar en muchas células, como quizá si sucedió en los cultivos inducidos con mayores concentraciones de droga. Podría decirse que a mayor concentración de mitomicina, mayor cantidad de células inducidas, lo que se reflejaría en una disminución de densidad celular. Esta sugerencia corresponde con las inducciones a 2.5 y 5 $\mu\text{g/ml}$ de droga, donde se observa claramente que, a mayor cantidad de droga, menor crecimiento de la cepa lisogénica. Además, hay que tomar en cuenta que en un cultivo bacteriano las células inducidas son opacadas por las células no inducidas, es decir, las células que no son afectadas por el inductor (ya sea porque este no penetra o porque son resistentes naturales) tienen la capacidad de reproducirse en fase logarítmica, de tal forma que si la inducción se dio solo en una porción pequeña de la población podría ser imperceptible mediante mediciones de densidad óptica. Por otro lado, es posible que la mitomicina también sea capaz de matar células *per se*, por lo que una menor concentración tendría la capacidad de matar un menor número de células.

Tomando en cuenta las explicaciones expuestas sería lógico pensar que a menor concentración de mitomicina C hubo una menor concentración de las células inducidas, por lo que el crecimiento bacteriano podría parecer normal, es decir, podría darse la impresión de que la inducción fue ineficiente, cosa que sería completamente falsa. Sin embargo, hay que recordar que comprobar la inducción mediante curvas de crecimiento es solo un ensayo prospectivo, y que los ensayos más determinantes son el ensayo en placa, la microscopía electrónica, la PCR, entre otros.

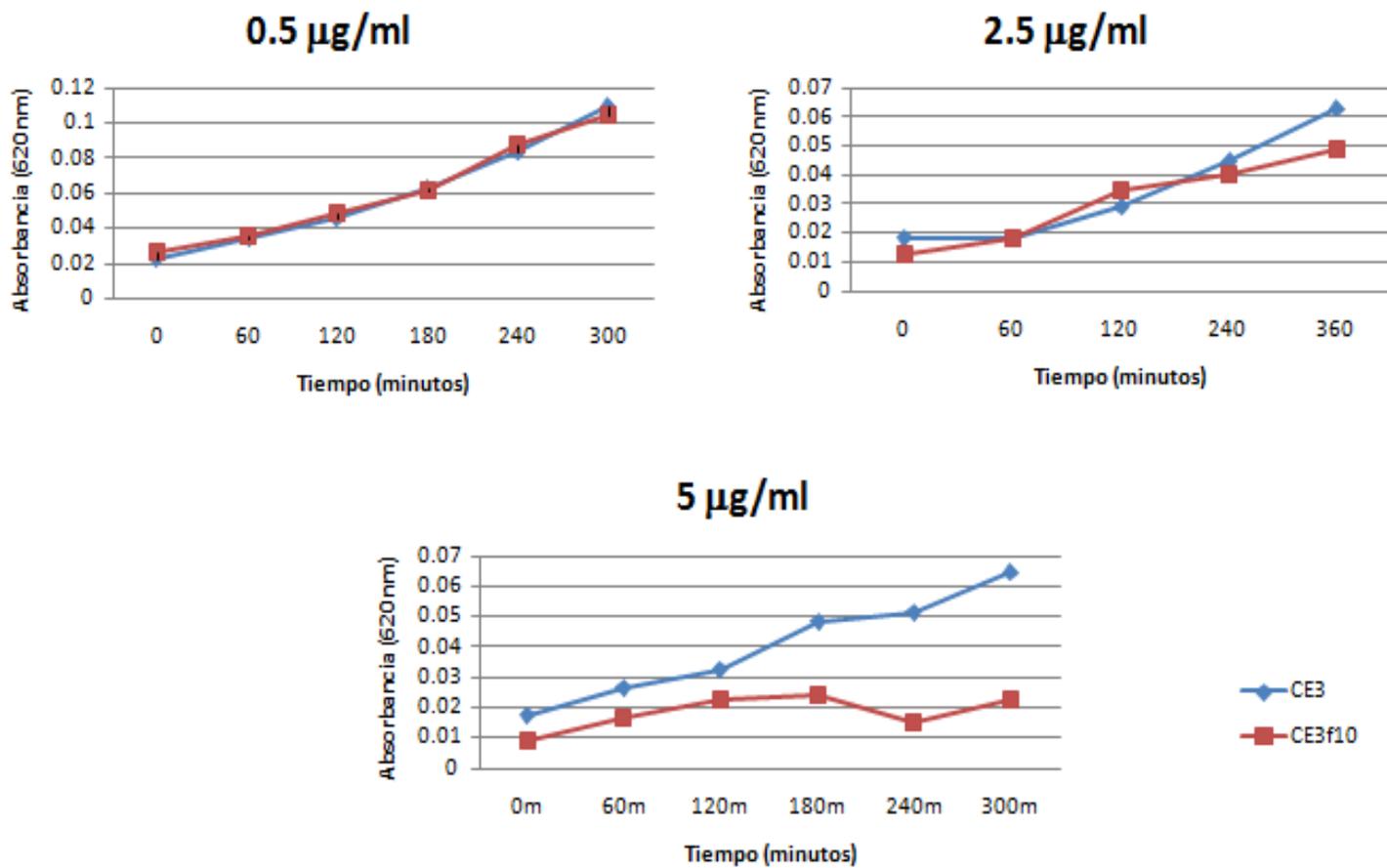


Figura R2. Curvas de crecimiento bacteriano después de agregar mitomicina C a diferentes concentraciones. A mayor concentración de mitomicina C hay menor crecimiento de la cepa CE3f10.

Tabla R1. Lecturas de densidades ópticas de las cepas CE3 lisogénicas y no lisogénicas inducidas con mitomicina C a una concentración de 0.5 µg/ml.

Inducción con Mit C. 0.5 µg/ml		
Tiempo (minutos)	ODs(620nm)	
	CE3	CE3f10
0	0.022	0.027
60	0.035	0.050
120	0.045	0.065
180	0.063	0.081
240	0.084	0.087
300	0.109	0.105

En esencia, el experimento demuestra que al crecimiento de la cepa CE3f10 lo afecta la presencia de mitomicina C. Dado que este antibiótico produce daño celular y modifica la viabilidad de las cepas, pudiera pensarse que la disminución del crecimiento de la cepa CE3f10 se debe a este efecto (como se menciona arriba) y no a la inducción del profago, que produciría partículas virales y lisis bacteriana. Sin embargo, si el efecto de la mitomicina únicamente se debiera a esto se observarían las mismas lecturas para las 2 cepas tratadas, pero el crecimiento bacteriano es claramente mayor en la cepa CE3, lo que sugiere una inducción en la cepa CE3f10, que muestra una disminución en su crecimiento.

Para corroborar la liberación del fago 10 se aplicó el método de extracción de fagos para las cepas tratadas con mitomicina C y posteriormente se realizó el método de detección de fagos por formación de placa lítica (ver metodología). En la figura R3 (goteo superior izquierdo de la caja de la izquierda) se observa una placa lítica sobre la cepa CE3, cuyo goteo fue hecho a partir de una muestra de CE3f10 después de 4h de ser inducida con mitomicina C a 0.5 µg/ml. En ese mismo césped se pueden apreciar 3 goteos adicionales (correspondientes a extractos de fagos de 3 cultivos de la cepa CE3 inducidos con 0.5, 2.5 y 5 µg/ml de mitomicina C) que no formaron placa lítica, indicando que lo que provoca la lisis en las células de la cepa CE3 proviene del extracto de fagos obtenido a partir de la cepa CE3f10 inducida.

En la caja de la derecha, que tiene un césped con la cepa CE3f10 se observa el mismo patrón de goteos que en la otra caja. En ellos no se ve la formación de placas líticas, lo que demuestra 2 cosas: 1) los goteos correspondientes a las extracciones de fagos de la cepa CE3 afectan de igual manera a las cepas CE3 y CE3f10, indicando que ninguno de ellos contiene al fago 10; 2) la cepa CE3f10 (césped derecho) fue inmune al extracto fágico obtenido a partir de la cepa CE3f10 (goteo superior izquierdo), por lo que se puede suponer que la cepa CE3f10 es inmune al fago 10, hecho que es muy común en cepas lisogénicas, como por ejemplo en la cepa de *Escherichia coli* lisogenizada con el fago λ, cuya inmunidad hacia el mismo fago se conoce ampliamente (Heinemann y Spiegelman, 1970).

Con respecto a las marcas “fantasma” de los sitios donde se realizaron los goteos, que no corresponden a placas líticas, lo más probable es que hayan sido provocadas por la acción toxica que tiene la mitomicina C hacia las bacterias (Tomasz, 1976), que en el caso de las cepas CE3 y CE3f10 se sugiere baja pero clara. Para corroborarlo se podrían hacer algunos experimentos, sin embargo sería algo sin sentido ya que lo que interesa es la inducción de profagos, y los experimentos anteriores ya la han comprobado.

Es importante señalar que las fotos de la figura R3 solo comprueban la inducción exitosa de extractos de cultivos inducidos con mitomicina C a 0.5 µg/ml. Sin embargo el mismo experimento se realizó para cultivos inducidos con concentraciones de mitomicina C de 2.5 y 5 µg/ml, en cuyos casos se comprobó la inducción de la misma manera. Se tomó como ejemplo la inducción con la menor concentración de droga para demostrar que, aunque su inducción no fue probada mediante la disminución del crecimiento de la cepa CE3f10 (figura R2), se verificó claramente en el ensayo de placa lítica (figura R3).

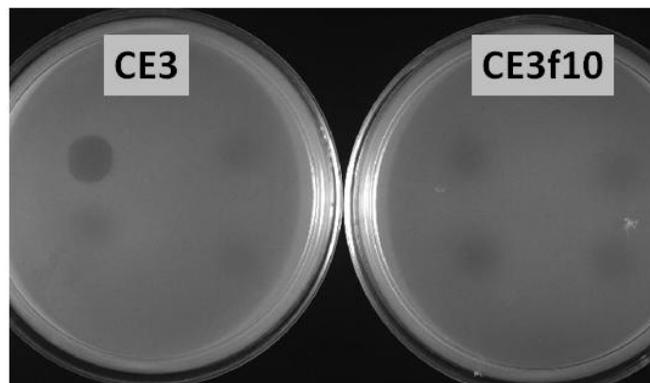


Figura R3. Ensayo de placa para comprobar la presencia de fagos. Cada cama (CE3 y CE3f10) fue goteada con 4 extractos: arriba a la izquierda (ext. de CE3f10 inducida con mitomicina C a 0.5 µg/ml); arriba derecha, abajo izquierda y abajo derecha (ext. de CE3 “inducidos” con mitomicina C a concentraciones de 0.5, 2.5 y 5 µg/ml respectivamente). En la cama de la cepa CE3 observamos una placa visible en el sitio donde se goteó el extracto de la cepa CE3f10 inducida con mitomicina C a 0.5 µg/ml. En la cama de la cepa CE3f10 no se observa ninguna placa.

Después de comprobar la inducción y presencia de fagos mediante las dos técnicas antes mencionadas se procedió a determinar el tiempo al que los fagos se liberan en mayor cantidad después de agregar el inductor. Para ello se aplicó el método de cuantificación de fagos liberados. Se obtuvieron muestras del cultivo CE3f10 (inducido con mit. C a 0.5 µg/ml) a las 1, 2, 4 y 23 horas, y se extrajeron los fagos por el método de “extracción de fagos” arriba descrito. Con cada una de las extracciones se hicieron series de diluciones desde 10^0 hasta 10^{-8} y cada una se añadió como césped sobre una capa de la cepa CE3 (ver método de cuantificación de fagos liberados), después se cultivaron toda la noche para la obtención de placas líticas, las cuales se contaron para estimar el número de unidades formadoras de placas (UFPs), en la figura R4 podemos ver el ejemplo de una caja. En la tabla R2 se observan los resultados de cada dilución plaqueada en cada tiempo, estimando las UFPs por cada 100 µl de cada dilución agregada a cada caja. Es importante señalar que cada UFP corresponde a un fago que se multiplicó, luego liso a su hospedero y después se dispersó a otras células.

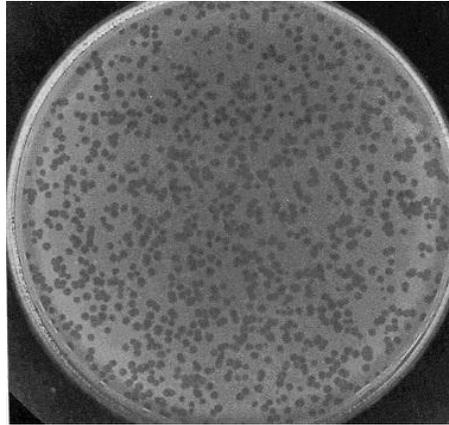


Figura R4. Plaqueo con una capa de 100 μ l de extracto de fagos de CE3f10 4 horas después de ser inducida con mitomicina C a 0.5 μ g/ml , diluida 1000 veces y probada sobre la cepa CE3.

Tabla R2. Unidades formadoras de placa (UFPs) de 100 μ l de extractos de fagos a partir de CE3f10 inducido con mitomicina C a 0.5 μ g/ml. En las filas observamos el tiempo en el cual se realizó la extracción y en las columnas de la dilución de las mismas.

