

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE DOCTORADO EN CIENCIAS BIOMÉDICAS

CENTRO DE CIENCIAS GENÓMICAS

PROGRAMA DE INGENIERÍA GENÓMICA

"CARACTERIZACIÓN DE UN NUEVO GRUPO DE FAGOS PROMISCUOS CAPAZ DE INFECTAR DIVERSAS GAMMAPROTEOBACTERIAS CO-AISLADAS DE COMUNIDADES DE RÍOS, EXHIBE ADAPTACIÓN DINÁMICA HACIA HOSPEDEROS DE DIFERENTES GÉNEROS".

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS

PRESENTA:

DANIEL CAZARES LÓPEZ

DIRECTOR DE TESIS:

PABLO VINUESA FLEISCHMANN

COMITÉ TUTOR:

GABRIEL GUARNEROS PEÑA

DEPARTAMENTO DE GENÉTICA Y BIOLOGÍA MOLECULAR, CINVESTAV

VÍCTOR MANUEL GONZÁLEZ ZÚÑIGA

PROGRAMA DE GENÓMICA EVOLUTIVA, CENTRO DE CIENCIAS GEÓMICAS

CUERNAVACA, MORELOS, MÉXICO, JUNIO DE 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Este trabajo se realizó bajo la dirección del Dr. Pablo Vinuesa Fleischmann en el laboratorio del programa de Ingeniería Genómica del Centro de Ciencias Genómicas de la Universidad Nacional Autónoma de México.

Agradecimientos

A la Universidad Nacional Autónoma de México por brindarme la invaluable oportunidad de desarrollarme profesionalmente de manera laica y gratuita, abrazándome e iluminando en sus prestigiosas aulas desde el nivel medio superior hasta el doctorado. Para mi es motivo de dicha y orgullo haber crecido en esta institución que considero única en el mundo, y ser parte de su gran comunidad.

Al Programa de Doctorado en Ciencias Biomédicas y al Centro de Ciencias Genómicas por haberme aceptado y proporcionado todo lo que necesitaba para consolidar mi formación científica.

Al CONACYT por haberme apoyado con una beca durante mi doctorado, la cual fue el sostén y estímulo necesario para fomentar mi desarrollo científico, permitiendo que me enfocara de tiempo completo en mis estudios y en la ciencia que tanto me apasiona.

A los proyectos DGAPA-PAPIIT/UNAM IN206318 and CONACYT 23_A1-S-11242 por haber proporcionado el financiamiento para la realización de este estudio.

Al doctor Pablo Vinuesa Fleischmann por haberme concedido el honor de ser su alumno y formar parte de su laboratorio y equipo de trabajo. Gracias por haber roto los moldes del rol de tutor y darme la confianza para tratarnos como amigos. Te estaré siempre agradecido por haber depositado tu confianza en mí y darme la apreciada oportunidad de desarrollarme siguiendo mis propias ideas, hipótesis y de aprender de mis propios errores, gracias por tu paciencia y por haberme apoyado en todo momento, aunque a veces fallará. Te doy las gracias por todas tus enseñanzas y por haberme inspirado en mi carrera científica, gracias por haberme guiado en los momentos difíciles y por compartirme tus experiencias de vida, así como por escucharme y alentarme a seguir adelante, gracias por todo.

A los doctores Gabriel Guarneros Peña y Víctor González Zúñiga por haber aceptado ser los miembros de mi comité tutor, gracias por su constante y valioso feedback y por acompañarme en este largo camino guiándome con sus sabios consejos.

A los doctores David Romero Camarena, Susana Brom Klanner y Alejandro García de los Santos por su constante asesoramiento y apoyo incondicional, gracias por su cordialidad y haberme hecho sentir en familia.

A los técnicos del laboratorio Javier Rivera Campos y Laura Cervantes de la Luz por su soporte técnico y asistencia con los experimentos.

A Denny Peralta Luna por haberme auxiliado y asesorado con todos los trámites durante mi doctorado.

A todo el personal y amigos del Programa de Ingeniería Genómica, gracias por haber compartido tantas experiencias y alegrías conmigo, y haber sido parte de mi vida en esta etapa tan decisiva de mi carrera.

A todos los miembros de mi comité sinodal: la doctora Gloria Soberón Chávez, el doctor Luis David Alcaraz Peraza, la doctora Blanca Itzel Taboada Ramirez, el doctor José Luis Puente Garcia y el doctor Pablo Vinuesa Fleischmann. Gracias por haber destinado parte de su valioso tiempo para revisar esta tesis y contribuir a mejorar este trabajo.

A mi madre por haber estado ahí en cada paso que he dado en la vida, procurando en todo momento que tuviera éxito en mis planes. Gracias por haberme apoyado incondicionalmente en cada decisión importante de mi vida. Gracias por todo tu cariño, comprensión y enseñanzas. Mi eterna gratitud hacia ti madre por haber dedicado todos tus esfuerzos y destinado tus ilusiones para que mi hermano y yo tuviéramos un por venir. Gracias por haber hecho posible que nuestros sueños fueran alcanzables, te lo debemos todo y sin ti no seríamos nada.

A mi padre por todo el afecto y amor que me brindo, gracias por cada momento que compartiste conmigo, siempre te llevaré vivo en mi memoria.

A Wendy que es como una hermana para mí, gracias por toda tu ayuda y apoyo incondicional, gracias por compartirme tu experiencia y siempre aconsejarme, nunca dejo de aprender de ti; eres una persona admirable, capaz de lograr cualquier cosa que se proponga, eres un gran ejemplo a seguir, muchas gracias por permitirme ser parte de tu vida.

A mi hermano que ha sido como un segundo padre para mí, así también como un maestro y mejor amigo, gracias por protegerme, guiarme, motivarme y por compartir sueños y aspiraciones conmigo. Me siento muy afortunado de poder seguir compartiendo experiencias contigo, de poder seguir aprendiendo a tu lado, siempre serás mi modelo a seguir, gracias por todo querido hermano.

Índice general

Índice de figuras	7
Índice de tablas	8
Resumen	9
Importancia	10
Introducción	11
Relevancia ecológica de los fagos y de sus interacciones con bacterias	11
El rango de hospedero y su implicación en las interacciones fago-bacteria	13
Objetivos	15
General	16
> Caracterizar de la dinámica de infección de fagos promiscuos	16
Particulares	16
> Generar una colección de aislados bacterianos ambientales susceptibles a la infección por los fagos coexistentes	16
> Aislar fagos promiscuos y caracterizar su eficiencia de infección a través de un panel diverso de hospederos	16
> Caracterizar los fagos promiscuos aislados mediante microscopía electrónica de transmisión	16
> Secuenciar y hacer un análisis genómico de los fagos promiscuos	16
> Explorar la evolución de la eficiencia de infección de los fagos promiscuos a través de hospederos de diferentes géneros	16
Resultados	17
Aislamiento de fagos promiscuos	17
Eficiencia de plaqueo de los fagos Atoyac	22
Morfología del virión y comparación genómica.	24
Búsqueda de fagos Atoyac-like en datos metagenómicos	27
Evolución experimental de la eficiencia de plaqueo	28
Discusión	32
Conclusiones	38
Perspectivas	38
Materiales y Métodos	39
Muestreo y conteo de coliformes	39
Aislamientos bacterianos	39
Aislamiento de fagos	39
Identificación taxonómica de los aislados bacterianos a nivel de género	40
Purificación de los fagos promiscuos	41
Determinación del rango de hospedero y eficiencia de plaqueo	41
Análisis estadístico	42

Determinación de la morfología del virión	42
Extracción del DNA de los fagos y secuenciación genómica	42
Análisis de los genomas de los fagos	42
Predicción in silico de promotores	43
Búsqueda de los genomas de los fagos Atoyac en datos metagenómicos	43
Evolución experimental de la eficiencia de plaqueo	43
Disponibilidad de los datos y material biológico	44
Publicación de los resultados de este estudio	44
Material suplementario	45
Figuras suplementarias	45
Tablas suplementarias	48
Sets de datos suplementarios descargables	49
Notas	49
Referencias	50
Anexos	59

Índice de figuras

Figura 1	
Diagrama de la estrategia empleada para el aislamiento de fagos promiscuos	11
Figura 2	
Lugares de aislamiento de los fagos Atoyac	13
Figura 3	
Aislados bacterianos susceptibles a la infección por los fagos promiscuos Atoyac	14
Figura 4	
Placas de los fagos Atoyac en cepas representativas de los diferentes grupos taxonómicos susceptibles a su infección	175
Figura 5	
Rango de hospedero y eficiencia de plaqueo de los fagos Atoyac	17
Figura 6	
Micrografías de los viriones de los fagos del grupo Atoyac	19
Figura 7	
Análisis genómico comparativo y filogenético del grupo de fagos Atoyac	20
Figura 8	
Diagrama de la estrategia usada para los ensayos de evolución experimental	23
Figura 9	
Evolución experimental de la eficiencia de plaqueo de dos fagos representativos del grupo Atoyac	25
Figura Suplementaria 1	
Comparación genómica de los fagos Atoyac y sus principales regiones variables del grupo Atoyac	38
Figura Suplementaria 2	
Evidencia de la pureza de los fagos Atoyac	39
Figura Suplementaria 3	
Comparación de la morfología de placa de los fagos Atoyac1 y Atoyac23 a través de los experimentos de evolución experimental	490

Índice de tablas

Tabla Suplementaria 1	
Sitios de muestreo y lugares de aislamiento de los fagos Atoyac	11
Tabla Suplementaria 2	
Parámetros ambientales de los sitios de muestreo	13
Table S1 (link al archivo descargable)	
Predicción in silico de promotores en la región intergénica de los fagos Atoyac	14

Sets de datos suplementarios descargables

Data Set S2 (link al archivo descargable)	
Medios y condiciones usadas en los ensayos de plaqueo. Títulos y EOPs de los fagos Atoyac en los aislados susceptibles. Test de Mann-Whitney U de las diferencias de las EOPs entre grupos taxonómicos y subgrupos para cada fago	175
Data Set S3 (link al archivo descargable)	
Mapeo de lecturas de metagenomas a los genomas del fago crAss (crAssphage) y del fago T4	17
Data Set S4 (link al archivo descargable)	
EOPs de los ensayos de evolución experimental y test de ANOVA	19

Resumen

Los fagos son generalmente descritos como especie específicos o incluso cepa específicos, implicando una inherente limitación para que algunos puedan ser mantenidos en las comunidades microbianas diversas. Sin embargo, el aislamiento de fagos y determinación del rango de hospedero raramente considera el contexto ecológico de los fagos, sesgando probablemente nuestra noción sobre la especificidad de los fagos. En este trabajo aislamos y caracterizamos un nuevo grupo de seis fagos promiscuos, nombrados Atoyac, presentes en ríos y aguas residuales mediante el uso de una colección diversa de más de 600 aislados bacterianos recuperados a partir de los mismos ambientes como potenciales hospederos. Estos podofagos aislados a partir de diferentes regiones en México mostraron un rango de hospedero notoriamente amplio, infectando bacterias de seis géneros de Gammaproteobacteria: *Aeromonas*, *Pseudomonas*, *Yersinia*, *Hafnia*, *Escherichia* y *Serratia*. Los genomas de los fagos Atoyac presentan una longitud ~42 kb y una alta similitud genómica entre ellos, pero no con fagos actualmente disponibles en bases de datos de genomas y metagenomas públicos. Una comparación detallada de la eficiencia de plaqueo (EOP) de los fagos reveló variación entre los géneros bacterianos, implicando un costo asociado con la infección de hospederos distantemente relacionados, y entre fagos, a pesar de su similitud de secuencia. Mostramos a través de evolución experimental usando un solo hospedero o alternando hospederos de diferentes géneros, que la eficiencia de la producción de placas es altamente dinámica y tiende hacia la optimización en hospederos que presentan baja formación de placas. No obstante, la adaptación a distintos hospederos difirió entre fagos fagos similares, mientras un fago optimizó su EOP en todos los hospederos probados, el otro redujo su producción de placas en un hospedero, sugiriendo que la propagación en múltiples bacterias puede ser clave para mantener el rango de hospedero promiscuo en algunos virus. Nuestro estudio expande nuestro conocimiento sobre la virosfera y revela las posibles interacciones fago-bacteria que normalmente están ocultas en los sistemas naturales.

Importancia

En los ambientes naturales, los fagos coexisten e interactúan con una amplia variedad de bacterias, planteando un enigma para el mantenimiento de los fagos de rango estrecho de hospedero en las comunidades microbianas diversas. Este contexto es raramente considerado en el estudio de las interacciones fago-bacteria, típicamente enfocadas en virus de rango de infección estrecho y en su infectividad sobre bacterias blanco provenientes de fuentes distintas de donde se aislaron los fagos. Mediante el estudio de las interacciones fago-hospedero en bacterias y virus aisladas de comunidades microbianas de ríos, nosotros mostramos que nuevos fagos con rango de hospedero amplio, abarcando múltiples géneros bacterianos, pueden ser encontrados en el ambiente. La evaluación de cientos de interacciones en diversos hospederos revelaron que fagos similares exhiben diferentes eficiencias de infección y patrones de adaptación. Comprender el rango de hospedero es fundamental para nuestro conocimiento de las interacciones fago-bacteria y su impacto en las comunidades microbianas. La dinámica natural de la promiscuidad de los fagos revelada en nuestro estudio tiene implicaciones sobre diferentes aspectos de la investigación de fagos, tales como la transferencia horizontal de genes o la terapia con fagos.