Unidades formadoras de placa de extractos de fago de CE3f10 inducida con Mit.C 0,5 μ g/ml									
Tiempos de extracción de fagos después de agregar el inductor	Diluciones								
	10 ⁰	10 ⁻¹	10 ⁻²	10 ⁻³	10 ⁻⁴	10 ⁻⁵	10 ⁻⁶	10 ⁻⁷	10 ⁻⁸
1h.	Incontables	775	26	3	1	0	0	0	0
2h.	Incontables	900	216	17	0	0	0	0	0
4h.	Incontables	Incontables	Incontables	1160	168	28	2	0	0
23h.	Incontables	Incontables	2000	404	24	2	0	0	0

Con la intención de darle uniformidad y una mayor coherencia a los datos se realizaron operaciones para obtener la cantidad de fagos por mililitro puro (sin diluir) que se obtienen al extraerlos después de 1, 2, 4 y 23 horas posteriores a la adición del inductor (tabla R3 y figura R5). Se observa que hay un crecimiento gradual de UFPs hasta las 4 horas después de que la cepa CE3f10 es expuesta a la mitomicina C. La mayor cantidad de UFPs la vemos a las 4 horas; sin embargo, a las 23 horas el título UFPs es marcadamente menor, lo que podría deberse principalmente a dos cosas: 1) los fagos (contenidos en las diluciones) lisogenizan a las cepas CE3, convirtiéndolas inmunes al fago, por lo que la cepa (ahora CE3f10) continua su crecimiento logarítmico normal; 2) probablemente sobrevivan células resistentes naturales al fago, que también se

multiplican logarítmicamente, pero al inicio no se notan porque son muy pocas en comparación con toda la población.

Tabla R3. UFP/ml de extractos puros de fagos recuperados a diferentes tiempos después de inducir la cepa CE3f10 con mitomicina C a 0.5 µg/ml.

Tiempo de extracción	UFP/ml de extracto de fagos
1h.	5.8×10^4
2h.	1.9×10^5
4h.	2×10^7
23h.	2.5×10^6

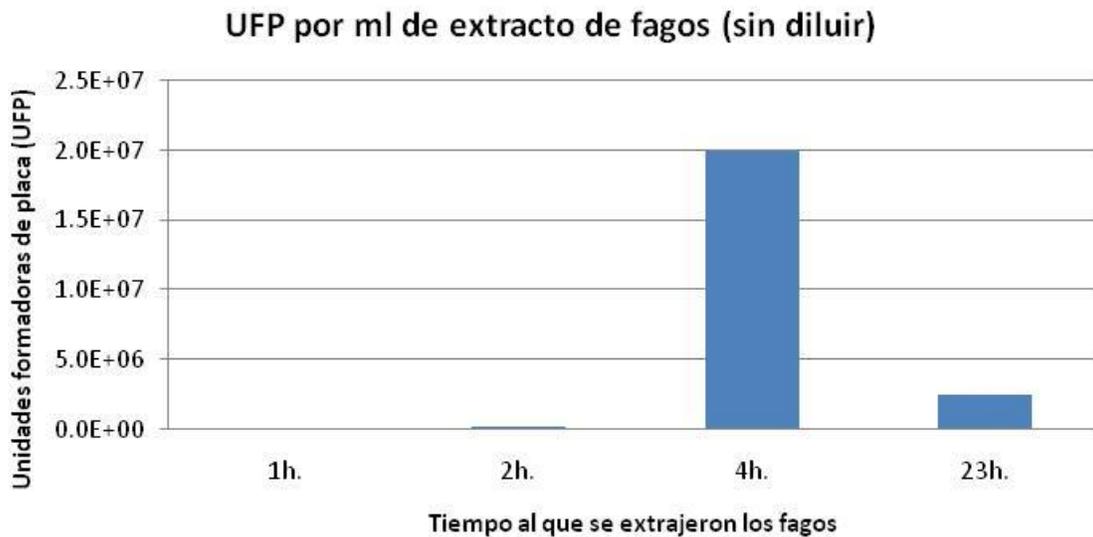


Figura R5. UFPs/ml de extractos fágicos puros extraídos a diferentes tiempos a partir de CE3f10 inducida con mitomicina C a 0.5 µg/ml. Nótese que la mayor cantidad de fagos liberados se presenta 4 horas después de agregar el inductor.

Como se puede apreciar, tanto los resultados de curvas de crecimiento, placas líticas y conteos de UFPs indican que después de inducir la cepa CE3f10 se libera una sustancia o partícula que muy probablemente sea el fago 10, que tiene la capacidad de lisar células de *R. etli* CE3. Es importante apreciar que estos resultados solo sucedieron después de agregar mitomicina C a la cepa CE3f10, ya que los mismos experimentos se realizaron sobre la cepa CE3 (que no contiene al profago 10) y los resultados fueron negativos para la lisis de células CE3. Por lo que se deduce que la partícula responsable de lisar a la cepa CE3 es el fago 10, que se liberó después de inducir la cepa CE3f10 con mitomicina C.

INDUCCIÓN DE PROFAGOS EN CEPAS DE *Rhizobium*.

En el laboratorio existe una colección de aproximadamente 120 cepas (Anexo) pertenecientes al género *Rhizobium*. El genoma de algunas de ellas está completamente secuenciado y el de otras están en proceso de serlo.

Después de comprobar la eficiencia de la mitomicina C como inductor de lisógenas en *Rhizobium*, así como de encontrar las mejores concentraciones y tiempos de exposición del inductor para lograr la mayor eficiencia en el proceso, se tomó la decisión de inducir un conjunto de cepas de *Rhizobium* con la hipótesis de que si estas contenían fagos podríamos obtener partículas virales infectivas (viriones) en el sobrenadante de cultivos inducidos con mitomicina C, además de otras sustancias químicas y condiciones físicas.

Las cepas que se eligieron para probar la inducción se dividieron en dos paquetes debido a que se les ubicaron profagos en diferentes momentos del desarrollo del proyecto. En el primer paquete se analizaron 22 genomas y se encontraron 2 “profagos intactos” pertenecientes a las cepas *R. etli* R630 y *R. etli* NXC14. En el segundo paquete 50 genomas fueron analizados y se encontraron 17 profagos probablemente funcionales en las cepas de *R. etli* 8C3, Bra5, NXC14, TAL182, N122, N324, N671, N771, NXC24, R339, R630, R693 y R711 (ver siguiente apartado).

Las cepas con profagos del primer grupo (R630 y NXC14) junto con los controles CE3f10 (positivo) y CE3 (negativo) fueron sometidas a una serie de experimentos de inducción donde no solamente se probó la mitomicina C a diferentes concentraciones (0.5, 2.5 y 5 µg/ml) sino también tetraciclina (a 0.5, 1 y 10 µg/ml), cloranfenicol (a 9.5, 19 y 25 µg/ml), estreptomina (a 6.5, 12.5 y 25 µg/ml), disminución de pH (de 7 a 5) y aumento de temperatura (de 30 a 37°C). Se realizaron curvas de crecimiento bacteriano sobre las cepas inducidas con mitomicina C, pH y 37°C (figura R6).

De ello se observa que: la cepa CE3 (control positivo) tiene un patrón de crecimiento similar con la mayoría de los inductores, desarrollándose un poco más rápido que las otras cepas, pero con respecto a la reducción de pH su crecimiento es menor, lo que puede deberse a que dicha cepa es muy sensible a los cambios de concentración de iones hidronio; la R630 crece uniforme en todas las condiciones excepto en la de pH donde crece un poco más que las otras cepas, sugiriéndole una mayor resistencia a los medios ácidos; la NXC14 crece similar en todas las condiciones y; la cepa CE3f10 tiene una disminución en el crecimiento cuando se agrega mitomicina C a 2.5 y 5 µg/ml, cuestión que ya se discutió en párrafos pasados.

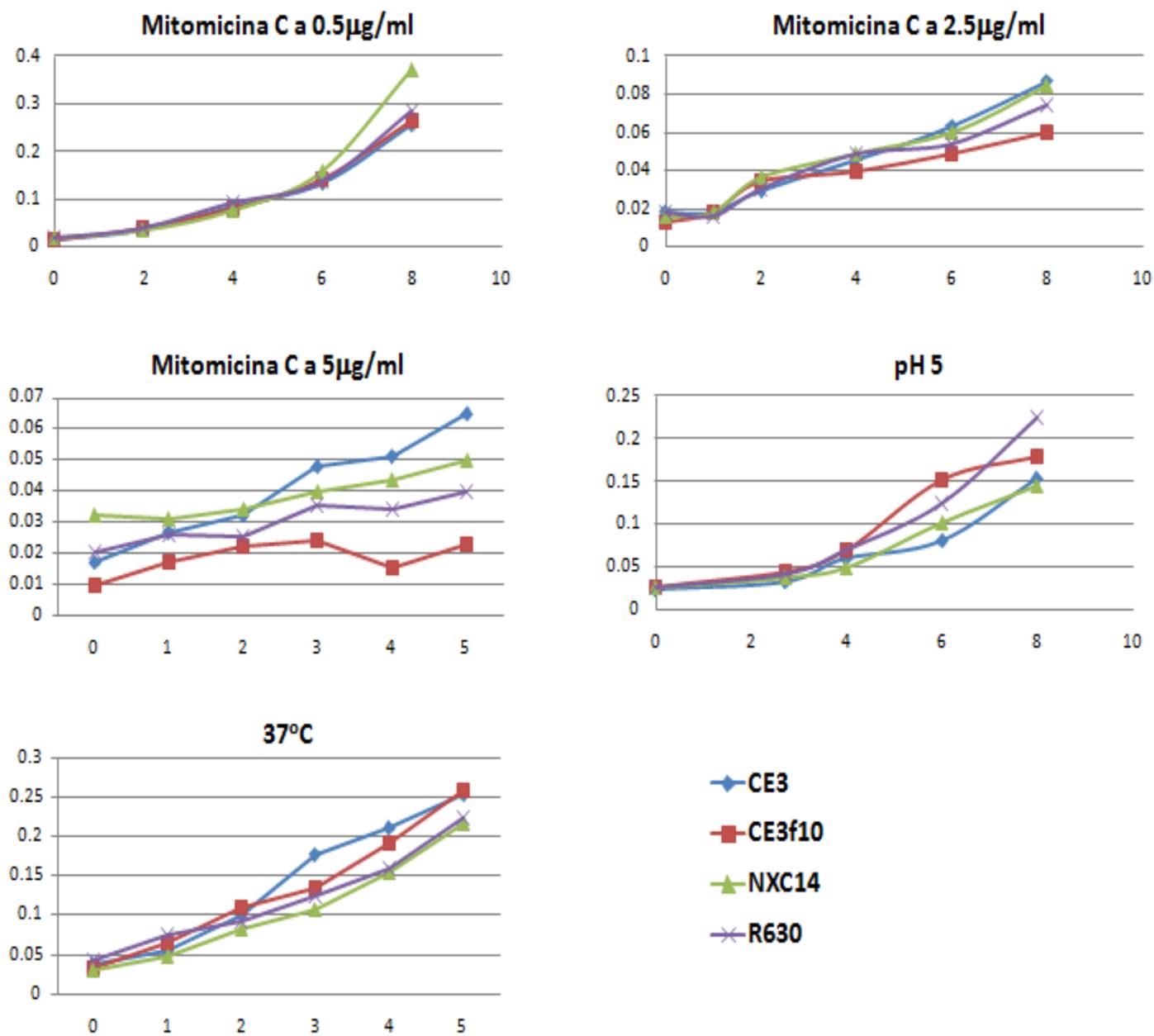


Figura R6. Curvas de crecimiento bacteriano de 4 cepas de *Rhizobium* que fueron inducidas con diferentes tratamientos (título de cada grafico). 3 de ellas contienen profagos, según el programa PHAST, y una corresponde al control negativo (CE3). El eje de las "Y" corresponde a la DO₆₂₀ y el eje de las "x" corresponde al tiempo (en horas) en que se leyeron las muestras.

Los resultados en el crecimiento bacteriano no sugieren ningún efecto de los inductores (salvo de la mit. C) sobre alguna(s) de las cepas citadas. Aun así, ya que las curvas son solo presuntivas, se tomó la decisión de hacer extractos de fagos (4 horas después de agregar el inductor) a partir de las cepas inducidas con pH5, 37°C y mitomicina C a 2.5 y 5 µg/ml, obteniendo un total de 16 extractos (4 inductores para c/u de 4 cepas). Cada uno de los 16 extractos fue probado (mediante la técnica de placa para la detección de fagos) sobre cada una de 118 cepas de *Rhizobium* de la colección del laboratorio, por lo que en total se realizaron 1888 pruebas.

En la tabla R4 se muestra que únicamente 2 extractos (C3f10 inducido con las 2 concentraciones de mitomicina C) produjeron placas líticas. Ya se sabía que los fagos extraídos de la cepa CE3f10 producen placa sobre CE3, lo que corresponde a un control positivo, pero fue nuevo descubrir que también lo hacen sobre MIM 1, CIAT652, 14C-1 y 14C-2. En todos los casos restantes (de las 1888 pruebas) no se observó la formación de placas líticas, por lo que se puede concluir que ninguna de las inducciones (salvo la del control) fue exitosa.

Tabla R4. Detección de fagos de 16 extractos fágicos mediante la técnica de formación de placa lítica. Cada extracto fue probado en 118 cepas de *Rhizobium*.

ENSAYOS DE PLACA PARA LA PRESENCIA DE FAGOS				
Inductor	Extracto fágico de la cepa inducida (FP=forma placa lítica sobre una o mas cepas y NFP=no forma placa)			
	CE3	CE3f10	NXC14	R630
Mit. C a 2.5µg/ml	NFP	FP (CE3, MIM1, CIAT652, 14C-1 y 14C-2)	NFP	NFP
Mit. C a 5µg/ml	NFP	FP (CE3, MIM 1, CIAT652, 14C-1 y 14C-2)	NFP	NFP
pH 5	NFP	NFP	NFP	NFP
37°C	NFP	NFP	NFP	NFP

Como se mencionó al principio de este apartado, los experimentos realizados para la inducción y detección de fagos se hicieron en 2 partes. Los resultados anteriormente descritos (figura R5 y tabla R3) corresponden a la primera parte. En la segunda parte, basados en posteriores búsquedas con PHAST, se indujeron las cepas *R. etli* 8C3, Bra5, NXC14, TAL182, N122, N324, N671, N771, NXC24, R339, R630, R693 y R711, además de los controles CE3 y CE3f10. En esta segunda parte se utilizaron los inductores de aumento de temperatura (de 30 a 37°C), aumento de acidez (de pH 7 a 5) y adición de mitomicina C (2.5 µg/ml). Es decir, se realizaron 3 tipos diferentes de inducción para 15 diferentes cepas de *Rhizobium*, siendo 45 inducciones en total.

En cada caso, 4 horas después de agregar el inductor se obtuvo el extracto de fagos y se probó sobre 120 cepas de *Rhizobium* (Anexo), por lo que en total se realizaron 5400 pruebas. Nuevamente todos los resultados fueron negativos ya que no se formó placa lítica en ninguno de los extractos probados, a excepción del control positivo, donde las cepas sensibles resultaron ser las mismas que las del experimento anterior. (Tabla R5).

Tabla R5. Experimentos de detección de fagos por el método de formación de placa. Cada celda corresponde a un extracto que fue probado sobre 120 cepas de *Rhizobium*. Únicamente el extracto de la CE3f10 inducida con mitomicina C formó placa sobre las camas de cultivo de 5 cepas diferentes.

ENSAYOS DE PLACA PARA LA PRESENCIA DE FAGOS			
Para cada extracto fágico: FP=forma placa lítica sobre una o mas cepas y NFP=no forma placa lítica			
CEPA	INDUCTOR		
	M. C 2.5 mg/ml	pH 5	37oC
CE3	NFP	NFP	NFP
CE3f10	FP (CE3, MIM1, CIAT652, 14C-1 y 14C-2)	NFP	NFP
8CE3	NFP	NFP	NFP
Bra5	NFP	NFP	NFP
NXC14	NFP	NFP	NFP
TAL182	NFP	NFP	NFP
N122	NFP	NFP	NFP
N324	NFP	NFP	NFP
N671	NFP	NFP	NFP
N771	NFP	NFP	NFP
NXC24	NFP	NFP	NFP
R339	NFP	NFP	NFP
R630	NFP	NFP	NFP
R693	NFP	NFP	NFP
R711	NFP	NFP	NFP

Los resultados de las inducciones mostraron claramente que únicamente el profago 10 de la cepa CE3f10 se induce con mitomicina C. Así mismo se demostró que todas las cepas con profagos detectados mediante el programa PHAST no se indujeron con ningún tratamiento. Lo anterior puede deberse a que los inductores no fueron los adecuados, o que la forma de inducción tuvo alguna falla, o inclusive a que los profagos no se encuentran en un estado completo (y por lo tanto funcional). Varios trabajos (Touchon *et al.*, 2016; Canchaya *et al.*, 2003; etc.) han probado la existencia de profagos incompletos, ya sea por falta genes o parte de los mismos, así como por la ausencia de elementos importantes de control, lo que se traduce en una formación y/o liberación incorrecta de partículas virales.

Otra explicación se centra en el comportamiento intrínseco de los fagos atemperados, el cual dista mucho de ser destructivo, como en el caso de los fagos líticos. Abedon, en el año 2008, menciona que existen profagos que forman partículas virales y las liberan espontáneamente al exterior de la célula hospedera, sin la necesidad de lisanlas, ya liberadas pueden infectar otras células y reproducirse de la misma forma. Dicho ciclo de vida es conocido como fase crónica y es incapaz de formar placas líticas, ya que en ningún momento se lisan células. El mismo autor menciona que los fagos tienen la capacidad de interactuar con células de su entorno ecológico y que muchas veces son incapaces de infectar células (aunque sean de la misma especie) que no pertenezcan a su nicho. Esto se debe a que las interacciones virus-célula han coevolucionado a través del tiempo y se han adaptado para funcionar solo con organismos del mismo ecosistema.

Ya que las cepas del laboratorio fueron obtenidas de diferentes ecosistemas, además de que ya están adaptadas a condiciones muy diferentes que las del ecosistema de donde provienen, es probable que los extractos fágicos (que probablemente si contenían fagos) no les afectaran. Sin embargo, estas explicaciones deben ser probadas. O, alternativamente, se puede comprobar o descartar la presencia de fagos en los extractos inducidos mediante otras técnicas, como por ejemplo microscopía electrónica, detección por anticuerpos, PCRs de secuencias conocidas, etcétera.

PREDICCIÓN DE PROFAGOS EN GENOMAS DE *Rhizobium*.

Como se menciona en el apartado anterior, paralelamente a las inducciones se realizaron búsquedas de profagos mediante un programa bioinformático. Existen muchos programas especializados en la búsqueda de profagos a partir de secuencias genómicas bacterianas. Algunos ejemplos son los siguientes: PHAST (ahora PHASTER), Phage classification tool set, Phage finder, Phage term, Phispy, virstorter, profhinder, prophage finder, etc. En este estudio se eligió el programa PHAST debido a la facilidad que tiene para usarse, la claridad de los resultados que arroja, su conexión con amplias bases de datos (genbank y EMBL) y su constante actualización (cada 6 meses aproximadamente) (Zhou *et al.*, 2010).

PHAST es un programa online al que se accede entrando a la página "phast.wishartlab.com". Acepta secuencias (fragmentadas o completas) en formato fasta o a través del número de acceso Genbank. Es un software que, al tener características heurísticas, arroja resultados de una manera muy rápida (aproximadamente de 1-5 minutos por genoma bacteriano). Trabaja comparando el genoma de interés con secuencias de profagos anotadas en las bases de datos (Zhou *et al.*, 2010).

Entre los resultados que arroja PHAST, los relevantes para este trabajo fueron:

- El la figura R7 se observa un gráfico circular que muestra el tamaño del genoma bacteriano y las ubicaciones genómicas aproximadas de los profagos encontrados. Así mismo se señala el tipo de profago, cuyo nombre y características, determinadas por el programa, se exponen a continuación:

- Profago intacto: Son las secuencias que presuntamente tienen todos los genes para que el profago entre en fase lítica, y por lo tanto, tenga la capacidad de producir y liberar partículas virales infectivas. Es importante señalar que el programa no detecta mutaciones genéticas, cuyo efecto podría traducirse en proteínas o sitios de regulación no funcionales, entre otras cosas, lo que evitaría la síntesis exitosa de viriones.
 - Profago cuestionable: Se refiere a secuencias de tamaños apreciables, generalmente mayores a 30 Kbs, cuya cantidad de CDSs es similar o mayor a la de los profagos intactos. Sin embargo no cuentan con algún(os) gen importante, como transposasas, integrasas o lisozimas. No obstante cuentan con un gran repertorio de genes hipotéticos y similares a los de fagos (PLPs, del inglés protein like phages), que aunque es poco probable, podrían sustituir la función de los genes antes mencionados.
 - Profago incompleto: Son secuencias pequeñas, generalmente 10 a 20 Kpbs, a las cuales les faltan varios genes importantes para la síntesis y liberación de viriones, pero que los genes que tienen son sin lugar a dudas provenientes de secuencias virales. Estos profagos son prácticamente incapaces de formar viriones.
- Gráfico lineal que muestra el mapa génico de cada profago. Además de su localización exacta, número de regiones codificantes (CDSs, del inglés coding sequence), porcentaje GC, tipo de genes (integrasas, de cápside, transposasas, lisozimas, etc.) y tipo de profago (intacto, incompleto o cuestionable). (Figura R8)
 - Archivo de texto con la secuencia genómica del profago.
 - Archivo de texto con el proteoma del profago. Donde también se menciona la ubicación del mismo, además del tamaño genómico y porcentaje GC de su hospedero.

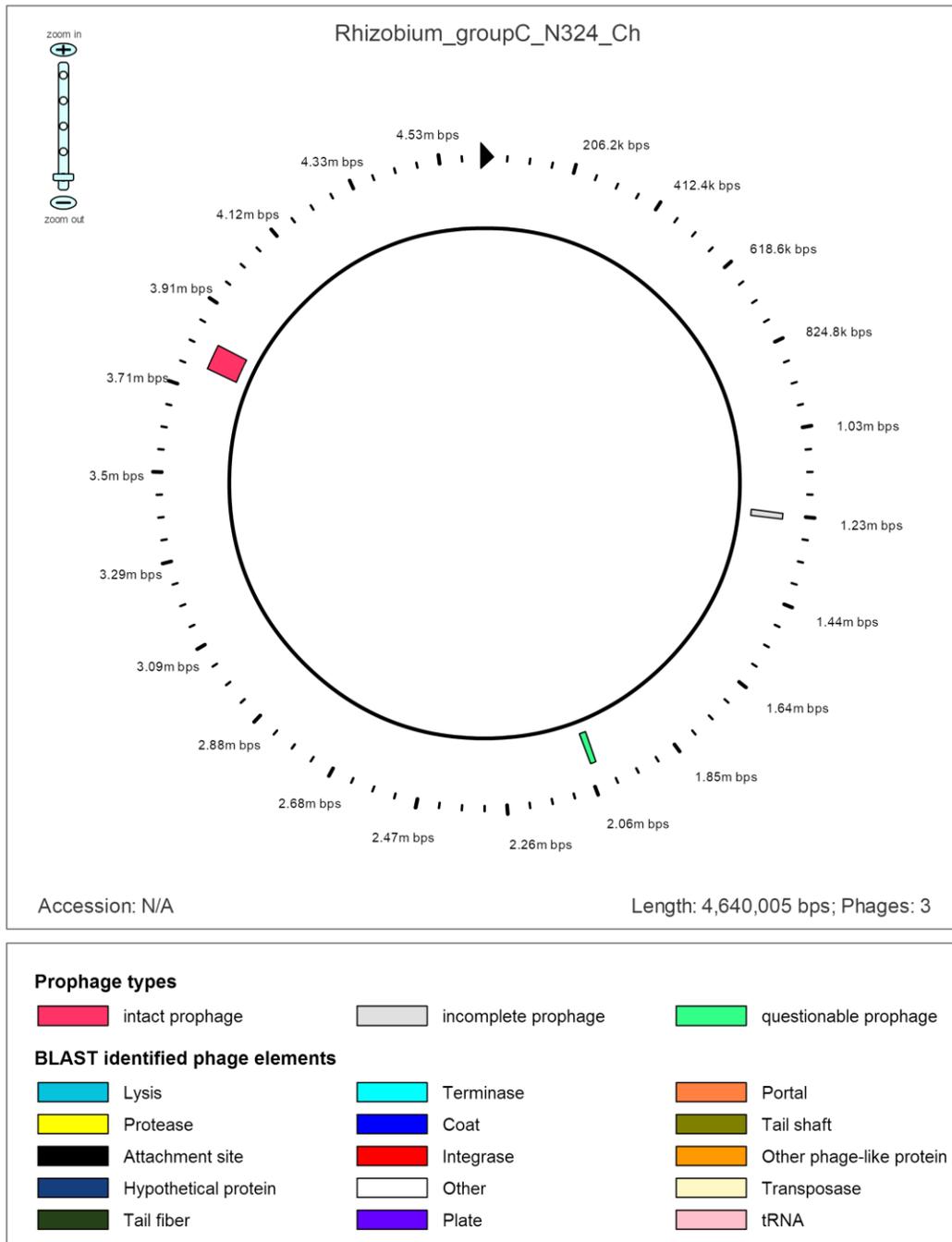


Figura R7. Representación gráfica del genoma de la cepa *R. etli* N324. Donde se observan los tres tipos profagos ubicado a las 1.23, 2.06 y 3.75 megabases, sucediéndose como incompleto, cuestionable e intacto (respectivamente). En la parte de arriba se observa el nombre de la cepa bacteriana, abajo del mapa circular, a la izquierda, se observa el numero de acceso (N/A en este caso ya que la secuencia no se ha subido a las bases de datos), a la derecha se observa la longitud del genoma bacteriano y el numero de profagos que contiene. En la parte hasta abajo se observan los colores con los que se puede ubicar cada tipo de profago dentro del mapa circular.