Introducción

Relevancia ecológica de los fagos y de sus interacciones con bacterias

A pesar de pasar comúnmente desapercibidos por su tamaño ultramicroscópico, los virus son las entidades biológicas dominantes en el planeta por su inmensa abundancia y diversidad. Entre ellos, destacan los virus que infectan y se replican en bacterias, conocidos comúnmente como bacteriófagos o fagos [1, 2]. A nivel global se estima que la abundancia de los fagos es 10^{31} (para tener una idea de su magnitud, esta cantidad es 1 000,000,000 de veces superior al número de galaxias que se estima existen en el universo). Su distribución abarca todo tipo de ambientes, desde los marinos, los de agua dulce, los terrestres (incluyendo el permafrost), los extremófilos y las microbiotas asociadas a las plantas y a los animales, incluida la del humano [3–6]. Se consideran por tanto ubicuos, y están presentes en cualquier lugar donde exista la vida microbiana. La concentración de los fagos suele variar entre los diversos ambientes, por ejemplo, en los acuáticos fluctúa entre 1 millón y 10 millones por mililitro, en los terrestres entre 1 millón y 1000 millones por gramo de suelo, mientras que en microbiotas como la del intestino humano se estima en un rango de entre 100 millones y 10 000 millones por gramo de materia fecal, etc. No obstante, sin importar el ambiente, la concentración de los fagos suele ser muy alta, y generalmente supera al menos 10 veces a la de las bacterias [3–7].

Como consecuencia de su gran abundancia, es reconocido que los fagos son la mayor fuente de diversidad genética en el mundo, representando un importante reservorio de genes movilizable y que es clave en la adaptabilidad de sus hospederos a nuevos ambientes [8–12]. Es de destacar, que los fagos también son una importante fuente de genes con potencial biotecnológico; como prueba de ello es que muchas de nuestras actuales herramientas moleculares de uso rutinario en el laboratorio, tales como: vectores, polimerasas, enzimas de restricción, promotores, etc., provienen de los fagos [13]. En conjunto, fagos y bacterias juegan un rol ecológico trascendental para el planeta, ya que sus interacciones impulsan desde los ciclos biogeoquímicos más importantes hasta la evolución microbiana. Por ejemplo, se estima que tan sólo en los océanos, ocurren 10^{23} infecciones de fagos por segundo, las cuales provocan la lisis de entre el 20 y 40% de las poblaciones bacterianas diariamente, permitiendo así la liberación y bio-disponibilidad de una inmensa cantidad de nutrientes para ser reutilizados por otros organismos [3, 8, 12, 14]. Mediante este proceso ecológico de regulación de la densidad poblacional bacteriana, los fagos promueven, tanto el equilibrio dentro los ecosistemas microbianos, como el mantenimiento de los principales ciclos de los nutrientes, tales como el del carbono y el del fósforo [3, 8, 12].

Las relaciones de los fagos con las bacterias se caracterizan por ser sumamente complejas, pues además de ser sus principales depredadores [14], paradójicamente, también suelen fungir como sus principales aliados, ya sea proporcionándoles inmunidad contra la infección por otros fagos (derivado de la homoinmunidad adquirida tras el proceso de lisogenización por fagos de tipo temperados) y/o la adquisición de nuevo material genético con potencial valor adaptativo, siendo así unos de los principales conductores de diversidad bacteriana [9, 12, 15, 16]. La transducción, mecanismo por el cual los fagos actúan como vectores de transferencia horizontal genética (Horizontal Gene Transfer, en inglés), es considerada el principal motor evolutivo bacteriano [16–

18], pues convierte a los fagos en plataformas versátiles capaces de transferir desde algunos cuantos genes hasta grandes cantidades de DNA entre bacterias [19]. Este mecanismo puede presentarse en dos diferentes variantes, la especializada o la generalizada; no obstante, ambas derivan de errores en el proceso de empaquetamiento del DNA del fago [20, 21]. En la transducción especializada, el fago (de tipo temperado) además de empaquetar su DNA, añade a su genoma algunos genes bacterianos adyacentes a su sitio de integración (el fago Lambda es un ejemplo de aquellos que presentan transducción de tipo especializada). Mientras que, en la transducción generalizada, el fago (pudiendo ser lítico o temperado) empaqueta exclusivamente DNA bacteriano, ya sea cromosomal o proveniente de plásmidos [22], generando así partículas virales defectivas capaces de movilizar desde algunas cuantas kilobases hasta decenas o centenas de kilobases de información genética (dependiendo del tamaño del fago) hacia otras bacterias receptoras, las cuales pueden incorporar este nuevo material genético mediante mecanismos de recombinación [16, 17, 21]. Por lo anterior, se considera que la transducción generalizada es uno de los procesos que más contribuye a la evolución bacteriana [17], pues además de las grandes cantidades y diversidad de DNA que puede movilizar, se estima que ocurre a mayores frecuencias que otros procesos de transferencia horizontal genética, con una frecuencia global de 20 mil billones de eventos cada segundo [12]. A pesar de su relevancia, evidenciar dicho fenómeno aún representa un importante desafío para la biología molecular, principalmente a gran escala y en modelos que asemejen el nivel de complejidad de las comunidades microbianas, lo cual en parte es debido a que por la naturaleza del proceso no es fácil rastrear el contenido genético de las partículas transductoras [23].

La participación de la transducción en la virulencia bacteriana es representativa del rol de los fagos en la potenciación del fitness de sus hospederos, pues ha sido documentado ampliamente que los fagos portan importantes toxinas que intensifican la virulencia de ciertas bacterias patógenas, fenómeno conocido como conversión lisogénica [18]. Casos referentes son el fago CTX Φ que codifica la toxina colérica (ctxAB) en el patógeno *Vibrio cholerae* [24], así como una serie de fagos (H-19B, VT1-Sakai, VT2-Sakai, 933W, Φ 3538) que acarrean la toxina Shiga (Stx) en los enteropatógenos *Shigella* y *Escherichia coli*, confiriéndoles el fenotipo enterohemorrágico [10, 18, 25, 26]. Similarmente, la potente toxina botulínica (*Botulinum toxin*) de *Clostridium botulinum* es codificada en sus diferentes versiones dentro de los genomas de varios fagos (c-st, CE β , d16 Φ) [18, 26]. No obstante, también se ha documentado ampliamente que los fagos ayudan a movilizar genes de resistencia a antibióticos [27–29]. Estos ejemplos ilustran la relevancia de los fagos y de la transducción, que junto a los plásmidos y la conjugación, promueven la evolución de bacterias de importancia clínica [19]. Así mismo, se ha demostrado que los fagos portan importantes genes para el metabolismo de sus hospederos, desde aquellos relacionados al proceso de fotosíntesis [30], hasta los involucrados en la comunicación bacteriana por medio del sistema de “quorum sensing” [31], e incluso aquellos que no parecen tener un valor adaptativo evidente para sus hospederos, como los genes del rRNA 16S [23], por lo cual es evidente que cualquier tipo de gen es movilizable por medio de la transducción. Sin embargo, se ha demostrado que, ante situaciones de estrés bacteriano por cambios en el ambiente, los fagos pueden fungir como aliados valiosos, ya que ayudan a movilizar genes que promueven la adaptabilidad y supervivencia de sus hospederos frente a la adversidad, haciendo a la comunidad más resiliente [17, 32, 33]. Por estos motivos, se ha sugerido que la relación

entre los fagos y las bacterias puede asemejarse en parte a aquellas relaciones de tipo mutualista, en las que los organismos involucrados procuran el beneficio mutuo.

El rango de hospedero y su implicación en las interacciones fago-bacteria

Diversos estudios se han enfocado en tratar de describir las interacciones fago-bacteria, proponiendo modelos generales que ayuden a comprender estas complejas relaciones biológicas. Las cuales se caracterizan por ser altamente dinámicas [3, 9], ya que a través de la implementación de medidas y contramedidas antagonistas, fagos y bacterias promueven su propia coevolución (dicha competencia comúnmente es referida como carrera armamentista o *arm race*) [34, 35]. No obstante, para aspirar a comprender dicha dinámica, resulta primordial analizar el rango de hospedero de los fagos, así como estudiar su evolución. Definido como la diversidad de hospederos bacterianos que un fago es capaz de infectar [36], el rango de hospedero en gran medida delimita las posibles interacciones entre fagos y bacterias de un ambiente determinado, así como el alcance de posibles eventos de transducción y sus consecuencias evolutivas.

La caracterización del rango de hospedero es un paso fundamental hacia la evaluación de la especificidad de los fagos y la identificación de especímenes con un amplio alcance de hospederos susceptibles. Así mismo, las estrategias usadas para definir una “infección exitosa” son clave para caracterizar este rasgo. Para poder infectar, los fagos deben adsorberse a la célula hospedera, transportar sus genomas dentro de la célula, superar los mecanismos de defensa bacteriana y secuestrar la maquinaria molecular del hospedero para replicar sus genomas, ensamblar y liberar la nueva progenie viral [36, 37]. Ensayos de infección usando cultivos en medio líquido (conocidos como curvas de lisis), goteo de lisados (*spotting*) o por ensayos de plaqueo (*phage plaquing*) representan métodos comúnmente usados para identificar infección por fagos [38]. Sin embargo, debido a que otros factores no asociados a fagos (por ejemplo, las bacteriocinas) y mecanismos independientes de la producción de fagos (por ejemplo, la *lysis from without*, que es la lisis bacteriana no asociada a una infección y derivada de la presencia de altas concentraciones extracelulares de fagos o colas de fagos) pueden conducir a la muerte celular por lisis. El aislamiento de placas líticas individuales, conocido como plaqueo, se considera una aproximación confiable para identificar infecciones productivas por fagos (es decir, ciclos de infección completos que derivan en la producción de la nueva progenie viral). Adicionalmente, la caracterización del rango de hospedero productivo mediante ensayos de plaqueo tienen la ventaja de permitir la determinación de la eficiencia de plaqueo de los fagos para comparar sus capacidades de infección a través de un panel de hospederos susceptibles. Por estos motivos, en este estudio favorecimos el uso del método de plaqueo para estimar el rango de hospedero de los fagos. Aún así, vale la pena señalar que los ensayos de plaqueo son susceptibles de generar resultados falsos negativos en casos donde la infección por fagos no conduce a la formación de placas debido a un reducido vigor de la infección o por infecciones de tipo crónicas que no derivan en la muerte celular, conduciendo a una subestimación de este rasgo [38].

De manera histórica y sistemática, los fagos han sido descritos como virus especie-específicos e incluso cepa-específicos [38–40]. Sin embargo, en los ambientes naturales los fagos coexisten e interactúan con una amplia diversidad de hospederos. Lo cual plantea un enigma para su mantenimiento en el contexto ecológico de las comunidades microbianas taxonómicamente diversas, pues la premisa sobre su rango de hospedero estrecho intrínseco implicaría una severa limitación para que algunos de ellos puedan encontrar hospederos receptivos y logren su propagación y supervivencia [41]. Alternativamente, ambientes caracterizados por poseer una amplia diversidad microbiana podrían promover la existencia de fagos con rangos de hospedero más amplios, permitiendo que en este contexto algunos fagos se conduzcan con mayor flexibilidad e interactúen con múltiples miembros de la comunidad.

Los estudios de las interacciones fago-bacteria típicamente se han centrado en el aislamiento de fagos de rango estrecho, principalmente en aquellos que son infectivos de bacterias patógenas blanco, con miras a su uso potencial en la terapia con fagos [39, 40]. En dichos estudios el rango de hospedero suele analizarse a partir del uso de un número reducido de cepas potencialmente susceptibles, usualmente de una sola especie y las cuales no guardan una relación ecológica con los fagos aislados (proviendo de ambientes de aislamiento distintos). Esto evidencia que el contexto ecológico no es comúnmente considerado en los protocolos de aislamiento de fagos, ni en el estudio de sus interacciones. Lo cual podría estar sesgando nuestra noción acerca de la especificidad de los fagos. Aun cuando a la fecha, diversos fagos capaces de infectar más allá del rango taxonómico de género han sido descritos, estos son considerados raros en comparación con la inmensa mayoría que han sido reportados como fagos específicos de especie o de género [38–40]. No obstante, el aislamiento y estudio de fagos con la habilidad de infectar múltiples especies y géneros puede contribuir significativamente a avanzar nuestro nivel de comprensión sobre la ecología de los fagos. Estos representan a su vez poderosos modelos que permitirían estudiar la evolución del rango de hospedero, así como la transducción entre especies distantemente relacionadas, la cual podría tener efectos evolutivos profundos. Recientemente, estudios metagenómicos han proporcionado información valiosa acerca de la distribución de fagos con amplio rango de hospedero, sugiriendo que ellos pueden ser más abundantes de lo postulado previamente [36, 42, 43]. Sin embargo, la naturaleza de este enfoque restringe la caracterización experimental de este tipo de virus, los cuales podrían tener como hospederos a bacterias no cultivables o que aún no han sido aisladas. Por otro lado, algunos estudios pioneros se han enfocado en el aislamiento de fagos de amplio rango mediante procedimientos de enriquecimiento, ya sea usando co-cultivos [44] o monocultivos secuenciales [45] de potenciales hospederos bacterianos con diferente origen taxonómico. Estos enfoques condujeron independientemente al aislamiento de diferentes fagos capaces de infectar múltiples especies, tales como *Pseudomonas aeruginosa* y *Escherichia coli*.

En este estudio investigamos la presencia de fagos de amplio rango de hospedero en escenarios ambientales, debido a su relevancia para la implementación de novedosos modelos para el estudio de las interacciones fago-bacteria entre diferentes especies y la evolución del rango de hospedero. Así como por su posible trascendencia en eventos de transducción intergenérica, los cuales podrían tener consecuencias evolutivas profundas. Además de por su potencial utilidad en la terapia con fagos, ya que este tipo de fagos o sus sistemas de lisis podría aumentar la versatilidad de los tratamientos, mejorando su eficacia para eliminar a más de un patógeno blanco de manera similar a como lo harían los antibióticos de amplio espectro. Usando como estrategia el aislamiento de una amplia colección de bacterias y fagos recuperados a partir de muestras de ríos y aguas residuales, se logró la identificación de fagos de amplio rango de hospedero a partir del estudio de sus interacciones. Nosotros hipotetizamos que usando un conjunto taxonómicamente diverso de bacterias nativas de estos ambientes como potenciales hospederos podría facilitar el aislamiento de fagos capaces de infectar múltiples especies. El uso de este enfoque condujo al aislamiento de un grupo de fagos estrechamente relacionados pertenecientes a la familia *Podoviridae*, los cuales fueron capaces de infectar productivamente seis diferentes géneros bacterianos y tres órdenes taxonómicos dentro de la clase Gammaproteobacteria. Estos fagos, nombrados Atoyac (río en Náhuatl) y que aquí referimos como “promiscuos” por su amplio rango de hospedero, representan un nuevo grupo viral, con homólogos distintamente relacionados y sin presencia detectable en bases de datos metagenómicos públicas. Los fagos Atoyac mostraron diferentes eficiencias de plaqueo entre los diversos géneros susceptibles, sugiriendo un costo biológico asociado a la infección de bacterias distantemente relacionadas. Sin embargo, mediante el uso de evolución experimental para explorar el potencial adaptativo de su eficiencia de plaqueo, se reveló que tales eficiencias cambian rápidamente tras su propagación sucesiva en un solo fondo genómico o en múltiples, mostrando una adaptación dinámica caracterizada por una tendencia hacia la optimización en hospederos poco eficientes. Nuestros hallazgos contribuyen a expandir nuestro conocimiento acerca de la ecología de los fagos y las interacciones fago-bacteria ocultas en los sistemas naturales, proporcionando nuevos modelos y perspectivas en experimentos de evolución a corto plazo usando múltiples hospederos diversos.

Objetivos

General

- ❖ Caracterizar la dinámica de infección de fagos promiscuos

Particulares

- Generar una colección de aislados bacterianos ambientales susceptibles a la infección por los fagos coexistentes
- Aislar fagos promiscuos y caracterizar su eficiencia de infección a través de un panel diverso de hospederos
- Caracterizar los fagos promiscuos aislados mediante microscopía electrónica de transmisión
- Secuenciar y hacer un análisis genómico de los fagos promiscuos
- Explorar la evolución de la eficiencia de infección de los fagos promiscuos a través de hospederos de diferentes géneros

Resultados

Estrategia para el aislamiento de fagos promiscuos

Nosotros dirigimos la búsqueda de fagos promiscuos, es decir, de fagos capaces de infectar hospederos de distinto origen taxonómico, mediante el ensamble de una amplia y diversa colección de más de 600 aislados bacterianos y fagos libres presentes en muestras de agua de ríos y aguas residuales de doce sitios diferentes en México (Figura 1, ver detalles en Materiales y Métodos). El goteo de alícuotas y diluciones seriadas de las muestras de agua sobre los tapices de los aislados bacterianos recuperados a partir de las mismas muestras, produjo en muchas de las ocasiones la formación de spots de lisis y placas líticas. El uso de medios de cultivo estándar tales como LB o T Φ no siempre permitieron la visualización de placas de fagos, sin embargo, la utilización de diversos medios y la adición de diferentes colorantes a los mismos, condujeron a la identificación de placas líticas (ver detalles en Materiales y Métodos). Un screening de ensayos de infección cruzada entre los aislados susceptibles, permitió la identificación de fagos capaces de producir placas en múltiples cepas (Figura 1, ver detalles en Materiales y Métodos).

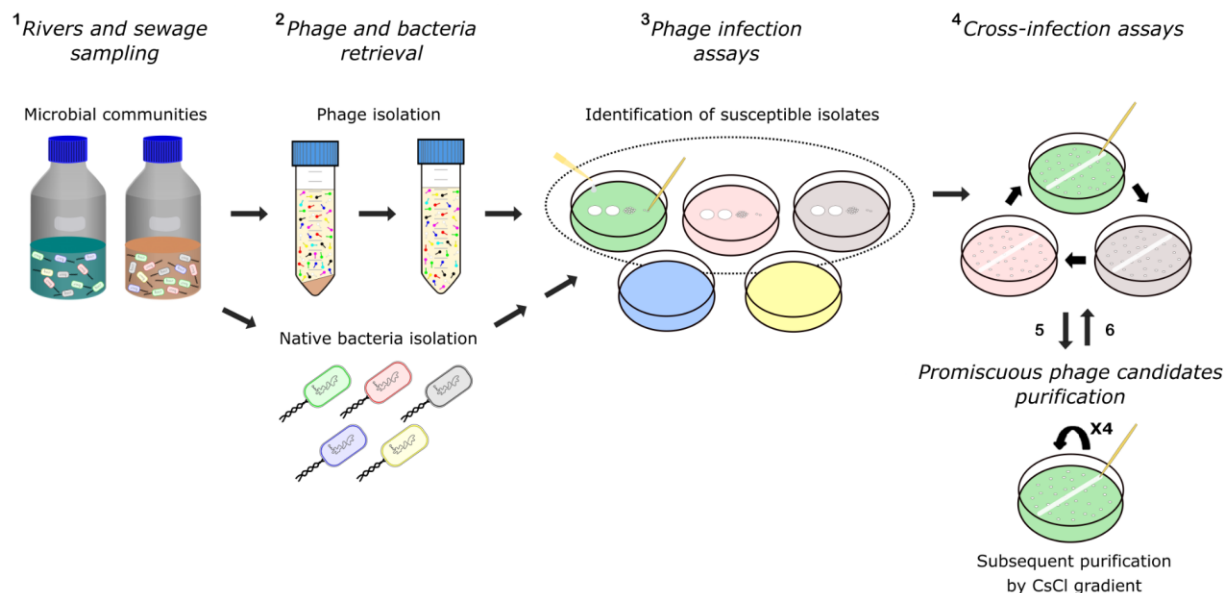


Figura 1. Diagrama de la estrategia empleada para el aislamiento de fagos promiscuos. 1. Diversas muestras de agua con diferentes niveles de contaminación fueron colectadas. 2. Los fagos libres de estas comunidades microbianas fueron recuperados mediante la centrifugación y filtración de las muestras. Simultáneamente, bacterias de estas comunidades fueron aisladas para generar una amplia y diversa colección de bacterias nativas. 3. La infectividad de los fagos presentes en las muestras sobre los aislados nativos fue evaluada, y los aislados susceptibles fueron identificados taxonómicamente. 4. Ensayos de infección cruzada; que consistieron en la selección de placas líticas individuales y su transferencia entre los tapices de los aislados bacterianos identificados como susceptibles en el paso anterior, fueron realizados para evaluar el rango de hospedero de los fagos (Notar que la transferencia de las placas de fagos sobre los tapices de los aislados susceptibles fue realizada tanto de manera simultánea así como secuencial). 5. Los candidatos a fagos promiscuos; correspondiendo a aquellos que podían infectar múltiples aislados, fueron propagados usando un solo hospedero (4 veces) y luego purificados por ultracentrifugación en gradientes de cloruro de cesio (CsCl). 6. Finalmente, los fagos purificados fueron

sometidos nuevamente a ensayos de infección cruzada para validar su amplio rango de hospedero (consultar sección Materiales y Métodos para más detalles).

De manera notable, seis fagos recuperados a partir de cinco de los sitios muestreados fueron capaces de infectar a 54 aislados bacterianos en el pánel (Figura 2 y Figura 3). Estos fagos, los cuales nombramos Atoyac, fueron seleccionados para posteriores análisis en este estudio. Las bacterias susceptibles incluían a miembros de los géneros *Aeromonas*, *Pseudomonas*, *Hafnia*, *Escherichia*, *Serratia* y *Yersinia*, confirmando la naturaleza promiscua del rango de hospedero de los fagos Atoyac, el cual cruza múltiples rangos taxonómicos dentro de la clase Gammaproteobacteria (Figura 3 y Figura 4). La morfología de placa de los fagos fue diversa a través de los distintos géneros susceptibles, siendo por lo general más grandes sobre los aislados de *Aeromonas* y de menor tamaño en los grupos de *Enterobacteriaceae*. Entre los fagos Atoyac las morfologías fueron relativamente similares, no obstante, algunas diferencias fueron apreciables y consistentemente reproducibles. Por ejemplo, los fagos Atoyac 1 y 14 produjeron las placas más grandes sobre el aislado de *Pseudomonas*, en contraste, los fagos Atoyac 15 y 23 produjeron placas diminutas sobre este aislado. Por otro lado, los fagos Atoyac 23 y 13 regularmente producían placas más grandes sobre los aislados de *Enterobacteriaceae* que las de los otros fagos, mientras que el fago Atoyac 15 en muchas ocasiones presentaba las placas de mayor tamaño sobre los aislados de *Aeromonas* (Figura 4). Los fagos Atoyac fueron aislados a partir de muestras provenientes de las regiones: central (Atoyac1, Atoyac10, Atoyac13 y Atoyac14), sureste (Atoyac23) y noroeste (Atoyac15) de México (Figura 4). La evaluación de la contaminación fecal de las muestras de agua mediante conteo de coliformes fecales o presencia del fago *crAssphage* (fago asociado a la microbiota intestinal humana comúnmente usado como marcador de contaminación fecal del agua [46]), reveló que los fagos Atoyac eran prevalentes en los sitios contaminados, los cuales representaron a muchas de nuestras muestras (Tabla S1). Dos fagos Atoyac fueron aislados a partir de aguas residuales, mientras que los otros cuatro eran provenientes de ríos contaminados con descargas residuales. Los sitios contaminados también presentaron el mayor título de fagos en los ensayos de plaqueo.

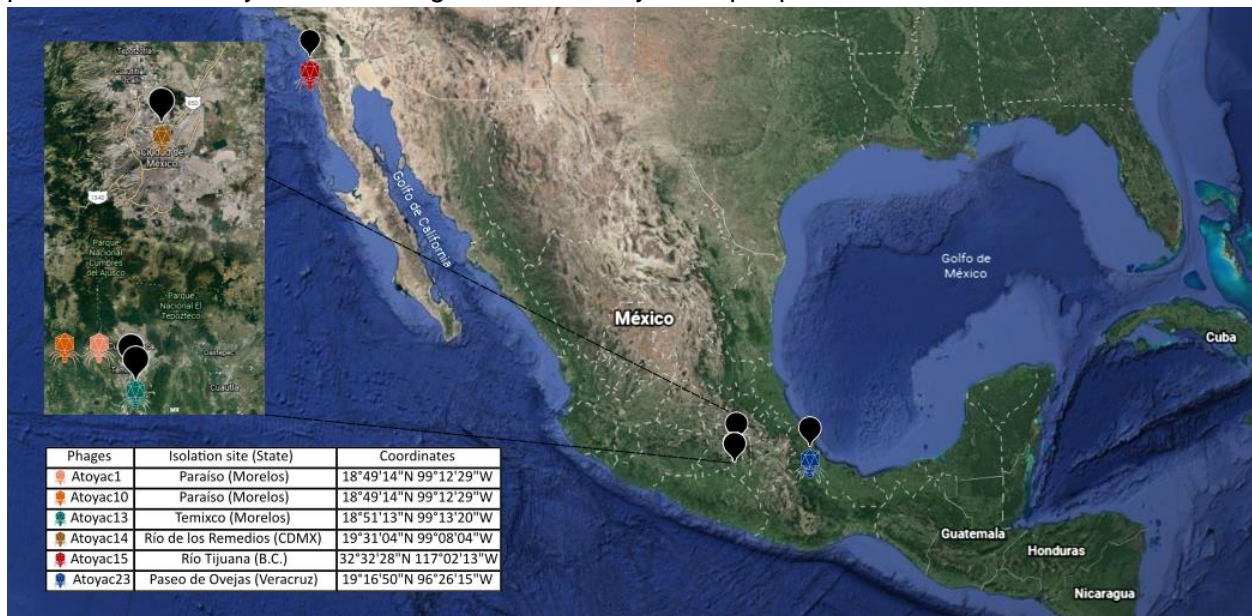


Figura 2. Lugares de aislamiento de los fagos Atoyac. La ubicación geográfica de los 6 fagos Atoyac aislados a partir de los sitios contaminados es indicada en el mapa usando un código de colores sobre los *pins* con forma de fagos. Cuatro de los fagos fueron recuperados a partir de muestras agua provenientes de los estados de Morelos y la Ciudad de México (fagos Atoyac1, Atoyac10, Atoyac13 y Atoyac14, respectivamente) ubicados en la región central de México (proyección del mapa con acercamiento), uno a partir de una muestra del estado de Veracruz (Atoyac23) ubicado en la región sureste y uno a partir de una muestra del estado de Baja California (Atoyac15) ubicado en la noroeste.