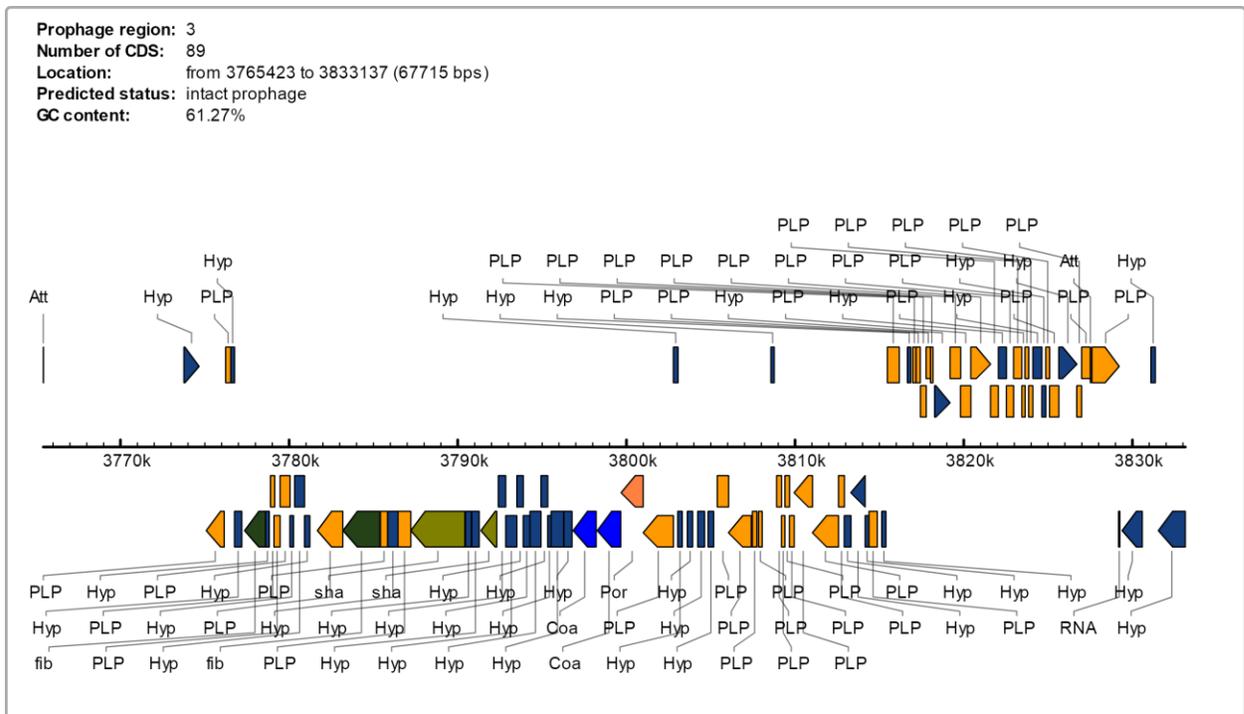


Figura R8. Gráfico que muestra el mapa génico (al centro) del profago intacto de la cepa *R. etli* N324. Arriba a la izquierda se observan el número de CDSs, la ubicación exacta con respecto al genoma bacteriano, el tipo de profago y su porcentaje GC. Abajo se observan los tipos de genes, diferenciados con colores correspondientes al mapa génico.

Se realizaron varias búsquedas bioinformáticas en diferentes momentos del desarrollo del proyecto, analizando los mismos genomas varias veces, ya que la base de datos de PAHST se actualiza regularmente, además de que los criterios mediante los cuales el programa busca profagos también se modifican regularmente. Por ejemplo, en segundas y terceras corridas algunos profagos que antes se mostraban intactos, se mostraron cuestionables; sin embargo, PHAST resultó ser una muy buena herramienta.

En total se analizaron 50 genomas de *Rhizobium*, y en 40 de ellos se encontraron 72 profagos: 10 intactos (profagos con los genes necesarios para formar viriones), 26 cuestionables (profagos grandes sin genes importantes, pero con secuencias muy similares a ellos) y 36 incompletos (profagos que no tienen genes importantes). En la tabla R6 se pueden observar las cepas que dieron positivo a profagos, cuyas especies y orígenes se detallan en el anexo.

En las figuras R7 y R8 se muestran los resultados correspondientes a la cepa *R. etli* N324, encontrada en el nódulo de una planta frijol de un campo de cultivo, en Tepoztlán,

Morelos. Las imágenes muestran una síntesis de los datos que arroja el programa, como el tipo de profagos encontrados y sus ubicaciones aproximadas (fig. R7), su tamaño y su mapa génico (fig. R8), entre otros aspectos. Además, el programa te permite bajar las secuencias nucleotídicas y proteicas de los fagos. Es importante señalar que este tipo de resultados fue analizado para cada una de las 50 cepas bacterianas analizadas.

En la tabla R6 pueden verse 40 cepas (de las 50 analizadas) y cada una contiene de 1-4 profagos. Allí mismo se ve el tipo de cada profago (intacto, cuestionable o incompleto) y su tamaño en pares de bases. Es interesante ver que el 80% de las cepas analizadas contienen profagos, lo que lleva a pensar que las cepas del genero *Rhizobium* tienen una transferencia horizontal de ADN muy activa, sugiriendo grandes interacciones genómicas fago-bacteria, y por ende bacteria-bacteria, debido al proceso de transducción. Esto lleva a pensar que la coevolución mediada transducción podría ser uno de los principales motores que estructuran los ecosistemas de las comunidades microbianas de los suelos.

Tabla R6. Profagos encontrados en 40 genomas pertenecientes a diferentes cepas de *Rhizobium*. En la primera columna observamos el nombre de la cepa, en la segunda el número de profagos encontrados en ella, en la tercera el tipo de profago (incompleto, cuestionable o intacto) y, en la cuarta, su tamaño en pares de bases pbs.

CORRIDA EN PHAST			
PROFAGO			
CEPA	NUMERO	CARACTERISTICA	TAMAÑO (pbs)
8C-3	1	Intacto	39144
	2	Cuestionable	23976
	3	Intacto	46897
Bra-5	1	Intacto	57379
	2	Cuestionable	36875
CIAT-894	1	Incompleto	20429
CNPAF	1	Incompleto	14266
GR56	1	Cuestionable	25042
Kim5	1	Incompleto	10689
	2	Cuestionable	24738
	3	Incompleto	8029
MIM1	1	Incompleto	10741
N122	1	Intacto	23034
	2	Incompleto	13646
	3	Cuestionable	22645
	4	Cuestionable	37829
N1314	1	Incompleto	9040
	2	Incompleto	22856
N1341	1	Cuestionable	20382
N161	1	Incompleto	16243

Tabla R6. Continuación

N261	1	Incompleto	16168
	2	Incompleto	7043
	3	Incompleto	21316
N324	1	Incompleto	16850
	2	Cuestionable	17878
	3	Intacto	67714
N4311	1	Incompleto	16850
	2	Cuestionable	18299
N541	1	Incompleto	16742
	2	Cuestionable	17878
N561	1	Incompleto	22053
	2	Incompleto	20128
N621	1	Cuestionable	20128
N6212	1	Cuestionable	20382
N671	1	Cuestionable	22195
	2	Cuestionable	41459
N731	1	Incompleto	9040
	2	Incompleto	22856
N741	1	Cuestionable	20382
N771	1	Cuestionable	22402
	2	Cuestionable	41459
N871	1	Incompleto	22053
	2	Incompleto	20128
N941	1	Incompleto	16742
	2	Incompleto	17980
NXC14	1	Cuestionable	10552
	2	Incompleto	9771
	3	Intacto	39174
	4	Incompleto	7665
NXC24	1	Incompleto	17977
	2	Intacto	35113
NXT3	1	Incompleto	20435
R339	1	Intacto	63081
	2	Incompleto	17902
R611	1	Incompleto	10726
R620	1	Cuestionable	27030
R630	1	Intacto	33059

Tabla R6. Continuación.

R634	1	Incompleto	24020
R635	1	Incompleto	30267
R650	1	Incompleto	10726
R693	1	Cuestionable	34039
	2	Cuestionable	9452
R711	1	Cuestionable	9357
	2	Cuestionable	46384
R72	1	Incompleto	10018
	2	Cuestionable	21710
R723	1	Cuestionable	19527
	2	Incompleto	7023
	3	Incompleto	21316
R744	1	Incompleto	8366
TAL182	1	Cuestionable	60429
	2	Intacto	23328

Por otro lado, al ver que la mayoría de los profagos son incompletos o cuestionables, podría pensarse que se insertaron hace mucho tiempo y han evolucionado como parte del genoma bacteriano, acoplándose a características propias de la bacteria hospedera, como por ejemplo el porcentaje GC, uso de codones, etc. (Lawrence and Ochman, 1998). Además, podría suponerse (ya que no hubo inducciones exitosas) que estas secuencias han perdido la capacidad de formar viriones debido a la mutación o delección (completa o fragmentada) de genes importantes, entre otras posibles explicaciones. Estas y otras cuestiones pueden investigarse a fondo para poder entender la evolución de dichas secuencias dentro de los genomas bacterianos y el impacto que han tenido (y que tienen) sobre los mismos, ya sea a nivel poblacional, comunitario y/o ecológico.

Haciendo hincapié en el número y tipo de profagos que se encuentran en los genomas analizados, la mayoría cuestionables e incompletos, que quizá se integraron hace bastante tiempo, se procedió a la comparación de algunos profagos (todos predichos por PHAST) encontrados en genomas de rizobiales y *E. coli*. Los resultados (figura R9) muestran una clara diferencia entre los de rizobiales y los de *E. coli*, en especial en los que tienen la característica de intacto (tabla derecha), apoyando la suposición de que los profagos de *E. coli* se han integrado recientemente, a diferencia de los encontrados en genomas de rizobiales.

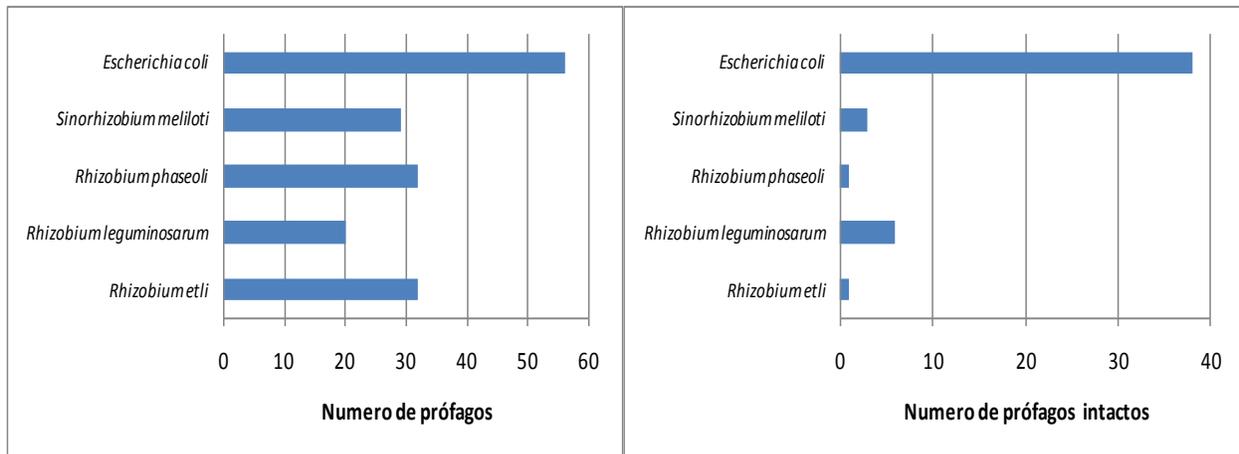


Figura R9. Tablas comparativas de profagos encontradas en 5 especies diferentes. A la izquierda se observa el conteo de fagos totales por especie, a la derecha (también por especie) se observa solamente el numero de profagos intactos.

Como se ha mencionado en párrafos anteriores, la elección de las cepas de *Rhizobium* para inducir se basó en el contenido y características de profagos encontrados con PHAST. De los 72 profagos encontrados se eligieron 17, los cuales podrían estar completos y ser funcionales. Dicha elección se basó en el tamaño genómico y en el numero de regiones codificantes (CDSs). Cada uno de estos 17 profagos se analizó por separado con PHAST ya que los resultados que arroja el programa son más detallados si la secuencia analizada es más pequeña, a diferencia de los resultados mostrados cuando se analizan genomas bacterianos completos que son mucho más grandes que la secuencia de un solo profago. De los análisis, a muchos profagos no se les detectaron proteínas importantes (transposasas, integrasas ó lisozimas) pero se encontró que cuentan con muchos genes hipotéticos (genes con función desconocida) y de proteínas parecidas a las de fagos (PLPs), cuya función podría ser importante para la producción de partículas virales funcionales. En la tabla R7 puede observarse el contenido génico de cada profago, además de su tipo (según PHAST), tamaño genómico, número de CDSs y su cepa de origen.

Las 13 cepas que contienen a los profagos antes mencionados (tabla R7) se sometieron a diferentes tratamientos de inducción (apartado “Inducción de cepas de *Rhizobium*”) y, como ya se mencionó, no se encontró ninguna evidencia de inducción, ni en curvas de crecimiento bacteriano ni en placas líticas. Sin embargo los resultados de PHAST fueron muy útiles para agrupar profagos, junto con algunos fagos de *Rhizobium* conocidos, y diseñar marcadores moleculares que podrían servir para determinar si cepas de *Rhizobium* no secuenciadas contengan secuencias propias de profagos.

Tabla R7. Proteomas de profagos probablemente completos y funcionales. Kpb (kilopares de bases), CDSs (regiones codificantes), Att (secuencias de unión al DNA del hospedero), Hip. (proteínas hipotéticas) y PLK (proteínas parecidas a las de otros fagos).

PROTEOMA DE PROFAGOS PROBABLES INTACTOS. Corridos en PHAST a partir de la secuencia del profago (no de la cepa)														
PROFAGOS	TIPO	TAMAÑO (Kpb)	CDSs	TIPOS DE PROTEINAS										
				Lisis	Proteasas	Att	Hip.	Fib. de cola	Terminasas	Cápside	Integrasas	Otras	De placa	
8C3-1	Intacto	39.1	54				33				1		1	3
8C3-3	Intacto	35.9	48			2	27		1		3	1	3	
Bra5-1	Intacto	57.4	92				43	2			2		5	
Bra5-2	Cuestionable	25.4	32				18	1	1		1	1		
NXC14-3	Cuestionable	25.7	29				12	1			2	1		
TAL182-1	Cuestionable	60.4	91			2	21				1		3	
TAL182-2	Incompleto	23.3	28				12		1		1			1
N122-1	Intacto	23	32		1		15				2		3	
N122-4	Cuestionable	29.3	41		1		26	1						
N324-3	Intacto	61.8	71			2	24	1			2		4	
N671-2	Incompleto	18	20			2	9					1	3	
N771-2	Incompleto	18	20			2	9					1	3	
NXC24-2	Cuestionable	27.3	36				20	1	1		1	1	1	
R339-1	Intacto	58.8	92			2	35	2			2	1	4	
R630-0	Cuestionable	33	47			2	28		1		1	1	1	
R693-1	Cuestionable	24.4	26			2	9		1		2	1	8	
R711-2	Incompleto	36.1	51				17				2		1	

AGRUPAMIENTO, POR SIMILITUD NUCLEOTÍDICA, DE PROFAGOS PREDICHOS POR PHAST.