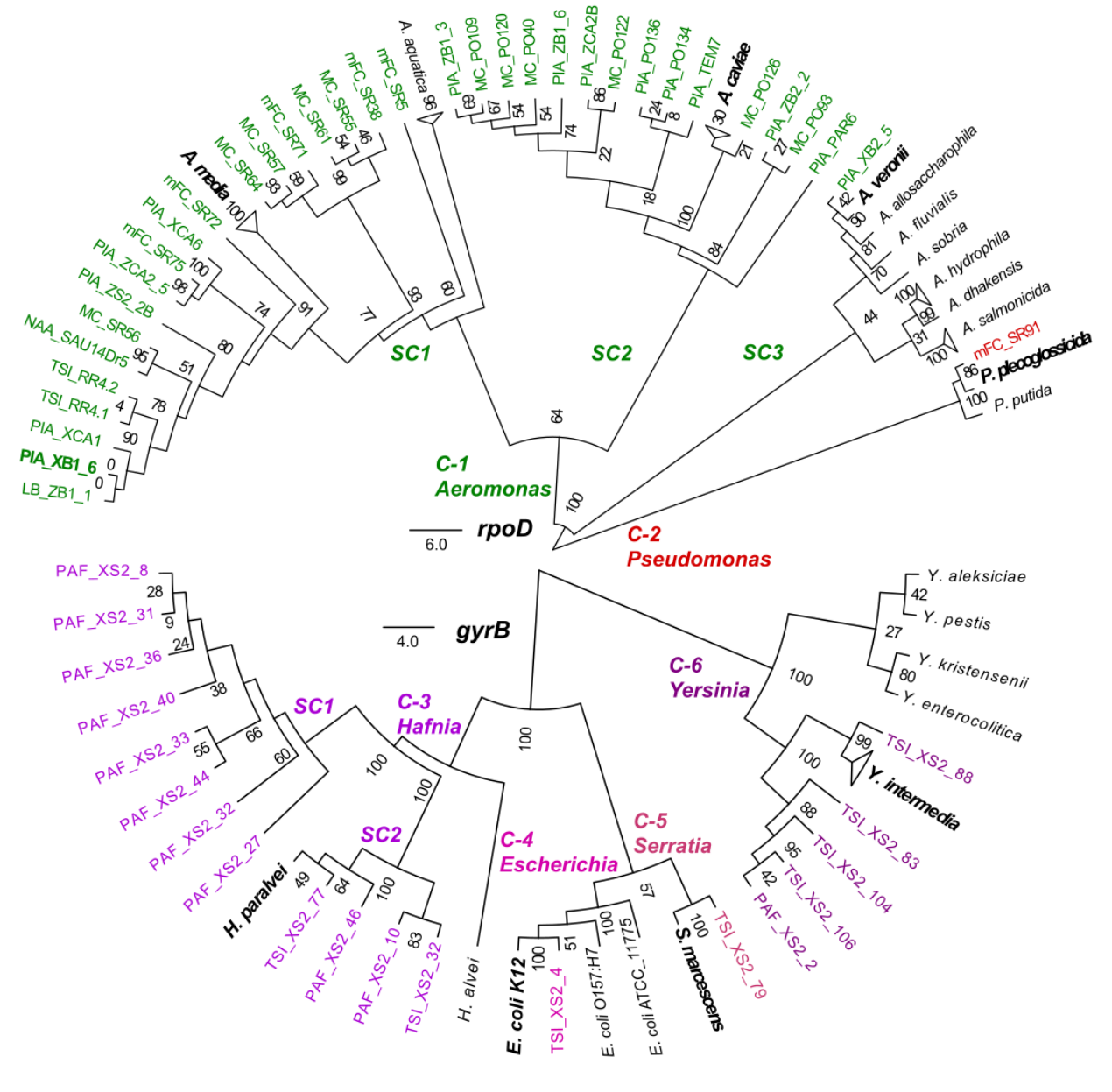


Figura 3. Aislados bacterianos susceptibles a la infección por los fagos promiscuos Atoyac. La relación de los aislados susceptibles a los fagos del grupo Atoyac es mostrada a partir de árboles filogenéticos Neighbor-joining construidos a partir de los alineamientos de la secuencia de nucleótidos de los genes *rpoD* (mostrado en la parte superior) o *gyrB* (mostrado en la parte inferior) bajo el modelo de sustitución J-C. Los aislados fueron considerados como susceptibles únicamente cuando los fagos

producían placas líticas en ellos (causando infecciones productivas, ver Figura 4). Los 54 aislados susceptibles de nuestra colección son indicados con etiquetas de colores acorde al género bacteriano al que pertenecen. Cepas de referencia de diferentes especies bacterianas incluidas en los árboles son indicadas con etiquetas negras, siendo destacadas (en *bold type*) aquellas que se agrupan con los aislados de nuestra colección. Los clusters y subclusters de los diferentes géneros bacterianos identificados son indicados por los prefijos “C-” y “SC”, respectivamente. Los valores de los soportes de rama a partir de 100 réplicas *bootstrap* son mostrados en los nodos de los árboles. Las escalas en barra representan el número esperado de sustituciones por sitio bajo este modelo.

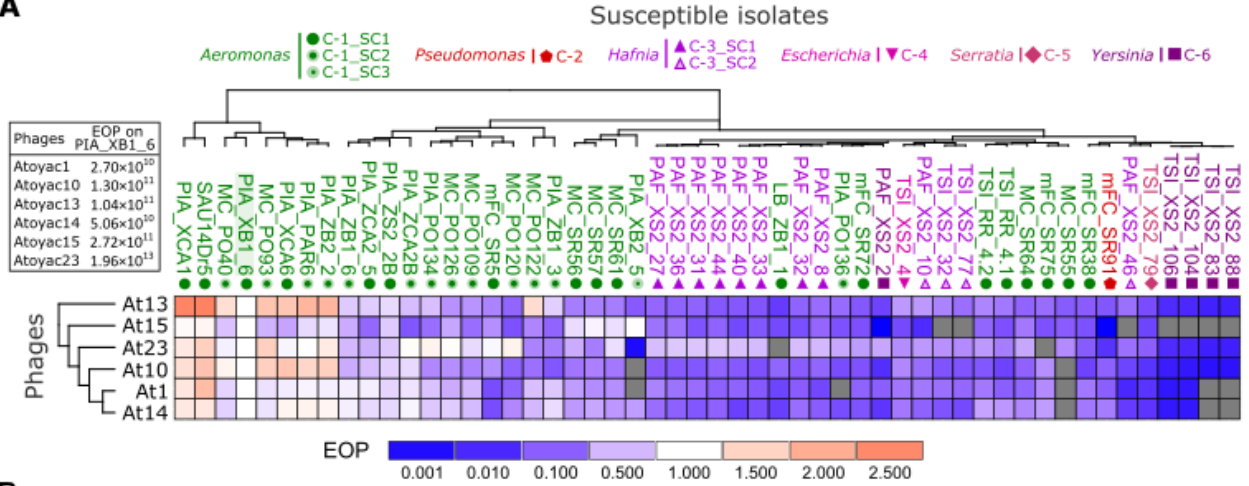
		Taxonomic groups						
		<i>Aeromonas</i>	<i>Pseudomonas</i>	<i>Hafnia</i>	<i>Escherichia</i>	<i>Serratia</i>	<i>Yersinia</i>	
		C1_SC1 XB1-6	C1_SC2 ZB2-2	C2 SR91	C3 XS2-36	C4 XS2-4	C5 XS2-79	C6 XS2-2
Phages	Atoyac1							
	Atoyac10							
	Atoyac13							
	Atoyac14							
	Atoyac15							
	Atoyac23							

Figura 4. Placas de los fagos Atoyac en cepas representativas de los diferentes grupos taxonómicos susceptibles a su infección. Las placas líticas muestran la infección productiva de diferentes géneros bacterianos por los fagos Atoyac y su rango promiscuo de infección. Los ensayos de plaqueo fueron realizados usando stocks purificados por CsCl. Los diferentes grupos taxonómicos bacterianos indicados corresponden a aquellos de la Figura 3.

Eficiencia de plaqueo de los fagos Atoyac

Los seis fagos Atoyac mostraron rangos de hospedero similares. Las diferencias detectadas en los rangos de hospedero correspondieron a su capacidad para infectar ciertos aislados de *Aeromonas*, *Hafnia* y *Yersinia* (Figura 5A). Por ejemplo, el fago Atoyac 15 fue el único miembro del grupo capaz de infectar a todos los aislados de *Aeromonas* en el panel, sin embargo, no fue capaz de infectar muchas de las cepas de *Yersinia* y algunas cepas del grupo *Hafnia*. Nosotros determinamos la eficiencia de plaqueo (EOP) [47] en panel de aislados susceptibles para obtener más información acerca de la capacidad de infección de los fagos Atoyac a través de los diferentes géneros bacterianos, para lo cual se comparó la EOP obtenida en cada cepa respecto al número de placas formadas en la cepa referencia *Aeromonas* sp. PIA_XB1_6 (aislada en este estudio y usada como cepa de referencia en la estimación de la EOP, ver Materiales y Métodos) (Figura 5). Los resultados revelaron complejos perfiles de EOP entre los aislados susceptibles, no obstante, surgió un claro patrón reconocible. Encontramos que la clusterización de los aislados bacterianos basada en las EOPs de los fagos refleja en gran medida su relación taxonómica a nivel de género (Figura 5A), sugiriendo que la eficiencia de los fagos Atoyac para producir unidades formadoras de placa (UFPs) es influenciada por el origen taxonómico del hospedero y que cruzar el rango de género impone un costo. En general, los fagos Atoyac produjeron más placas en los aislados de *Aeromonas*, seguido por las cepas de los géneros *Hafnia* y *Yersinia*, respectivamente (Figura 5, los datos detallados pueden ser descargados para su consulta a partir del material suplementario asociado al siguiente link: [Data Set S2](#)). Diferencias estadísticamente significativas en las EOP por fago entre los subclusters de *Aeromonas* no fueron encontradas (test de *Mann-Whitney U* con *P value* de 0.05), indicando que las mayores diferencias en las EOP ocurren entre géneros (Figura 5B, [Data Set S2](#)). Algunas de las particularidades encontradas a partir de esta caracterización que son de destacar fueron que los fagos Atoyac 13 y 23 presentaban los rangos de hospedero más amplios y también las mayores EOP sobre los grupos de las Enterobacterias (Figura 5). No obstante, el fago Atoyac 23 y el fago Atoyac 15 mostraron EOP considerablemente menores sobre el aislado de *Pseudomonas*, el cual fue infectado a su vez con mayor EOP por los fagos Atoyac 1, 14 y 10. Por otro lado, algunos aislados de *Aeromonas* se caracterizaron por presentar EOP menores que las registradas en cepas de otros géneros bacterianos. Los valores de EOP oscilaron entre -1000 y 2.5 veces a través de la colección de cepas susceptibles respecto a la EOP de la cepa de referencia *Aeromonas* sp. PIA_XB1_6 (usada originalmente como la cepa propagativa). Por ejemplo, el stock del fago Atoyac13 produjo 6.72×10^8 UFP/ml sobre el tapiz bacteriano de la cepa de *Yersinia* sp. TSI_XS2_88, 2.68×10^{11} UFP/ml sobre el tapiz de la cepa de *Aeromonas* PIA_XCA1 y 1.04×10^{11} UFP/ml en la cepa de referencia *Aeromonas* sp. PIA_XB1_6 ([Data Set S2](#)).