Para obtener datos sobre la diversidad de secuencias de profagos predichas por PHAST, las 72 secuencias nucleotídicas de profagos, además de las 11 secuencias de fagos completos, se compararon y agruparon mediante un alineamiento múltiple realizado con el programa MAUVE (MAUVE, on line), cuyo resultado se muestra en las figuras R10 y R11.

En la figura R10 se observa una matriz arrojada por MAUVE. Se ve que cada secuencia es comparada contra cada una de las 83 secuencias, visualizando un total de 6889 comparaciones (83*83). Los colores representan el porcentaje de identidad nucleotídica, es decir, en que porcentaje de secuencias nucleotídicas se parece un par de secuencias; mientras más amarillo, el parecido es mayor. Las secuencias correspondientes a los 72 profagos vienen nombradas por el nombre de la cepa y la mayoría tienen un guion y después el número 1, 2, 3 o 4, que representa el número de profago que se encontró en esa cepa, por ejemplo, el nombre TAL182-2 representa al segundo profago de la cepa TAL182, ó el N741 (sin guión ni numero sucesivo) representaría al único profago perteneciente a dicha cepa. Las secuencias de fagos completos se nombran con las letras RHEph y un número que representa al número de fago, por ejemplo, RHEph10 representa al fago 10 de *R. etli* CE3.

En dicha matriz comparativa (figura R10) se visualiza el parecido entre secuencias. Por ejemplo, el profago N874-2 es muy similar a los profagos N6212, N74, N341 y N871-1, pero muy diferente a los profagos TAL182-2, X8C3-2 y X8C3-3, así como a los fagos RHEph04, RHEph05 y RHEph06. Por otro lado, observamos que el profago N541-1 (el primero de los verticales) es muy diferente a las primeras 50 secuencias horizontales y muy similar a las últimas 25. Es importante mencionar que en este tipo de matriz los únicos resultados relevantes corresponden a las secuencias parecidas (color amarillo), ya que las secuencias alejadas podrían parecerse muy poco o nada, es decir, el programa solo considera homologías, las diferencias no son consideradas y se acomodan al azar.

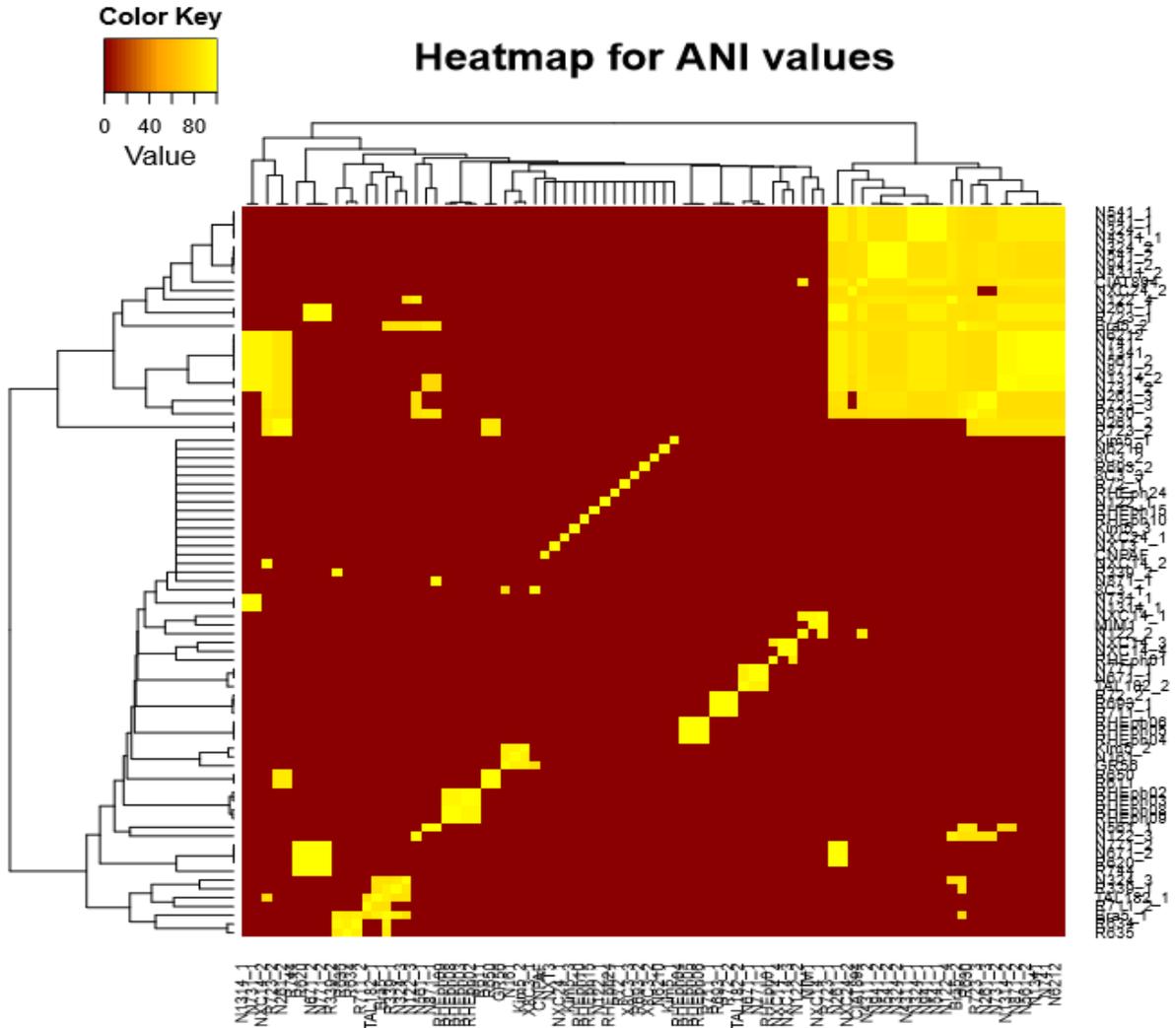


Figura R10. Dendrograma tipo matriz que muestra el porcentaje aproximado de similitudes ó identidades nucleotídicas (ANI) entre pares de secuencias de las 83 analizadas. Entre más amarillo sea el cuadro el porcentaje de identidad nucleotídica es mayor.

La figura R11 muestra un dendrograma tipo árbol, también obtenido a partir de MAUVE, el cuál se modificó para agrupar las secuencias más cercanas entre sí en 14 grupos. En cada grupo se detallan 3 aspectos: 1) número de pares de bases compartidas; 2) porcentaje que cubre la secuencia compartida con respecto a un promedio del tamaño de todas las secuencias dentro del grupo; y 3) porcentaje de identidad nucleotídica (ANIm) correspondiente al fragmento compartido, es decir, el ANIm de las pbs compartidas. Adicional a ello se agregó un código de colores que representa las especies bacterianas de *Rhizobium* donde se encontraron los profagos o, en el caso de los nombres en rojo, los fagos completos de la colección del laboratorio.

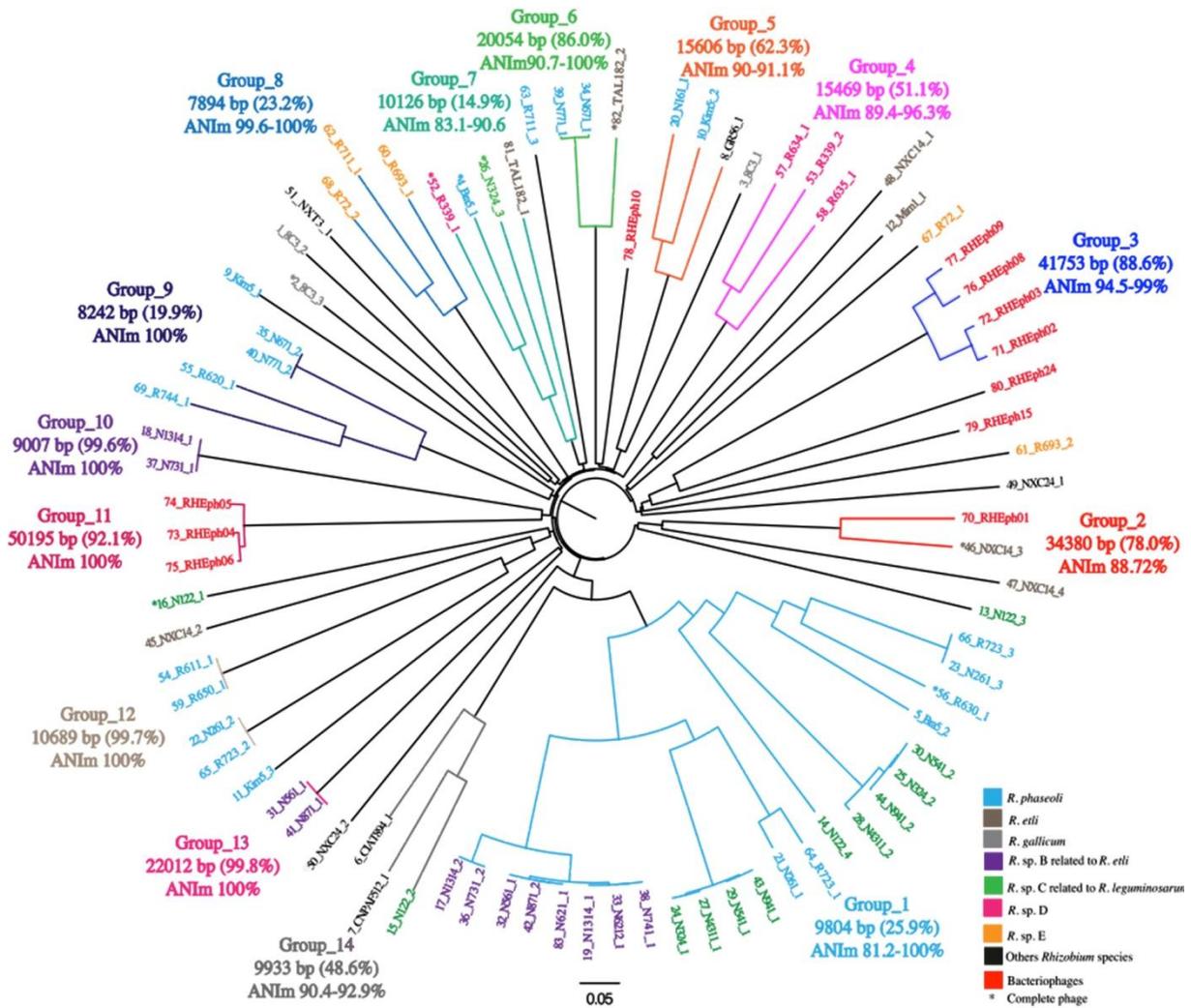


Figura R11. Agrupamiento de 72 profagos y 11 fagos de bacterias del genero *Rhizobium*.

En la tabla R8 se detalla la distribución de los 14 grupos. Por ejemplo, el grupo 1 (G1) cuenta con 23 profagos, de los cuales 22 pertenecen a cepas colectadas en Tepoztlán, Morelos y 1 a una cepa proveniente de Brasil (Bra5). En la parte de arriba del cada grupo se encuentra el número de pares de bases compartidas por las 23 secuencias, 9805 para dicho grupo. Adicional a los 14 grupos se puede observar un apartado llamado “Sin grupo”, donde se encuentran las secuencias que no emparejaron con ninguna otra, 19 en total (16 profagos y 3 fagos). Con respecto al origen geográfico de los profagos y fagos, es interesante ver que hay grupos más heterogéneos que otros, por ejemplo, en el grupo 7 observamos 4 profagos con origen en EUA, Brasil y Tepoztlán (México). En el grupo 5 se encuentran secuencias procedentes de España, EUA y Tepoztlán. Las secuencias resaltadas en gris representan a los fagos completos y las secuencias con letras rojas corresponden a los profagos que se piensa tendrían la capacidad de ser inducidos a fase lítica (ver tabla R7).

Tabla R8. Fagos y profagos agrupados de acuerdo con las secuencias nucleotídicas que comparten.

G1 (9805pb)		G2 (68794pb)		G8 (7895pb)		G14 (9934pb)	
BRAS-2	BRASIL	NXC14-3	MORELOS	R693-1	TEPOZTLAN	CIAT894-0	COSTA RICA
N122-4	TEPOZTLAN	RHEph01	ZACATECAS	R711-1	TEPOZTLAN	CNPAF-0	BRASIL
N1314-2	TEPOZTLAN	G3 (41754pb)		R72-2	TEPOZTLAN	N122-2	TEPOZTLAN
N1341-0	TEPOZTLAN	RHEph02	EDO. MEX.	G9 (8243pb)		Sin grupo (0pb)	
N261-1	TEPOZTLAN	RHEph03	MORELOS	N671-2	TEPOZTLAN	N122-3	TEPOZTLAN
N261-3	TEPOZTLAN	RHEph08	MORELOS	N771-2	TEPOZTLAN	RHEph10	EDO. MEX.
N324-1	TEPOZTLAN	RHEph09	MORELOS	R620-0	TEPOZTLAN	RHEph15	TEPOZTLAN
N324-2	TEPOZTLAN	G4 (15471pb)		R744-0	TEPOZTLAN	RHEph24	TEPOZTLAN
N4311-1	TEPOZTLAN	R635-0	TEPOZTLAN	G10 (9008pb)		8C3-1	ESPAÑA
N4311-2	TEPOZTLAN	R339-2	TEPOZTLAN	N1314-1	TEPOZTLAN	8C3-2	ESPAÑA
N541-1	TEPOZTLAN	R634-0	TEPOZTLAN	N731-1	TEPOZTLAN	8C3-3	ESPAÑA
N541-2	TEPOZTLAN	G5 (15607pb)		G11 (50196pb)		KIM5-1	USA
N561-2	TEPOZTLAN	GR56-0	ESPAÑA	RHEph04	MORELOS	KIM5-3	USA
N621-0	TEPOZTLAN	KIM5-2	USA	RHEph05	ZACATECAS	MIM1-0	MORELOS
N6212-0	TEPOZTLAN	N161-0	TEPOZTLAN	RHEph06	EDO. MEX.	N122-1	TEPOZTLAN
N731-2	TEPOZTLAN	G6 (20055pb)		G12 (10690pb)		NXC14-1	MORELOS
N741-0	TEPOZTLAN	TAL182-2	USA	R611-0	TEPOZTLAN	NXC14-2	MORELOS
N871-2	TEPOZTLAN	N671-1	TEPOZTLAN	R650-0	TEPOZTLAN	NXC14-4	MORELOS
N941-1	TEPOZTLAN	N771-1	TEPOZTLAN	N261-2	TEPOZTLAN	NXC24-2	MORELOS
N941-2	TEPOZTLAN	G7 (10127pb)		R723-2	TEPOZTLAN	NXT3-0	MORELOS
R630-0	TEPOZTLAN	TAL182-1	USA	G13 (22013pb)		R693-2	TEPOZTLAN
R723-1	TEPOZTLAN	N324-3	TEPOZTLAN	N561-1	TEPOZTLAN	R711-2	TEPOZTLAN
R723-3	TEPOZTLAN	BRAS-1	BRASIL	N871-1	TEPOZTLAN	R72-1	TEPOZTLAN
		R339-1	TEPOZTLAN				

IDENTIFICACIÓN DE SECUENCIAS DE PROFAGOS EN CEPAS DE *RHIZOBIUM* NO SECUENCIADAS

Lo interesante de agrupar las secuencias mencionadas radicó en la construcción de marcadores moleculares (segmentos de ADN) que representan grupos de fagos y profagos de *Rhizobium*. De tal forma que se puedan buscar profagos en cepas de *Rhizobium* no secuenciadas. De hecho, en la colección del laboratorio, de aproximadamente 122 cepas de *Rhizobium*, cerca de 70 no han sido secuenciadas. La ubicación de secuencias profágicas en cepas no secuenciadas puede ser de gran utilidad para posteriores experimentos de inducción o para elegir cuales genomas de *Rhizobium* vale la pena secuenciar, agilizando el proceso de investigación del impacto que tienen los profagos en el metabolismo, evolución y ecología de bacterias del genero *Rhizobium*.

Para la generación de marcadores se eligieron los grupos más heterogéneos en cuanto a origen (1, 2, 5, 6, 7, 8 y 14), asegurando una mayor diversidad de profagos y cepas de *Rhizobium*, por lo que la búsqueda en cepas no secuenciadas podría ser más exitosa. Por ejemplo, en el grupo 7 encontramos profagos de Brasil, EUA y México, y en el grupo 5 los encontramos de España, EUA y México (tabla R7).

Para la construcción de los marcadores se utilizó un script en Phyton (PHYTON, on line) que convirtió a formato fasta las secuencias de los grupos con los fagos y profagos de los grupos mencionados. Después se analizaron las secuencias fasta mediante la plataforma del programa MEGA (MEGA, on line), donde se alinearon las secuencias de cada grupo para elegir los fragmentos comunes. Este programa tiene la característica de alinear

Tabla R9. Tamaño y procedencia de los marcadores de profagos de *Rhizobium*.

Marcadores moleculares de fagos y prófagos de <i>Rhizobium</i>					
Grupo	Cepas del grupo	Origen	Pares de bases de la secuencia nucleotídica en común	Oligonucleotidos (de 5' a 3')	
				Forward	Reverse
1	BRA5-2	BRASIL	898-901pbs	GGCGGCCGCACCATCGT	TCGCGCCAAGACGATCGG
	N122-4	TEPOZTLAN			
	N1314-2	TEPOZTLAN			
	N1341-0	TEPOZTLAN			
	N261-1	TEPOZTLAN			
	N261-3	TEPOZTLAN			
	N324-1	TEPOZTLAN			
	N324-2	TEPOZTLAN			
	N4311-1	TEPOZTLAN			
	N4311-2	TEPOZTLAN			
	N541-1	TEPOZTLAN			
	N541-2	TEPOZTLAN			
	N561-2	TEPOZTLAN			
	N621-0	TEPOZTLAN			
	N6212-0	TEPOZTLAN			
	N731-2	TEPOZTLAN			
	N741-0	TEPOZTLAN			
N871-2	TEPOZTLAN				
N941-1	TEPOZTLAN				
N941-2	TEPOZTLAN				
R630-0	TEPOZTLAN				
R723-1	TEPOZTLAN				
R723-3	TEPOZTLAN				
2	NXC14-3	MORELOS	773-786pbs	CAGCTCTTCTACCAGCTC	CACTGGTAGTCGTAGCTG
	RHEph01	ZACATECAS			
5	GR56-0	ESPAÑA	765pbs	GCGTCGATGGTGCCGATC	CTGGTGACGGCCATCCAC
	KIM5-2	USA			
	N161-0	TEPOZTLAN			
6	TAL182-2	USA	958pbs	GAGCAGTTGTCGGCGGGC	CACTGCCCGTTGACGGG
	N671-1	TEPOZTLAN			
	N771-1	TEPOZTLAN			
7	TAL182-1	USA	946-970pbs	ATGGAAATTAATCTACATTC	ATCTGAACGGCAAGATTG
	N324-3	TEPOZTLAN			
	BRA5-1	BRASIL			
R339-1	TEPOZTLAN				
8	R693-1	TEPOZTLAN	578-579pbs	GCGAGCGACATGAAGACG	GGCGAATTCGACGATGAG
	R711-1	TEPOZTLAN			
	R72-2	TEPOZTLAN			
14	CIAT894-0	COSTA RICA	706pbs	CTGACGGCGCCGAGCAG	GGATCGCCGGATTGACCTTG
	CNPAF-0	BRASIL			
	N122-2	TEPOZTLAN			

De los 14 marcadores diseñados, el que representa al grupo 1 se probó en 70 genomas de *Rhizobium* no secuenciados. Los resultados (figuras R13-R16) mostraron que la región (de 887 a 899 pbs) amplificó en las siguientes 13 cepas: Inc1-2, Inc2-5, Tux10P, 343, B27, D13, D21, 4C-2, 16NJ-2, 17NJ-2, 6PR-1, GR10 y GR62. Es importante señalar que se utilizó la cepa Bra-5 como control positivo.

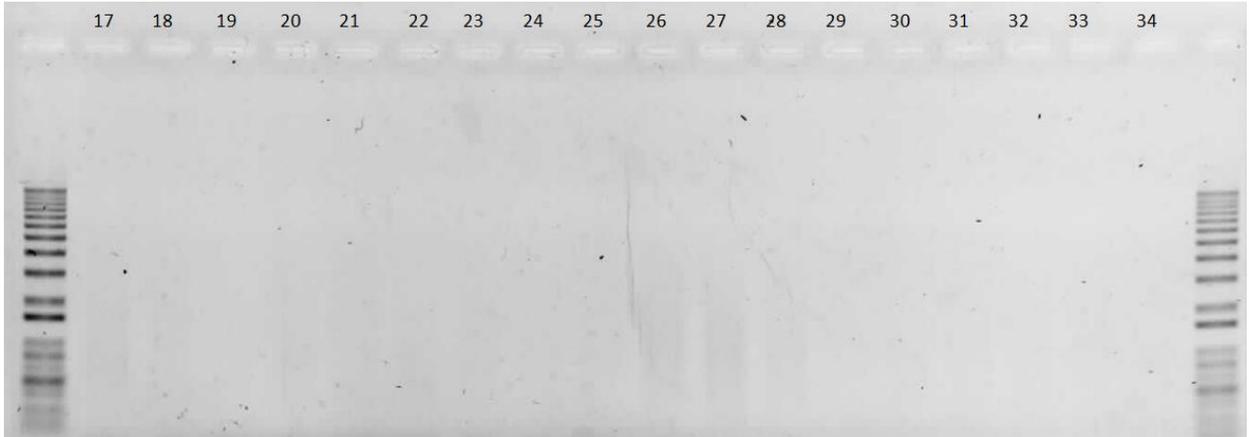


Figura R13. Amplificaciones por PCR del marcador del grupo 1, cepas 17-34.

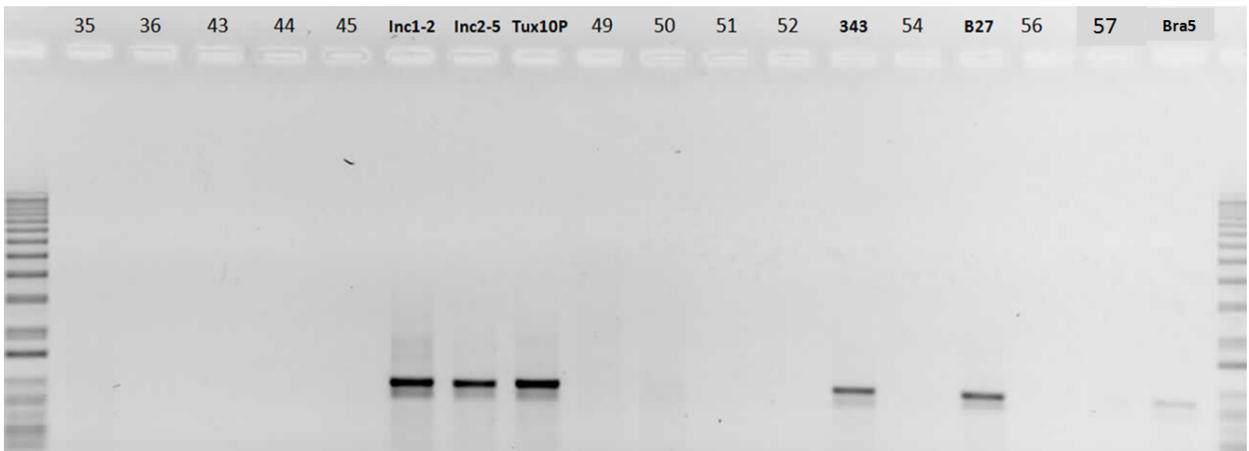


Figura R14. Amplificaciones por PCR del marcador del grupo 1, cepas 35-36 y 43-57.

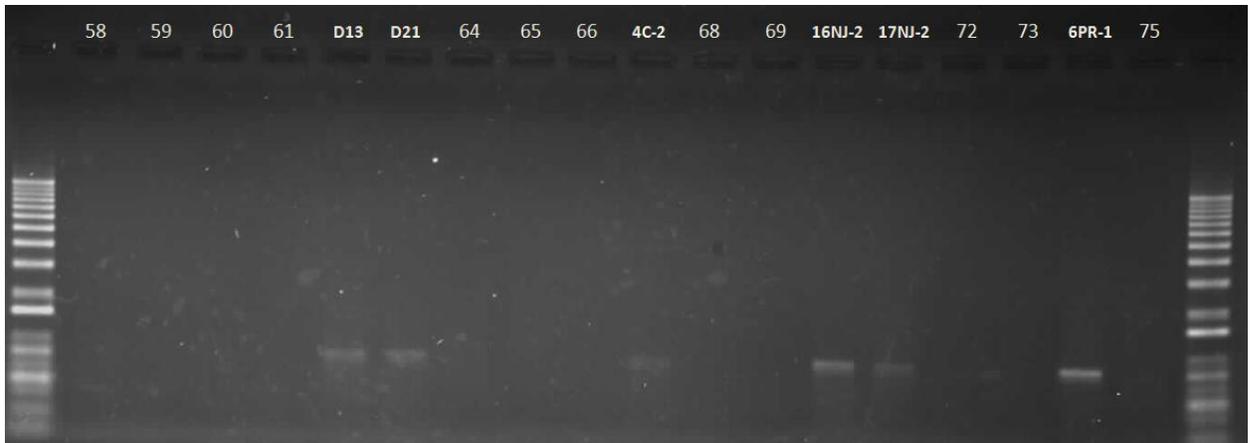


Figura R15. Amplificaciones por PCR del marcador del grupo 1, cepas 58-75.

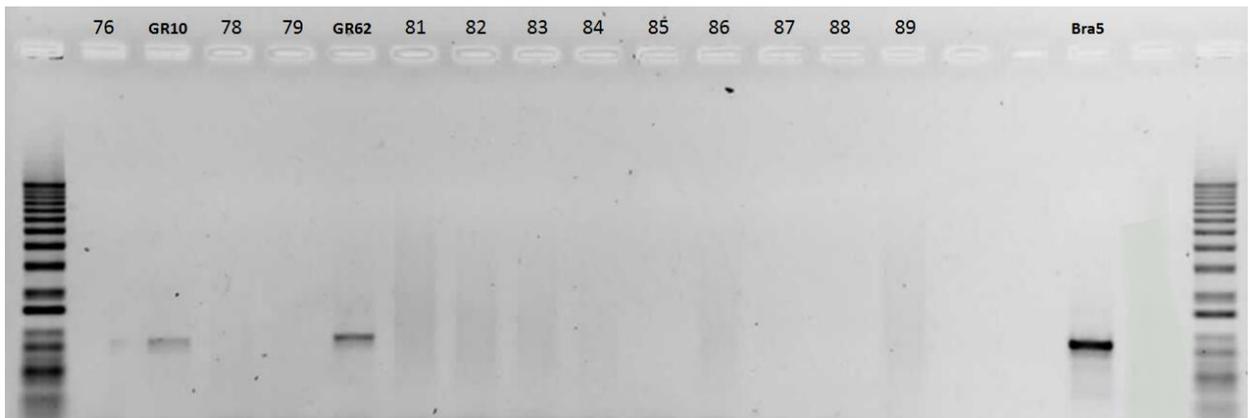


Figura R16. Amplificaciones por PCR del marcador del grupo 1, cepas 76-89.