A



B

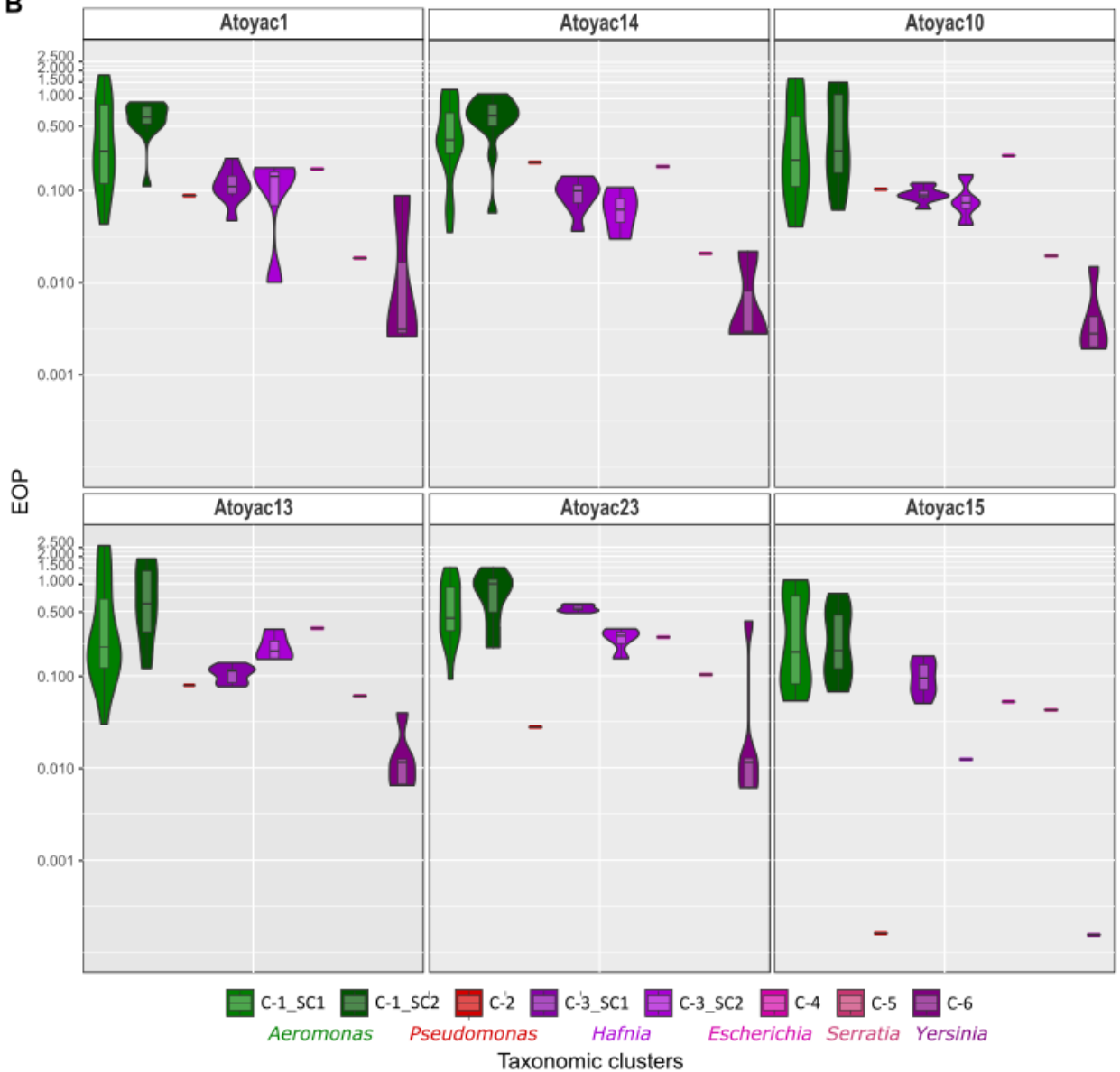


Figura 5. Rango de hospedero y eficiencia de plaqueo de los fagos Atoyac. (A) El mapa de calor ilustra la eficiencia de plaqueo (EOP) registrada para los seis fagos Atoyac (filas: At phages) sobre un panel de 54 cepas ambientales aisladas en este estudio (columnas). Los nombres de los aislados bacterianos son mostrados en la parte superior de la figura con un código de color acorde al género taxonómico al que pertenecen. Los símbolos debajo de los nombres de los aislados indican el cluster o subcluster taxonómico inferido a través de un árbol filogenético (ver Figura 1). El título determinado para cada fago en la cepa propagativa PIA_XB1_6 (aislada en este estudio), indicado en el recuadro del lado izquierdo, fue usado como el valor de referencia para calcular la EOP. La escala en el fondo del heatmap representa la desviación respecto la EOP registrada en la cepa de referencia (EOP= 1.0). Las celdas grises indican las interacciones que no generaron placas líticas detectables, correspondiendo por tanto, a cepas no susceptibles. Los aislados bacterianos y los fagos están agrupados jerárquicamente basados en los valores de EOP mediante el método de distancia Euclidiana. **(B)** Los valores de EOP registrados para cada fago Atoyac agrupados por el cluster taxonómico (género) o subcluster de sus hospederos son mostrados usando un gráfico de violín. Los clusters se muestran coloreados como en el panel A y la Figura 1. Los valores en el gráfico de violín se muestran en escala logarítmica. Las EOP fueron calculadas a partir de cinco réplicas biológicas independientes y los valores promedios se muestran graficados.

Morfología del virión y comparación genómica.

Nosotros caracterizamos los fagos Atoyac a nivel de morfología del virión y de secuencia genómica para investigar sus similitudes entre ellos y con otros fagos. Los análisis por microscopía electrónica de transmisión de las partículas virales purificadas mediante gradiente de CsCl reveló una morfología compartida y típica de los fagos de la familia Podoviridae, caracterizada por presentar una cabeza icosaédrica de ~50 nm unida a una cola corta (Figura 6). Los genomas de los fagos Atoyac presentaron un tamaño promedio de 41.7 kb y un contenido de GC promedio de 59%. El análisis comparativo de sus genomas reveló que ellos comprenden un grupo monofilético, compartiendo entre el 98.4 y el 85.9% de la identidad global de secuencia de nucleótidos (Figura 7). Identificamos dos subgrupos de fagos Atoyac (subgrupo 1: fagos Atoyac1, 14 y 10; subgrupo 2: fagos Atoyac 13 y 23) que estaban más cercanamente relacionados entre ellos a partir del árbol inferido con la secuencia de todo su genoma y que mostraron perfiles de infección muy similares (Figura 7 y Figura 5A). El fago Atoyac 15 fue el más divergente en la comparación genómica y también exhibió uno de los perfiles de infección más disímiles. La mayor fuente de diversidad entre los genomas de los fagos Atoyac correspondió a una larga región no codificante caracterizada por presentar un menor contenido de GC que el promedio global y un cluster adyacente de marcos de lectura abiertos (ORFs) codificando muchas proteínas hipotéticas pequeñas (< 100 aminoácidos) de función desconocida (Figura 7 y FS1). Una marcada variación de secuencia fue también detectada en ORFs contiguos codificando putativamente una fibra de la cola y una proteína de la superfamilia glucosidasa. De los 46 a 50 ORFs identificados en los genomas de los fagos Atoyac, 23 fueron asignados a una función (46% - 50%) basados en la homología de secuencia y presencia de dominios conservados. Estos ORFs se encuentran organizados en dos módulos funcionales principales codificando proteínas relacionadas a (i) la morfogénesis del virión y la emisión de estructuras como la cápside, la cola y la lisis del hospedero, y (ii) a la transcripción y replicación, incluyendo las polimerasas de DNA y RNA, una DNA primasa, una helicasa, una ligasa, una exonucleasa y una endonucleasa con potencial rol en la recombinación (Figura 7A).

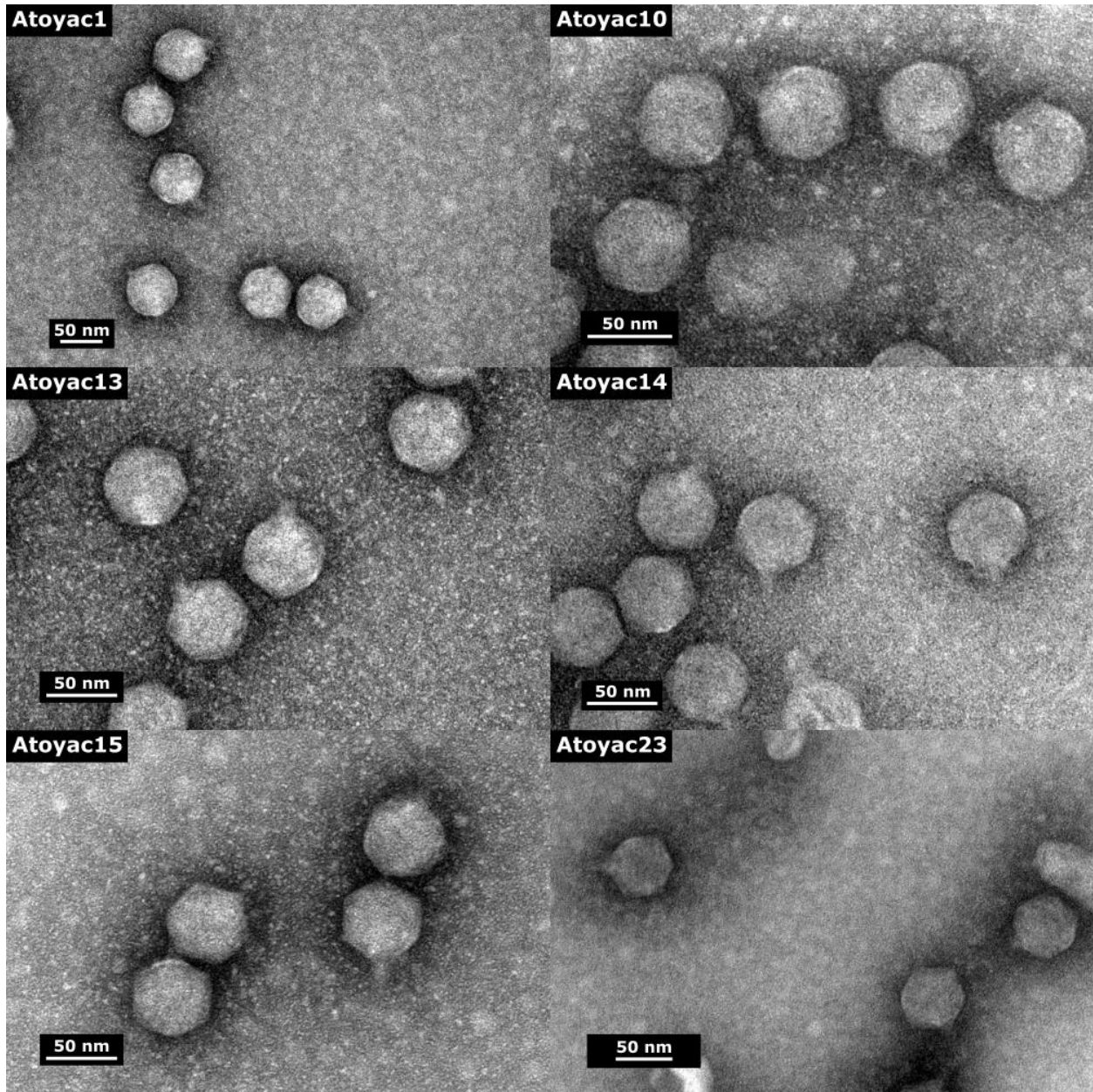


Figura 6. Micrografías de los viriones de los fagos del grupo Atoyac. Las partículas virales purificadas por CsCl fueron teñidas negativamente con acetato de uranilo al 2% y visualizadas por microscopía electrónica de transmisión con un aumento de 150,000 veces. Los nombres correspondientes a los diferentes fagos y la escala de tamaño (50 nm) son mostrados en cada micrografía.

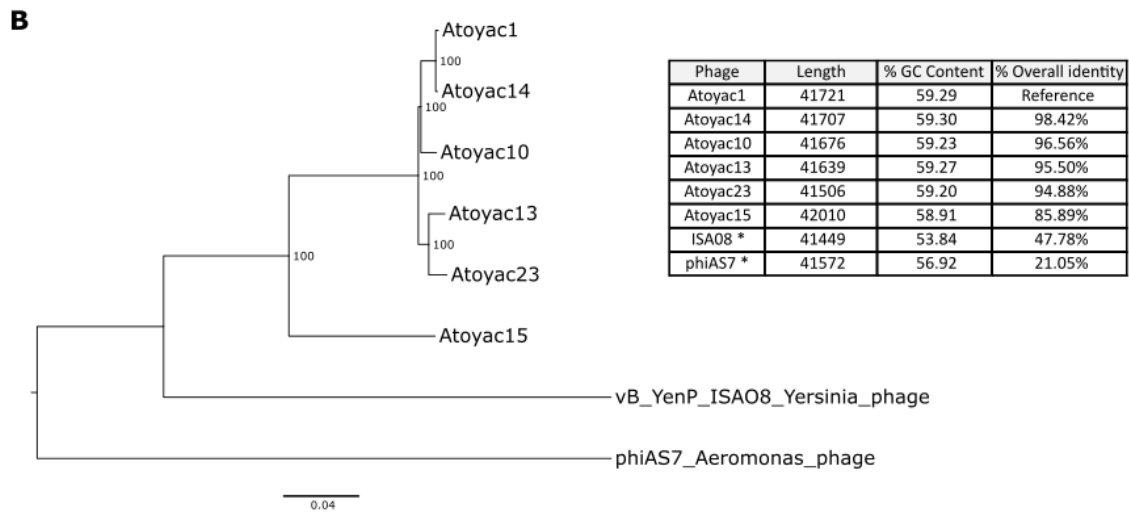
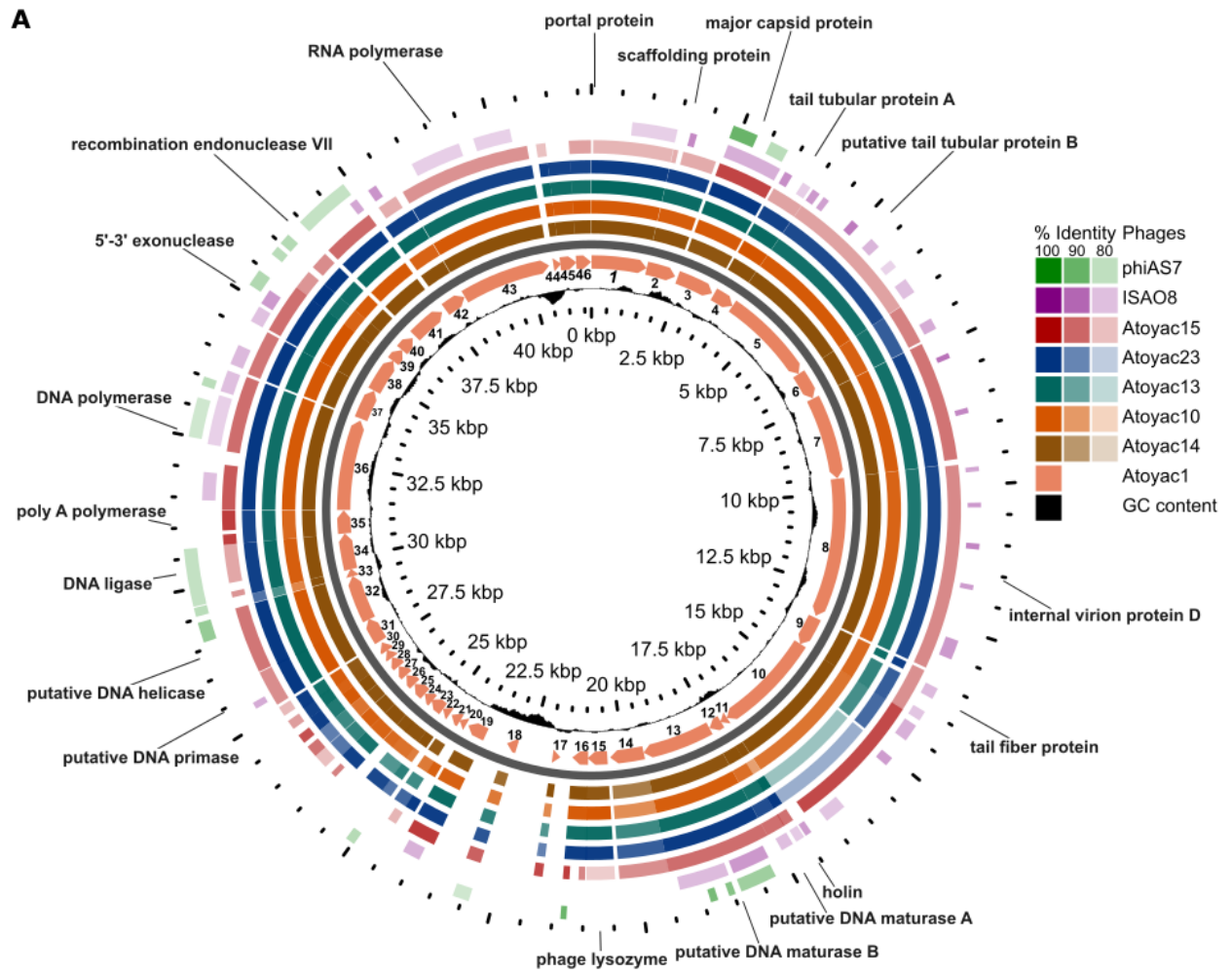


Figura 7. Análisis genómico comparativo y filogenético del grupo de fagos Atoyac. (A) Los genomas de los seis fagos Atoyac y dos homólogos distantes están representados como anillos de colores en el mapa genómico circular comparativo y sus nombres y código de color son indicados a la derecha de la figura. El genoma del fago Atoyac1, representado en el mapa con el anillo de color más interno, fue usado como referencia para comparar las secuencias de los ORFs de los otros fagos a nivel de nucleótidos. Las flechas numeradas en el mapa y la orientación en la que apuntan corresponden al número de ORFs identificados y a su dirección de transcripción, respectivamente. Las funciones identificadas en el genoma son indicadas afuera de los anillos, correspondiendo a los ORFs del fago Atoyac1. Las coordenadas en el genoma de referencia y la distribución de su contenido de GC respecto al promedio (59%) son indicadas en el anillo interno negro correspondiente. El nivel de identidad de secuencia detectado en los genomas (>79%) respecto a los ORFs del fago Atoyac1 es mostrado en código de intensidad de color en la figura. Los dos anillos más externos corresponden a representantes de los grupos de fagos que mostraron la más alta similitud de secuencia a los genomas del grupo Atoyac. **(B)** El árbol genómico de los fagos Atoyac y dos homólogos distantes (neighbor joining) fue construido a partir del alineamiento de sus genomas a nivel de nucleótidos. El valor de soporte de bootstrap (100 réplicas) de las ramas es mostrado en cada nodo. El porcentaje de identidad de secuencia global (es decir, sobre la longitud total del genoma) de los fagos Atoyac y sus homólogos distantes (marcados con un asterisco) respecto al fago Atoyac1 (referencia) es mostrado en la tabla, así también como la longitud y el contenido de GC de cada fago.

No se identificaron en los genomas de los fagos Atoyac genes asociados con el ciclo de vida temperado, implicando que estos fagos son virulentos. Esta observación es consistente con nuestros intentos no exitosos para aislar cepas lisógenas de los fagos Atoyac sobre los fondos genómicos de *Aeromonas* y *Yersinia*. Las búsquedas de homólogos en bases de datos revelaron que los fagos Atoyac constituyen un nuevo grupo, con dos grupos de fagos homólogos distantemente relacionados al clado Atoyac. Los representantes con mayor similitud de estos grupos corresponden al podofago 4vB_YenP_ISAO8 (aquí referido sólo como ISA08, número de acceso GenBank: [NC_028850.1](#)) capaz de infectar a *Yersinia enterocolitica* y al podofago phiAS7 (número de acceso GenBank: [JN651747.1](#) [48]) capaz de infectar a *Aeromonas salmonicida*, con los cuales el fago Atoyac1 comparte el 47.78 y 21.05% de identidad global de secuencia, respectivamente (Figura 7). Nosotros también comparamos los genomas de los fagos Atoyac con el del fago T7, un arquetipo virulento de la familia *Podoviridae* (originalmente aislado usando a *E. coli* como hospedero). A pesar de la falta de similitud de secuencia extensiva entre los genomas a nivel de nucleótidos o de proteínas, la presencia de una larga región no homóloga y no codificante fue reconocida en una posición del genoma relativamente similar. Una región similar identificada en los genomas de fagos T7-like de diferentes especies bacterianas, ha sido reportada por contener múltiples promotores transcripcionales [49]. En concordancia con esta observación, usando tres diferentes programas de predicción *in silico* (ver Materiales y Métodos), nosotros identificamos una serie de promotores putativos en ambos sentidos de una larga región no codificante de los fagos Atoyac ubicada entre los nucleótidos 21553 y 23775 (Figura 7 y material suplementario descargable: [Table S1](#)). Siete promotores fueron predichos en la cadena positiva, 5 de los cuales fueron predichos por dos o más programas, mientras que en la cadena negativa, 8 promotores fueron predichos, tres de ellos por más de un programa ([Table S1](#)).

Búsqueda de fagos Atoyac-like en datos metagenómicos

Los fagos Atoyac fueron aislados a partir de diferentes regiones geográficas en México, las cuales estaban separadas entre ~3.9 a 2,525 kilómetros, sugiriendo una amplia distribución geográfica para este nuevo grupo de fagos promiscuos (Figura 2, Tabla Suplementaria 1). No obstante, ya que no pudieron ser encontrados otros miembros del grupo a través de la búsqueda de homólogos en bases de datos genómicos, decidimos investigar la distribución de los fagos Atoyac mediante la búsqueda de homología en datos metagenómicos disponibles. Para ello intentamos recuperar lecturas de secuenciación homólogas al genoma del fago Atoyac1 analizando conjuntos de datos metagenómicos depositados en el *Sequence Read Archive* (SRA) del NCBI. Sorpresivamente, nuestro análisis de más de 65,000 sets de datos usando una estrategia de búsqueda recientemente reportada (ver Materiales y Métodos) [50], no identificó metagenomas con lecturas que hicieran match con nuestro fago Atoyac de referencia. Este resultado contrasta con los obtenidos cuando usamos como referencia y control para la búsqueda los genomas de otros dos fagos no relacionados, el fago *crAss* (*crAssphage*) y el fago T4. En tales casos, miles de lecturas cubriendo entre el 86.5 y hasta el 100% de los genomas el fago *crAss* y el fago T4 fueron recuperadas a partir de múltiples muestras (set de datos suplementarios descargables: [Data Set S3](#)), sugiriendo la ausencia o baja abundancia de fagos Atoyac-like en el set de datos analizado. Alternativamente, una sub-exploración de los ambientes habitados por los fagos Atoyac (por ejemplo, los ríos) a través de enfoques metagenómicos podría también explicar este resultado [43].

Evolución experimental de la eficiencia de plaqueo

Debido a que los fagos Atoyac mostraron mayores EOPs en los aislados del grupo *Aeromonas* que en cepas de los otros géneros susceptibles, decidimos adoptar un enfoque de evolución experimental para investigar la posibilidad de revertir esta tendencia en hospederos caracterizados por presentar una baja eficiencia producción de placas. Para diversificar el conjunto de hospederos en el experimento, elegimos a *Yersinia* como representante de la familia *Enterobacteriaceae* y a *Pseudomonas* como el único género identificado como susceptible y no perteneciente a *Enterobacteriaceae*, además de *Aeromonas*. Aislados representantes de estos géneros registraron algunos de los valores de EOP más bajos en la colección (Figura 2A). Los fagos Atoyac 1 y 23, que mostraron distintas EOPs en los aislados bacterianos elegidos para el experimento (cepas *Pseudomonas* sp. mFC_SR91 y *Yersinia* sp. PAF_XS2_2; Figura 2A), fueron seleccionados como representativos de los subclusters identificados dentro del grupo Atoyac (Figura 4B). Nosotros permitimos que los fagos evolucionaran en las cepas seleccionadas usando en todo momento sus versiones ancestrales. Tres placas individuales de los fagos Atoyac 1 y 23, cada una representando un linaje, fueron serialmente propagadas a través de 10 pases usando un solo hospedero, ya sea *Pseudomonas* o *Yersinia* (Figura 8, esquemas A y B). Para evaluar cambios en la producción de placas sobre el curso del experimento, la EOP en *Pseudomonas* y *Yersinia* fue calculada antes del experimento y en los pasos de propagación 5 y 10, usando *Aeromonas* como referencia (ver detalles en Materiales y Métodos).

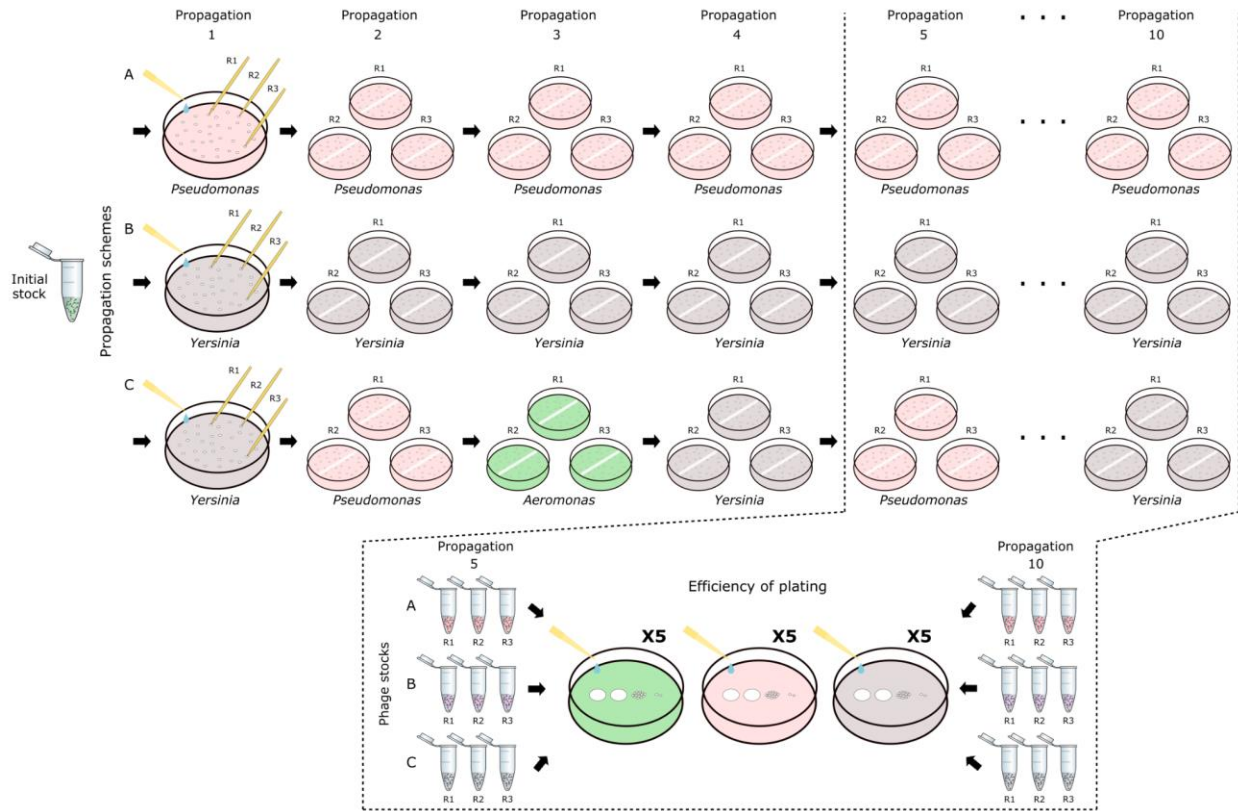


Figura 8. Diagrama de la estrategia usada para los ensayos de evolución experimental. Los fagos Atoyac 1 y 23 fueron sometidos a diferentes esquemas de propagación mediante ensayos de infección en placa para evolucionar su EOP, dos de ellos usando propagación secuencial con un solo hospedero (A y B) y el otro alternando su propagación entre tres diferentes hospederos (C). Stocks de fagos purificados por CsCl (representando el stock ancestral; paso de propagación 0) fueron propagados en *Pseudomonas* (tapices en color rosa) (A) o *Yersinia* (tapices en color gris) (B). Después de su crecimiento, tres placas líticas individuales (R1-R3, linajes derivados del stock ancestral) fueron seleccionadas aleatoriamente desde los tapices bacterianos y propagadas independientemente con un palillo estéril a través de 10 pases en *Pseudomonas* (A) o *Yersinia* (B) (esquemas de propagación usando un solo hospedero) o cíclicamente en *Yersinia*, *Pseudomonas* y *Aeromonas* (tapices en color verde) (C) (esquema de propagación usando múltiples hospederos). Nuevos stocks de fagos fueron generados para cada uno de los tres linajes en los diferentes esquemas desde los pasos de propagación 5 y 10, y estos nuevos stocks de los fagos evolucionados fueron goteados (así también como los stocks iniciales correspondiendo al paso de propagación 0) en los tapices de *Aeromonas*, *Pseudomonas* y *Yersinia* para estimar sus EOPs (a partir de 5 réplicas técnicas) respecto a la eficiencia en *Aeromonas*.