Los resultados anteriores no son prueba concluyente de que las cepas que dieron positivo cuenten con profagos; sin embargo, es muy probable que los tengan, y que además sean muy similares a los encontrados en el grupo 1. Además, hace falta la búsqueda con los 6 marcadores restantes, y posteriormente comprobar la presencia de todos los posibles profagos ubicados con estos marcadores, lo que podría realizarse mediante la secuenciación de los fragmentos genómicos que contengan la secuencia marcadora.

CONCLUSIÓN

Los profagos que se encuentran en las cepas de género *Rhizobium* no han mostrado ser funcionales. Aunque la metodología de inducción se estandarizó cuidadosamente, utilizándose una buena cantidad de inductores (físicos y químicos) y comprobándose la efectividad de algunos en la cepa control (CE3 lisogenizada con el fago RHEph10), ningún profago bioinformáticamente tuvo la capacidad de inducirse. Y remarco, aunque el control positivo siempre fue indicador de que la inducción sucedía, quizá la metodología no era la adecuada para inducir a los otros profagos; tal vez por incompletos o mutados, ó tal vez porque no podían ser transcritos, no lo sabemos aún; incluso pudo haber sido algo que no imaginemos. El hecho fue que las inducciones no fueron exitosas.

Por otro lado, la gran cantidad de secuencias virales, tanto de fagos como de profagos, fueron utilizadas para la construcción de marcadores moleculares que podrían utilizarse para la identificación de prófagos (o segmentos de ellos) en cepas de *Rhizobium* no secuenciadas.

Por el momento, dichos marcadores solo serán útiles en el campo de la investigación básica, ubicando cepas candidatas a estudios más profundos, como secuenciación de genomas de cepas candidatas (o fragmentos de los mismos), inducción con diferentes metodologías, combinación con otras secuencias para darle a la cepa la capacidad de ser inducida, entre otros. Y tal vez en un futuro las aplicaciones podrían expandirse hacia el campo de la ciencia aplicada.

PERSPECTIVAS

Es indudable que este trabajo tiene mucho futuro. Apenas estamos comenzando a ubicar y caracterizar las secuencias profágicas de bacterias del Género *Rhizobium*. Y el impacto que podrían tener sobre los ecosistemas y la evolución de genomas bacterianos, ni siquiera se ha comenzado a vislumbrar.

Los siguientes puntos sugieren algunos estudios que, a mi consideración, son necesarios para continuar con el estudio de la lisogenia en *Rhizobium*:

- Amplificar la búsqueda bioinformática de profagos en mas genomas de *Rhizobium*, y añadir otros Géneros rhizobiales (*Sinorhizobium* por ejemplo).
- Utilizar otros programas de búsqueda para confirmar los resultados arrojados por PHAST (ahora PHASTER).
- Ubicar las zonas genómicas donde se encuentran los profagos. Además de organizarlas y analizarlas para encontrar posibles patrones (p. ej. “sitios calientes”) e implicaciones metabólicas (conversión lisogénica).
- Realizar análisis transcriptómicos de genes profágicos en diferentes cepas de bacterias rhizobiales para ubicar los CDSs que se expresan. Y si es que los hay, analizarlos para encontrar posibles patrones. Esto llevaría, por ejemplo, a encontrar secuencias profágicas que han mantenido una expresión constante a lo largo de la evolución.
- Realizar los estudios antes mencionados, además de las inducciones de cepas lisogénicas, en grupos bacterianos recién colectados de diferentes ambientes. Esto porque probablemente los profagos se comportan diferente en células de laboratorio que en células silvestres; inclusive podría haber diferencias entre las cepas extraídas de un campo de cultivo y las extraídas de un ecosistema natural.

BIBLIOGRAFÍA

- Abedon S.** (2008). *Bacteriophage ecology. Population growth, evolution and impact of bacterial viruses*. AMCM press, Cambridge, U.K.
- Adams M.**, Lefkowitz E., King A., Harrach B., Harrison R., Knowles N., et. al. (2017). *Changes to taxonomy and the international code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses*. Archives of Virology, 74: 2391-2399.
- Ackerman H.** (2011). *Bacteriophage taxonomy*. Microbiology Australia (Under the microscope), pp. 90-94.
- Anderson R.**, Sogin M. y Baross J. (2014). *Evolutionary strategies of viruses, archea and bacteria in hydrothermal vents ecosystems revealed through metagenomics*. PLoS ONE, 9:10.
- Arkhipova I.**, Batzer M., Brosius., Feschotte C. y Schmitz J. (2012). *Genomic impact of eukaryotic transposable elements*. Mobile DNA, 3:19.
- Barnet Y.** y Vincent J. (1970). *Lysogenic conversion of Rhizobium trifolii*. Journal of General Microbiology, 61(3), 319-25.
- Barrangou N.**, Fremaux C., Deveau H., Richards M., Boyaval P., Moineau S., et al. (2007). *CRISPR provides acquired resistance against viruses in prokaryotes*. Science, 315: 1709-1712.
- Beijerinck M.** (1898). *Concerning a contagium vivum fluidum as cause of the spot disease of tobacco leaves*. Phytopathological Classics (American Phytopathological Society), 7.
- Beiko R.**, Harlow T. y Ragan M. (2005). *Highways of gene sharing in prokaryotes*. Proceedings of the National Academy of Sciences of the United States of America, 102: 14332–14337.
- Bertani G.** (1953). *Lysogenic versus lytic cycle of phage multiplication*. Cold Spring Harbor Symposium Quantitative Biology, 18: 65-70.
- Brazelton W.** y Baross. (2009). *Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm*. The ISME Journal, 3: 1420-1424.
- Brussow H.** y Hendrix R. (2002). *Phage genomics: Small is beautiful*. Cell, 108:13-16.
- Campbell A.** (2003). *The future of bacteriophage biology*. Nature Reviews genetics, 4: 471-477.
- Canchaya C.**, Proux C., Fournous G., Bruttin A. y Brussow H. (2003). *Prophage genomics*. Microbiology and Molecular Biology Reviews, 67, 2: 238-276.

Canchaya C., Fournous G. y Brussow H. (2004). *The impact of prophages on bacterial chromosomes*. *Molecular microbiology*, 53(1): 9-18.

Cann A. (2005). *Principles of molecular virology*. 4ª edición, Elsevier Academic Press, Oxford U.K. 332p.

Chibani-Chennoufi S., Bruttin A., Dillmann M. y Brussow H. (2004). *Phage-Host interaction, an ecological perspective*. *Journal of Bacteriology*, 186(12): 3677-3686.

Choi J., Kotay S. y Goel R. (2010). *Various physico-chemical stress factors cause prophage induction in Nitrosospira multiformis 25196--an ammonia oxidizing bacteria*. *Water Research*, 44(15): 4550-4558.

Coleman D., Knights J., Russell R., Shanley D., Birkbeck TH., Dougan G., *et al.* (1991). *Insertional inactivation of the Staphylococcus aureus β -toxin by bacterio- phage phi 13 occurs by site- and orientation-specific integration of the phi 13 genome*. *Molecular Microbiology*, 5: 933-939.

DeMarini D. y Lawrence B. (1992). *Prophage induction by DNA topoisomerase II poisons and reactive-oxygen species: Role of DNA breaks*. *Mutation Research*, 267: 1-17.

Deng W., Monks L. y Neuer S. (2015). *Effects of clay minerals on the aggregation and subsequent settling of marine Synechococcus*. *Limnology and Oceanography*, 60(3): 805-816.

D'Herelle F. (2007). *On an invisible microbe antagonistic toward dysenteric bacilli*. *Research in microbiology*, 158(7): 153-154.

Dimmock N., Easton A. y Leppard K. (2007). *Introducción to modern virology*. 6ª edición, Blackwell publishing, Oxford, UK. 531p.

Elsaied H., Stokes H., Nakamura T., Kitamura K., Fuse H. y Maruyama A. (2007). *Novel and diverse integron integrase genes and integron-like gene cassettes are prevalent in deep-sea hydrothermal vents*. *Environmental Microbiology*, 9: 2298-2312.

Evans J., Barnet Y. y Vincent J. (1979). *Effect of a bacteriophage on the colonisation and nodulation of clover roots by a strain of Rhizobium trifolii*. *Can J. Microbiol.*, 25(9): 968-973.

Faruque M., Asadhulgani P., Abdul A., Albert M., Nasirul K. y Mekalanos J. (1998). *Induction of the lysogenic phage encoding cholera toxin in naturally occurring strains of toxigenic Vibrio cholerae O1 y O139*. *Infection Immunity*, 66(8): 3752-3757.

Fasunon O. y Uwaifo A. (1989). *Prophage induction by 4 antimalaria drugs (Daraprim, Fansidar, Nivaquine and Camoquine) and in combination with aflatoxin B₁*. *Mutation Research (Genetic Toxicology)*, 222(4): 311-316.

Flint S., Racaniello V., Enquist L. y Skalka A. (2009). *Principles of virology*. 3ª edición, ASM press, Washington, E.U.A. 1034p.

Forterre P. (2013). *The common ancestor of archaea and eukarya was not an archaeon*. *Archea*, 2013: 372396.

Franklin R. (1956). *Structure of tobacco mosaic virus*. *Nature*, 177: 928-30.

Friedman D. y Gottesman M. (1983). *Lambda II (Chapter: Lytic model of lambda development)*. Cold Spring Harbor Laboratory Press, New York, E.U.A.

Gelderblom H. (1996). *Structure and Classification of Viruses (chapter 41)*. 4ª edición, In: Baron S, editor, Texas, U.S.A.

Grose J. y Casjens S. (2014). *Understanding the enormous diversity of bacteriophages: The tailed phages that infect the bacteria family Enterobacteriaceae*. *Virology*, 468-470: 421-443.

Hatfull G. (2008). *Bacteriophage genomics*. *Current opinion in microbiology*, 11: 447-453.

Hatfull G. y Hendrix R. (2011). *Bacteriophages and their genomes*. *Current Opinion in Virology*., 1(4): 298-303.

Heinemann S. y Spiegelman W. (1970). *Role of the gene N product in phage lambda*. *Cold Spring Harbor Symposium of Quantitative Biology*, 35: 315-318.

Hendrix R., Smith M., Burns R., Ford M. y Hatfull G. (1999). *Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage*. *Proceedings of the National Academy of Sciences of the United States of America*, 96: 2192-2197.

ICTV: Oficial web side. <http://ictvonline.org/>

Iwanowski D. (1892). *Concerning the mosaic disease of the tobacco plant*. Ithaca press, New York, 27-32pp.

Jaiswal S., Dhar B. (2011). *Identification and assessment of symbiotic effectiveness of phage-typed Rhizobium leguminosarum strains on lentil (*Lens culinaris* Medik) cultivars*. *Current Microbiology*, 62(5), 1503-9.

Jiang S. y Paul J. (1998). *Gene transfer by transduction in the marine environment*. *Applied of Environmental Microbiology*, 64: 2780-2787.

Kimura M., Jia Z., Nakayama N. y Asakawa S. (2008). *Viral ecology in soil: past, present, and future perspectives*. *Soil Science and Plant Nutrition*, 54: 1–32.

Los J., Los M., Wegrzyn G. y Wegrzyn A. (2009). *Differential efficiency of induction of various lambdaoid prophages responsible of production of shiga toxins in response to different induction agents*. *Microbial Pathogenesis*, 47(6): 289-298.

Louis-Marie B., Rocha E. y Touchon M. (2012). *The adaptation of temperate bacteriophages to their host genomes*. *Molecular Biology and Evolution*, 30(4): 737-751.

Koonin E., Makarova K. y Aravind L. (2001). *Horizontal gene transfer in prokaryotes: Quantification and classification*. *Annual Review in Microbiology*, 55: 709–742.

Koonin E., Senkevich T. y Dolja V. (2006). *The ancient virus world an evolution of cells*. *Biology Direct*, 1:29.

Lawrence J. Ochman H. (1998). *Molecular archaeology of the *Escherichia coli* genome*. *Proceedings of the National Academy of Sciences of the United States of America*, 95(16): 9413-9417.

Lewin B. 2008. *Genes IX*. 9ª edición, Mc Graw Hill, Cambridge, E.U.A. 863p.

Lodish H., Berk A. y Zipursky S. (2000). *Molecular Cell Biology (sec. 6.3: Viruses; Structure, Function and Uses)*. 4ª edition, New York.

Loeffler F. y Frosch P. (1898). *Berichte der Kommission zur Erforschung der Maul- und Klauenseuche bei dem Institut für Infektionskrankheiten in Berlin*. *Zentralblatt für Bakteriologie, Parasitenkunde und Infektionskrankheiten*, 1(23): 371–391.

Lu L., Li C. y Hagedorn Ch. (2006). *Phylogenetical analysis of global hepatitis E virus sequences. Genetic diversity, subtypes and zoonosis*. *Reviews in Medical Virology*, 16(1): 5-36.