Como se esperaba, el stock ancestral de ambos fagos produjo más placas en la cepa de *Aeromonas* que en los aislados de los otros 2 géneros bacterianos. Sin embargo, esta diferencia en la eficiencia de plaqueo se hizo más estrecha tanto en *Pseudomonas* como *Yersinia* cuando los fagos fueron consecutivamente replicados en el aislado de *Pseudomonas* (Figura 9, paneles A1 and B1), sugiriendo una expansión del rango de hospedero (eficiencia de plaqueo) a través de evolución colateral (idea discutida directamente con el Profesor Mike Bruckost [[Prof Michael Brockhurst | The University of Manchester](#)], investigador experto en tópicos relacionados con la evolución de fagos y otros elementos genéticos móviles). Un resultado similar se observó cuando el fago Atoyac1 fue serialmente propagado en el aislado de *Yersinia* (Figura 9, panel A2); sin embargo, bajo las mismas condiciones de propagación, el fago Atoyac23 no mostró un cambio sustancial en la EOP sobre la cepa de *Yersinia* y redujo considerablemente su virulencia sobre el aislado de *Pseudomonas* (decrecimiento en el número y tamaño de las placas) (Figura 9, panel B2). En paralelo a la propagación serial usando un solo hospedero, nosotros también propagamos los fagos Atoyac cíclicamente alternando sus hospederos entre *Yersinia*, *Pseudomonas* y *Aeromonas* (Figura 8, esquema C) para explorar si esta condición muestra una alteración en la dinámica de la EOP distinta a la obtenida a través de la replicación continua de los fagos usando un solo hospedero. Encontramos que los cambios en la EOP para ambos fagos sobre el curso del experimento se parecen a aquellos registrados cuando los fagos son exclusivamente propagados en la cepa de *Pseudomonas* o cuando el fago Atoyac1 es continuamente replicado en *Yersinia*, es decir, presentan una marcada reducción en el gap de EOP respecto al aislado de referencia de *Aeromonas* (Figura 9, paneles A3 y B3). Interesantemente, bajo este esquema de propagación, el fago Atoyac23 no perdió infectividad hacia la cepa de *Pseudomonas* a pesar de haber sido replicado en el aislado de *Yersinia* (en 4 de los 10 pases de propagación), sugiriendo por lo tanto que la alternancia de hospederos puede ser clave en la retención de la promiscuidad.

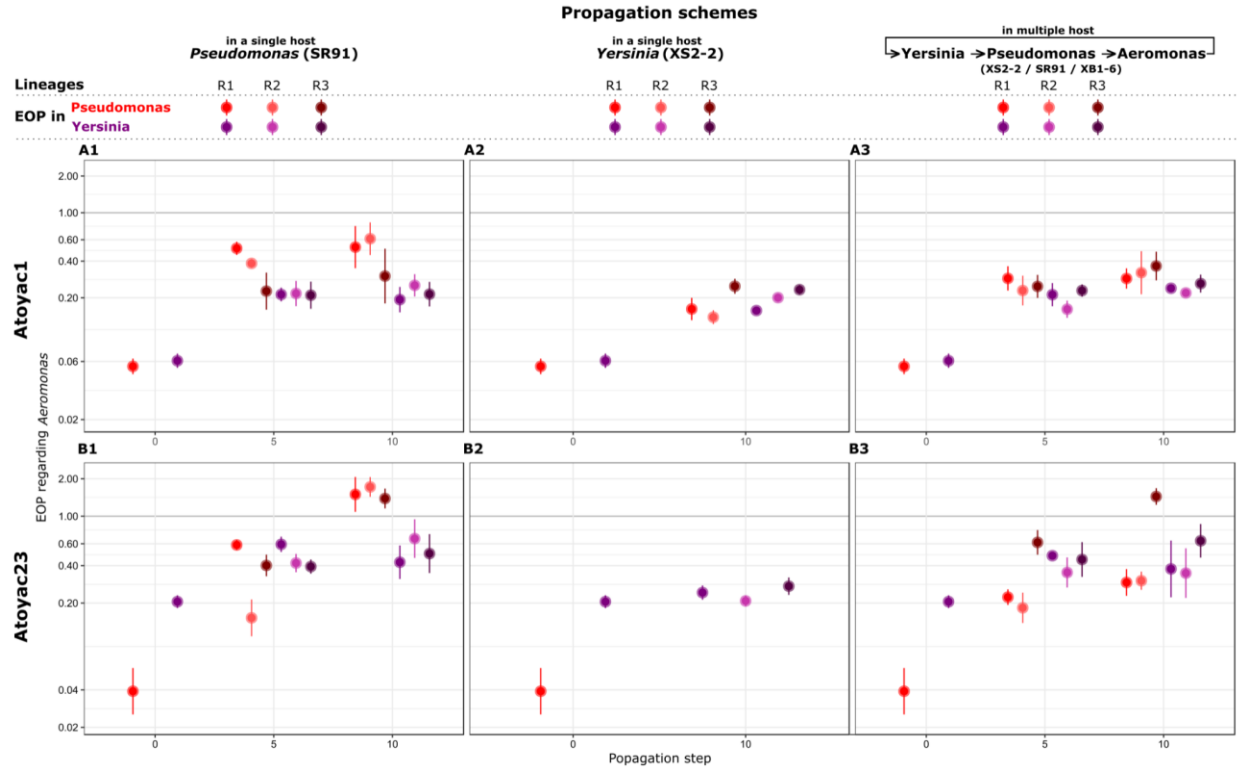


Figura 9. Evolución experimental de la eficiencia de plaqueo de dos fagos representativos del grupo *Atoyac*. Los gráficos representan las trayectorias de las dinámicas de eficiencia de plaqueo registradas durante la evolución experimental de los fagos *Atoyac1* y *Atoyac23* bajo diferentes esquemas de propagación (ver Figura 8). Nótese que en estos experimentos las bacterias usadas en cada propagación corresponden a los hospederos ancestrales; por lo tanto, únicamente se permitió evolucionar a los fagos (ver detalles en Materiales y Métodos). La fila de arriba (paneles A1, A2 y A3) muestra los resultados registrados para el fago *Atoyac1*. La fila de abajo (paneles B1, B2 y B3) muestra los resultados para el fago *Atoyac23*. En la primera columna (A1 y B1), los gráficos muestran las EOPs calculadas en *Pseudomonas* (en rojo) o *Yersinia* (en morado), usando a *Aeromonas* XB1-6 como referencia, cuando los fagos fueron serialmente propagados en *Pseudomonas* por 0, 5 y 10 pases. Los valores graficados en el paso de propagación 0 corresponden a las EOPs de los stocks de fagos ancestrales. La segunda columna (A2 y B2) indica las trayectorias de las EOPs registradas cuando los fagos fueron secuencialmente propagados en *Yersinia*. Valores de EOP para el fago 23 en *Pseudomonas* no son mostrados en el panel B2 y paso de propagación 10 porque ya no fue posible contar sus placas en este hospedero. Finalmente, la tercera columna (A3 y B3) muestra los resultados obtenidos cuando los fagos fueron propagados alternando hospederos entre *Yersinia* XS2-2, *Pseudomonas* SR91 y *Aeromonas* XB1-6, en cada pase. R1, R2 y R3 representan linajes independientes a partir del mismo stock de fago ancestral. Las EOPs para cada linaje fueron calculadas a partir de cinco réplicas técnicas. Los círculos rellenos muestran los promedios, y las barras representan las desviaciones estándar para los linajes. El eje “y” en todos los gráficos son mostrados en escala logarítmica. Todos los datos graficados son presentados para su consulta mediante su descarga en el siguiente link: [Data Set S4](#).

Discusión

Los fagos son las entidades biológicas más abundantes y diversas en el planeta [3]. Aunque el aislamiento de nuevo tipos de fagos representa uno de los propósitos más frecuentes de la investigación en el campo, la mayoría de estos estudios están sesgados o tienen como blanco el aislamiento de fagos con un rango de hospedero estrecho [41], limitando por lo tanto nuestro conocimiento sobre los fagos con capacidades de infección que les permiten cruzar entre diferentes rangos taxonómicos. Estos virus son usualmente referidos como fagos de amplio rango de hospedero; sin embargo, nosotros favorecemos el uso de “promiscuo” siguiendo el término aplicado a plásmidos con capacidades de movilización y replicación que cruzan múltiples rangos taxonómicos [51]. Adicionalmente, el uso del término promiscuo permite la distinción entre virus con hospederos de diferente origen taxonómico y aquellos que infectan varias cepas de la misma especie, comúnmente referidos también como de amplio rango de hospedero [36, 49]. Los fagos promiscuos representan poderosos modelos para ampliar nuestro conocimiento sobre las interacciones fago-bacteria y la evolución del rango de hospedero; sin embargo, aún no es claro que tan prevalentes ellos pueden ser [36].

Estudios recientes enfocados en predecir las relaciones fago-hospedero usando diversas estrategias (por ejemplo, mediante el uso de spacers provenientes de sistemas CRISPR-Cas bacterianos para hacer búsquedas de homología en genomas o contigs de fagos y así establecer asociaciones) han inferido interacciones más allá del nivel de especie o género, sugiriendo que la abundancia de fagos promiscuos puede ser mayor de lo postulado actualmente [36, 41–43]. Intentos previos para aislar fagos promiscuos se basan principalmente en el uso de métodos estándar de enriquecimiento, con la sutil modificación del uso de múltiples bacterias de diferentes especies (y no sólo una durante el procedimiento), difiriendo únicamente en su uso ya sea de manera simultánea o adicionando secuencialmente a los potenciales hospederos. Estos enfoques condujeron al aislamiento de fagos que infectan cepas de dos diferentes géneros o múltiples especies dentro de un mismo género [44, 45]. En estos reportes, se usaron como hospederos cepas de colección (ATCC) o de referencia (por ejemplo: *E. coli* K-12 y *P. aeruginosa* PAO1), sin embargo, aunque se demuestra la eficacia en su uso, esto deriva en el aislamiento de fagos que infectan cepas que probablemente no son representativas de aquellas que existen en los diferentes ambientes.

En contraste, nuestra estrategia de aislamiento exploró el uso de una colección taxonómicamente diversa de aislados bacterianos nativos de los sitios muestreados como hospederos potenciales de los fagos con los que coexisten (Figura 1, ver detalles Materiales y Métodos), bajo la lógica de que es más probable evidenciar entre ellos a múltiples cepas indicadoras (susceptibles a la infección por fagos) e identificar a fagos promiscuos cuando estos provienen de las mismas comunidades microbianas. Nuestra estrategia de aislamiento permitió la identificación de seis fagos, nombrados Atoyac (río en lenguaje Náhuatl), capaces de infectar bacterias de los géneros *Aeromonas*, *Hafnia*, *Yersinia*, *Escherichia*, *Serratia* y *Pseudomonas*, los cuales son parte de clase Gammaproteobacteria. Interesantemente, los fagos Atoyac fueron recuperados a partir de ríos contaminados. Nosotros especulamos que en estos ambientes puede haber una mayor diversidad y densidad poblacional microbiana debido a la abundante disponibilidad de fuentes de

nutrientes introducidos por las constantes descargas residuales humanas, lo cual a su vez podría favorecer la presencia de fagos capaces de infectar diferentes hospederos potenciales. No obstante, no descartamos la posibilidad de que fagos promiscuos puedan ser también recuperados a partir de ambientes prístinos, en los cuales ellos podrían estar presentes quizás en una menor concentración que en los ambientes contaminados. Pensamos que las condiciones de crecimiento que usamos en el laboratorio (ver Materiales y Métodos) podrían ser aplicadas por otros estudios que tengan como propósito el aislamiento de fagos promiscuos, especialmente si las muestras obtenidas provienen de ecosistemas similares a los que exploramos en este estudio. En un escenario distinto, alentamos a los investigadores a explorar condiciones de crecimiento que imiten el entorno natural de los ambientes muestreados, con la finalidad de recuperar aislados bacterianos representativos de la diversidad microbiana de los sitios de interés y maximizar la posibilidad de encontrar fagos promiscuos. Usando un enfoque similar respecto al uso de bacterias nativas como potenciales hospederos, otros autores fueron capaces de aislar un fago proveniente de muestras del mar Antártico capaz de infectar cepas de dos géneros pertenecientes a diferentes clases de proteobacteria [52].