MAUVE. <http://darlinglab.org/mauve/mauve.html>

McDaniel L., Young E., Delaney J., Ruhnau F., Ritchie K. y Paul J. (2010). *High frequency of horizontal gene transfer in the oceans*. *Science*, 330(6000): 50.

MEGA. <https://www.megasoftware.net/>

Meier-Kolthoff J. y Goker M. (2017). *VICTOR: genome-based phylogeny and classification of prokaryotic viruses*. *Bioinformatics*, 33(21): 3396-3404.

Mesyanzhinov V., Leiman P., Kostyuchenko V., Kurochkina L., Miroshnikov K., Sykilinda N. and Shneider M. (2004). *Molecular architecture of bacteriophage T4*. *Biochemistry (Moscow)*, 69(11): 1190-1202.

Nasir A., Forterre P., Kim K. y Caetano-Anolles G. (2014). *The distribution and impact of viral lineages in domains of life*. *Frontiers in Microbiology*, 5: 194.

OLIGO 7.0. <https://www.oligo.net/>

Orlova E (Editor). (2012). *Bacteriophages (Chapter: "Bacteriophages and Their Structural Organisation", by Dr. Ipek Kurtboke)*. ISBN: 978-953-51-0272-4, InTech, Available from: <http://www.intechopen.com/books/bacteriophages/bacteriophages-and-their-structural-organisation>

Ohnishi M., Kurokawa K. y Hayashi T. (2001). *Diversification of Escherichia coli genomes: are bacteriophages the major contributors?*. Trends in Microbiology, 9:481-485.

Pendulla M., Ford M., Houtz J., Karthikeyan T., Wadsworth C., Lewis J., *et al.* (2003). *Origins of highly mosaic Mycobacteriophage genomes*. Cell, 113: 171-182.

PHYTON. <https://www.python.org/psf/>

Rastelli E., Corinaldesi C., Dell'Anno A., Tangherlini M., Martorelli E., Ingrassia M., Chiocci F., Martire M. y Danovaro R. (2017). *High potential for temperate viruses to drive carbon cycling in chemoautotrophy-dominated shallow-water hydrothermal vents*. Environmental Microbiology, 19(11): 4432-4446.

Rohwer F., Prangishvili D. y Lindell D. (2009). *Roles of viruses in the environment*. Environmental Microbiology, 11(11): 2771-2774.

Santamaria R., Bustos P., Sepúlveda-Robles O., Lozano L., Rodríguez C., Fernández J., *et al.* (2014). *Genomic analysis of narrow host-range bacteriophages that infect Rhizobium etli*. Applied and Environmental Microbiology, 80(2):446-454.

Semsey S., Papp I., Buzas Z., Patthy A., Orosz L. y Papp P. (1999). *Identification of site-specific recombination genes *int* and *xis* of the *Rhizobium temperate* phage 16-3*. Journal of Bacteriology, 181(14): 4185-4192.

Shwinghamer E. y Reinhardt D. (1963). *Lysogeny in Rhizobium leguminosarum and Rhizobium trifolii*. Australian Journal of Biological Sciences, 16: 597-605.

Siddelle S., Wege H. y Meulen V. (1983). *The biology of coronavirus*. The Journal of General Virology, 64: 761-776.

Simmonds P., Holmes E., Cha T., Chan S., McOmish F., Irvine B., Beall E., Yap P., Kolberg J. y Urdea M. (1993). *Classification of hepatitis c virus into six major genotypes and a series of subtypes by phylogenetic analysis of NS-5 region*. The Journal of General Virology, 74: 2391-2399.

Simmonds P., Adams M., Benko M., Breitbart M., Brister J., Carstens E., *et al.* (2017). *Virus taxonomy in the age of metagenomics*. Nature Reviews Microbiology, 15: 161-168.

Tomasz M. (1976). *H₂O₂ generation during the redox cycle of mitomycin C and DNA-bound mitomycin*. Chemico-Biological interactions, 13(1): 89-97.

Touchon M., Bernheim A. y Rocha E. (2016). *Genetic and life-history traits associated with the distribution of prophages in bacteria*. Multidisciplinary Journal of Microbial Ecology, 10: 2744-2754.

Uchiumi T., Yasuhiro O., Mikiko A. y Shiro H. (1989). *Phage induction of lysogenic Rhizobium leguminosarum biovar trifolii y both the free-living and the symbiotic form*. Journal of General Microbiology, 135: 3133-3141.

Valentine R. y **Pereira H.** (1965). *Antigens and Structure of the Adenovirus*. Journal of Molecular Biology, 13: 13-20.

Waldor M. y **Mekalanos J.** (1996). *Lysogenic conversion by a filamentous phage encoding cholera toxin*. Science, 272:1910-1914.

Werkin M., **Hans-Wolfgang A.** y **Levensque R.** (1988). *A study of 33 prophages in Rhizobium meliloti*. Applied and Environmental Microbiology, 54: 188-196.

Zhou Y., **Liang Y.**, **Karlene H.**, **Linch J.**, **Whishart D.** (2010). *PHAST: A fast phage search tool*. Nucleic acid research, 39(2): 347-352.

ANEXO

En el siguiente cuadro se recopilan todas las cepas de *Rhizobium* utilizadas en los experimentos. En los resultados y discusión se hace referencia a ellas, generalmente con el “Nombre de la cepa”.

Cepas del genero <i>Rhizobium</i>			
Numero de cepa	Nombre de cepa	Especie	Lugar de origen
1	CE3	<i>R.etli</i>	México, Gto.
2	CR652	<i>R.phaseoli</i>	Colombia
3	GR56	<i>R.phaseoli</i>	España
4	KIM5	<i>R.phaseoli</i>	México
5	BRASIL5	<i>R.phaseoli</i>	Brasil
6	8C-3	<i>R.?</i>	España
7	S20	<i>R.phaseoli</i>	
8	CIAT894	<i>R.sp.</i>	Colombia
9	MIM1	<i>R.etli</i>	Sierra de Huautla Mor.
10	IE4771	<i>R.phaseoli</i>	Puebla
11	IE4803	<i>R.phaseoli</i>	Puebla
12	IE4872	<i>R.gallicum</i>	
13	IE951	<i>R.phaseoli</i>	Puebla
14	IE954	<i>R.?</i>	
15	IE1004	<i>R.phaseoli</i>	Puebla
16	IE1006	<i>R.phaseoli</i>	Puebla
17	IE2737	<i>R.phaseoli</i>	Puebla
18	IE4777	<i>R.phaseoli</i>	Puebla
19	IE4810	<i>R.phaseoli</i>	Puebla
20	IE6766	<i>R.gallicum</i>	Michoacán
21	IE6794	<i>R.phaseoli</i>	Michoacán
22	IE6795	<i>R.?</i>	Michoacán
23	IE6833	<i>R.etli</i>	Guanajuato
24	IE6840	<i>R.etli</i>	Guanajuato
25	IE6845	<i>R.etli</i>	Guanajuato
26	IE6854	<i>R.etli</i>	Guanajuato
27	IE6868	<i>R.etli</i>	Guanajuato
28	IE6896	<i>R.phaseoli</i>	
29	IE988	<i>R.gallicum</i>	
30	IE2735	<i>R.gallicum</i>	
31	IE4770	<i>R.gallicum</i>	
32	IE4845	<i>R.gallicum</i>	
33	IE4868	<i>R.gallicum</i>	
34	R.fabae	<i>R.fabae</i>	
35	CIAT899	<i>R.tropici</i>	
36	CNPAF512	<i>R.phaseoli</i>	

37	TAL182	<i>R.etli</i>	Hawaii
38	NXC12	<i>R.etli</i>	
39	NXC14	<i>R.etli</i>	
40	NXC24	<i>R.leucaenae</i>	
41	NXT3	<i>S.americanum</i>	
42	MIM2	<i>R.phaseoli</i>	Sierra de Huautla Mor.
43	MIM5-4	<i>R.?</i>	Sierra de Huautla Mor.
44	MIM7-4	<i>R.etli</i>	Sierra de Huautla Mor.
45	INC 1-2	<i>R.phaseoli</i>	Guanajuato
46	INC 2-5	<i>R.etli/R.phaseoli</i>	Guanajuato
47	TUX10P	<i>R.phaseoli</i>	Los Tuxtlas, Veracruz
48	TUX16P	<i>R.?</i>	Los Tuxtlas, Veracruz
49	TUX1712m	<i>R.phaseoli</i>	Los Tuxtlas, Veracruz
50	TUX250P	<i>R.phaseoli</i>	Los Tuxtlas, Veracruz
51	CH24-10	<i>R.phaseoli</i>	Puebla
52	343	<i>R.etli</i>	
53	F6	<i>R.etli</i>	
54	B27	<i>R.phaseoli</i>	
55	BAT-4	<i>R.etli</i>	Zacatecas
56	BAT-32	<i>R.etli</i>	Zacatecas
57	BAT-49	<i>R.phaseoli</i>	Zacatecas
58	BAT-95	<i>R.phaseoli</i>	Zacatecas
59	D5	<i>R.etli/R.phaseoli</i>	Durango
60	D11	<i>R.etli/R.phaseoli</i>	Durango
61	D13	<i>R.etli</i>	Durango
65	D21	<i>R.etli</i>	Durango
66	PS14	<i>R.phaseoli</i>	
67	6C-1	<i>R.gallicum</i>	España
68	14C-1	<i>R.etli</i>	España
69	14C-2	<i>R.phaseoli</i>	España
70	8NJ-2	<i>R.etli</i>	España
71	16NJ-2	<i>R.?</i>	España
72	17NJ-2	<i>R.gallicum</i>	España
73	21NJ-2	<i>R.gallicum</i>	España
74	6PR-1	<i>R.phaseoli</i>	España
75	4PR-2	<i>R.?</i>	España
76	14PR-2	<i>R.?</i>	España
77	GR10	<i>R.phaseoli</i>	España
78	GR14	<i>R.etli</i>	España
79	GR62	<i>R.phaseoli</i>	España
80	GR75	<i>R.phaseoli</i>	España
81	GR64	<i>S.fredii</i>	España
82	GR84	<i>R.leguminosarum</i>	España
83	GR03	<i>R.giardini</i>	España

84	GR93	<i>R.giardini</i>	España
85	GR18	<i>R.gallicum</i>	España
86	GR42	<i>R.gallicum</i>	España
87	GR45	<i>R.gallicum</i>	España
88	GR60	<i>R.gallicum</i>	España
89	N1314 (1-N1P3TC14)	<i>R.etli</i>	Tepoztlán, Morélos
90	N741 (2-N7P4TC1)	<i>R.etli</i>	Tepoztlán, Morélos
91	N561 (3-N5P6TC1)	<i>R.etli</i>	Tepoztlán, Morélos
92	N771 (4-N7P7TC1)	<i>R.phaseoli</i>	Tepoztlán, Morélos
93	N161 (5-N1P6TC1)	<i>R.phaseoli</i>	Tepoztlán, Morélos
94	R72 (6-R7C2)	<i>R.?</i>	Tepoztlán, Morélos
95	R634 (7-R6C34)	<i>R.?</i>	Tepoztlán, Morélos
96	R744 (8-R7C44)	<i>R.phaseoli</i>	Tepoztlán, Morélos
97	R635 (9-R6C35)	<i>R.?</i>	Tepoztlán, Morélos
98	N671 (10-N6P7TC1)	<i>R.phaseoli</i>	Tepoztlán, Morélos
99	N261 (11-N2P6TC1)	<i>R.phaseoli</i>	Tepoztlán, Morélos
100	R650 (12-R6C50)	<i>R.phaseoli</i>	Tepoztlán, Morélos
101	R723 (13-R7C23)	<i>R.phaseoli</i>	Tepoztlán, Morélos
102	R611 (1-R6C11)	<i>R.phaseoli</i>	Tepoztlán, Morélos
103	R620 (2-R6C20)	<i>R.phaseoli</i>	Tepoztlán, Morélos
104	R693 (3-R6C93)	<i>R.?</i>	Tepoztlán, Morélos
105	R630 (4-R6C30)	<i>R.phaseoli</i>	Tepoztlán, Morélos
106	R711 (5-R7C11)	<i>R.?</i>	Tepoztlán, Morélos
107	N122 (6-N1P2TC2)	<i>R.?</i>	Tepoztlán, Morélos
108	N324 (7-N3P2TC4)	<i>R.?</i>	Tepoztlán, Morélos
109	N841 (8-N8P4TC1)	<i>R.phaseoli</i>	Tepoztlán, Morélos
110	N541 (9-N5P4TC1)	<i>R.?</i>	Tepoztlán, Morélos
111	N621 (10-N6P2TC1)	<i>R.etli</i>	Tepoztlán, Morélos
112	N113 (11-N1P1TC3)	<i>R.etli</i>	Tepoztlán, Morélos
113	N6212 (1*-N6P2TC12)	<i>R.etli</i>	Tepoztlán, Morélos
114	N4311 (2*-N4P3TC11)	<i>R.?</i>	Tepoztlán, Morélos
115	N831 (3*-N8P3TC1)	<i>R.phaseoli</i>	Tepoztlán, Morélos
116	N731 (4*-N7P3TC1)	<i>R.etli</i>	Tepoztlán, Morélos
117	N931 (5*-N9P3TC1)	<i>R.phaseoli</i>	Tepoztlán, Morélos
118	N941 (6*-N9P4TC1)	<i>R.?</i>	Tepoztlán, Morélos
119	N1341 (7*-N13P4TC1)	<i>R.etli</i>	Tepoztlán, Morélos
120	N871 (8*-N8P7TC1)	<i>R.etli</i>	Tepoztlán, Morélos
121	R339 (9*-RP3TC39)	<i>R.?</i>	Tepoztlán, Morélos