La detección de fagos que infectan diferentes especies puede ser el resultado de una mezcla de múltiples fagos que no han sido separados apropiadamente, conduciendo a una identificación errónea de un rango de hospedero promiscuo [38]. De manera similar, la evaluación del rango de hospedero basada en un efecto bactericida (por ejemplo, a través de spots de lisis) en lugar de una infección productiva puede conducir también a su sobreestimación [38]. En este estudio abordamos estas cuestiones a través de una evaluación exhaustiva de la pureza de los fagos (Figura S2) y la determinación del rango de hospedero basada únicamente en la generación de infecciones productivas (producción de placas líticas). Como ha sido recomendado previamente [38], nuestras preparaciones de stocks de fagos incluyeron más de tres rondas de purificación de placas; adicionalmente los stocks fueron purificados por gradiente de densidad de cloruro de cesio (CsCl) antes de la determinación final del rango de hospedero (Figura 1). Nuestros ensayos de evolución experimental también involucraron 10 pases de propagación secuencial de placas individuales usando un solo hospedero o múltiples hospederos (Figura 8), demostrando concluyentemente la capacidad de propagación intergenérica de los fagos. En nuestros experimentos, nosotros observamos consistentemente evidencia sobre la pureza de los stocks de fagos (Figura S2), incluyendo uniformidad en la morfología de placas, la presencia de una banda única bien definida en los gradientes de CsCl, la inferencia de tamaños similares para los genomas de los fagos a partir de los patrones de restricción de DNA y los ensamblajes de los genomas, y homogeneidad en la morfología de los viriones dentro de una misma muestra y entre muestras. Nuestros resultados mostraron que los seis fagos Atoyac aislados tienen un rango de hospedero similar, lo cual es consistente con su similitud genómica e indicativo de que incluso el miembro más divergente del grupo; correspondiendo al fago Atoyac15 que mostró 86% de similitud de secuencia global respecto al fago Atoyac1, es capaz de infectar los mismos seis géneros bacterianos que sus homólogos. A pesar de la alta similitud genómica entre la mayoría de los miembros del grupo Atoyac, nosotros observamos variedad en los fenotipos de infección entre los fagos, mostrando así que poca variabilidad entre genomas de fagos relacionados es suficiente para presentar diversidad en las capacidades de infección. De manera interesante, encontramos una buena correlación entre el agrupamiento de los fagos Atoyac basado en la

comparación de todo el genoma (Figura 7B) y el inferido a partir de los perfiles de rango de hospedero y EOP (Figura 5A), lo cual sugiere que la variabilidad de la secuencia explica en gran medida la capacidad de infección diferencial mostrada por los fagos. Nosotros identificamos un par de genes estructurales que codifican putativamente para la fibra de la cola y una proteína de la superfamilia glucosidasa, y dos regiones adyacentes; una intergénica caracterizada por la presencia de promotores y otra codificando un cluster de pequeños ORFs putativos de función desconocida que, dada su variación en la secuencia dentro del grupo, podrían ser las responsables de las diferencias observadas en los fenotipos de infección. Mutaciones en las proteínas estructurales del virión, particularmente aquellas que forman parte de la cola, han sido reportadas como responsables de cambios en el rango de hospedero de varios fagos, incluyendo estudios de evolución experimental [53–55]. Por su parte, las regiones variables que codifican ORFs accesorios han sido documentadas en los genomas de los fagos; sin embargo, estas no han sido caracterizadas extensamente en términos de su impacto funcional en la biología de los fagos [56].

Aunque los fagos Atoyac son capaces de infectar los seis géneros bacterianos, los análisis de EOP mostraron que ellos producen más placas líticas en los aislados de *Aeromonas* que en las otras bacterias susceptibles, sugiriendo que este grupo de fagos está mejor adaptado para infectar el género *Aeromonas* (Figura 5). La alta similitud entre el contenido de GC de los genomas de los fagos Atoyac (59%) y el de *Aeromonas* (58 to 62% [57]) y el mayor tamaño de las placas líticas producidas sobre las cepas de *Aeromonas* soportan adicionalmente esta noción (Figura 4). En ciertos casos, cepas de *Aeromonas* no mostraron susceptibilidad a todos los miembros del grupo Atoyac o presentaron bajas EOPs, incluso menores que las observadas en los aislados de otros géneros bacterianos (por ejemplo: las cepas MC_SR55, PIA_XB2_5, mFC_SR71 o PIA_TEM7) (Figura 5, [Data Set S2](#)). Nosotros creemos que esto puede deberse a que esos aislados de *Aeromonas* pueden contar con mecanismos de defensa proficientes (por ejemplo, un sistema CRISPR-Cas) para contender contra la infección de algunos miembros específicos del grupo. Por otro lado, también es posible que los fagos Atoyac que sí pueden infectar tales aislados o que presentan mejores EOPs cuenten con genes que les permitan superar ciertos mecanismos de defensa bacterianos (tales como genes anti-CRISPR).

Una cepa de *Aeromonas* (PIA_XB1_6) fue seleccionada como el hospedero propagativo para los fagos Atoyac debido a que fue identificada como altamente sensible a la infección por fagos, junto con otras *Aeromonas*, en nuestro screening inicial para detectar cepas indicadoras provenientes de las muestras de agua. Vale la pena señalar que la propagación de los fagos en PIA_XB1_6 antes de la determinación final del rango de hospedero podría también haber mejorado la EOP en este género como consecuencia de un potencial efecto de domesticación por el uso repetido de este fondo genómico. Aunque nosotros tratamos de limitar el número de pases de propagación para evitarlo, este efecto potencial representa un artefacto inherente a los protocolos estándar para obtener stocks de fagos puros, un requerimiento clave previo a la evaluación del rango de hospedero [38]. Debido a que el género *Aeromonas* es conocido ser abundante en diversos ambientes acuáticos como los ríos [58], nosotros hipotetizamos que este sirve como el hospedero primario o recurrente de los fagos Atoyac, mientras que bacterias de los otros géneros susceptibles, también presentes en ríos, pueden actuar como hospederos

secundarios [59, 60]. Especulamos que la presencia de fagos promiscuos en ríos puede estar relacionada a la dinámica de este tipo de ambientes, la cual es comúnmente influenciada por una fuerte estacionalidad (temporadas marcadas de sequía y lluvias intensas) que impacta la composición de las comunidades microbianas presentes diversificándolas regularmente [60, 61]. Dentro de este contexto, un rango de hospedero promiscuo podría representar una ventaja para los fagos, permitiéndoles el uso de hospederos alternativos para su propagación y permanencia en la comunidad, incluso si los hospederos primarios no están a su alcance.

Nosotros observamos diferencias estadísticamente significativas entre las EOPs de los fagos Atoyac obtenidas sobre *Aeromonas* y las obtenidas sobre otros géneros susceptibles, pero no entre subclusters de *Aeromonas* ([Data Set S2](#)), sugiriendo que la infección de bacterias distantemente relacionadas involucra un costo en términos de eficiencia de producción de placas. Esta noción es consistente con lo observado en otros fagos de amplio rango de hospedero, los cuales exhiben una variación sustancial en las EOPs, infectando hospederos distantemente relacionados [44, 45]. Variación en la eficiencia de transferencia o replicación entre hospederos distantemente relacionados también ha sido reportada para plásmidos promiscuos [62], sugiriendo que la habilidad para propagarse entre múltiples fondos genómicos viene acompañada de un costo biológico y que este *trade-off* no es exclusivo de un cierto tipo de elemento genético móvil. Sin embargo, a pesar de la variación en la eficiencia de plaqueo, los fagos Atoyac no mostraron una considerable reducción en muchos de sus hospederos, incluso entre diferentes géneros (~85% de los aislados mostraron una variación de EOP menor de 100 veces; Figura 5). Las bacterias a menudo tienen estrategias de defensa que les permiten limitar la infección por fagos y el ingreso intracelular de otros elementos genéticos móviles, siendo los sistemas de restricción-modificación y los sistemas CRISPR-Cas algunos de los más frecuentes mecanismos de defensa anti-fagos encontrados en los genomas bacterianos. Debido a que estos mecanismos moleculares suelen impactar dramáticamente la EOP en varios órdenes de magnitud [38, 47], especulamos que es posible que alguno o múltiples de los ORFs putativos de función desconocida de los fagos Atoyac, encontrados durante su comparación y anotación genómica, codifican mecanismos para contender contra sistemas de defensa anti-fagos. En un estudio previo, Jensen y colaboradores [44] mostraron que fagos de amplio rango de hospedero que no presentaban un decremento significativo en sus EOPs en diferentes hospederos (tales como *Sphaerotilus natans* y *P. aeruginosa*) son insensibles a la acción de enzimas de restricción tipo I y II. Similarmente, el fago de amplio rango de hospedero T7 codifica y usa la proteína Orc para inhibir enzimas de restricción tipo I como mecanismo para desarmar las defensas bacterianas [63]. Por lo tanto, creemos que, a través del estudio de este tipo de fagos, se podrían encontrar proteínas interesantes con potencial aplicación biotecnológica, y es de nuestro interés el estudiar a mayor profundidad los ORFs de función desconocida de los fagos Atoyac para indagar qué funciones codifican. Entre dichas proteínas interesantes también se encontrarían aquellas involucradas en la lisis bacteriana (sistemas de lisis), ya que podrían representar valiosas herramientas moleculares con capacidades antimicrobianas más versátiles que el uso de fagos específicos, con un potencial similar al uso de los antibióticos de amplio espectro. Similarmente, el uso de fagos promiscuos completos también podría tener un efecto positivo en la terapia con fagos, mejorando la eficacia de los cócteles de fagos para eliminar a más de un patógeno blanco. No obstante, es de resaltar que su implementación podría también conducir a

resultados inesperados en la terapia, eliminando a bacterias que no son blanco del tratamiento (como lo hacen muchos de los antibióticos que actualmente usamos). En consecuencia, es necesario que se realicen estudios exhaustivos implementando el uso de fagos promiscuos en la terapia, para poder así evaluar sus efectos.

Los fagos Atoyac fueron aislados a partir de muestras de agua de río provenientes de diferentes regiones geográficas de México, sugiriendo una amplia distribución de este grupo en ambientes similares. Sorpresivamente, nuestra búsqueda en datos metagenómicos no identificó lecturas de secuencia que mapearán al genoma del fago Atoyac1. Este resultado sugiere la ausencia de homólogos al genoma *query* en el set de los datos explorados, siendo indicativo de la baja abundancia de fagos Atoyac-like en las muestras analizadas. La estrategia de búsqueda usada en este trabajo involucra un paso de submuestreo del conjunto de datos para hacer más eficiente la búsqueda a través de los datos metagenómicos del SRA (Sequence Read Archive) [50] con el costo de ser susceptible a resultados falsos negativos, particularmente cuando las lecturas de secuencia de ciertos taxa están en baja abundancia. Por otro lado, nuestro estudio muestra que los enfoques basados en técnicas de cultivo aún representan una buena alternativa para seguir descubriendo la diversidad de la virosfera, con la ventaja de que conducen al aislamiento de especímenes que adicionalmente pueden ser estudiados en el laboratorio.

El análisis comparativo de los genomas permitió identificar a los fagos Atoyac como un nuevo grupo viral. Interesantemente, los genomas más similares detectados fuera del grupo pertenecen a fagos distantemente relacionados que infectan de manera independiente especies de *Yersinia* o de *Aeromonas*, correspondiendo a dos de los grupos bacterianos identificados como susceptibles a la infección por los fagos Atoyac en este trabajo. La comparación de secuencias de proteínas conservadas de los fagos Atoyac asociadas con el módulo estructural o de replicación permite la identificación de *matches* con similitud a ORFs de fagos con hospederos de diferentes géneros (incluyendo algunos de los identificados ser susceptibles a los fagos Atoyac, tales como: *Pseudomonas*, *Serratia* y *Escherichia*); sin embargo estos *matches* se caracterizan con bajos niveles de similitud de secuencia, indicando de nuevo relaciones distantes entre los fagos. Aunque nosotros hipotetizamos que *Aeromonas* funge como hospedero primario de los fagos Atoyac, ellos muestran mayor similitud de secuencia al genoma del fago de *Yersinia enterocolitica* ISAO8 (~48%) que al genoma del fago de *Aeromonas salmonicida* phiAS7 (~21%). En consecuencia, no es claro si los fagos ISAO8 y phiAS7 son especializados para infectar sus hospederos reportados o tienen un rango de hospedero promiscuo como el grupo de fagos Atoyac. Una perspectiva de esta tesis es la caracterización adicional de ISAO8 y phiAS7, la cual podría proporcionar información valiosa acerca de la evolución del rango de hospedero promiscuo de los fagos Atoyac. No obstante, nuestro estudio destaca la importancia de una caracterización integral del rango de hospedero de un fago, incluyendo hospederos taxonómicamente diversos en el panel de cepas potencialmente susceptibles a evaluar. En el caso presentado en este estudio, los fagos Atoyac podrían haber sido reportados como “específicos” de *Aeromonas* dado el gran número de cepas hospederas detectadas pertenecientes a este género, ocultando así el amplio rango de infección intergenérico del grupo.

El aislamiento de fagos promiscuos representa una oportunidad única para explorar la evolución del rango de hospedero a través de hospederos distantemente relacionados; en consecuencia, nosotros explotamos este modelo para investigar la dinámica de infección de dos fagos Atoyac en hospederos no óptimos (*Pseudomonas* y *Yersinia*) usando diferentes esquemas de propagación. Nuestros resultados mostraron que la propagación sucesiva usando un solo hospedero o alternándolos tienen un fuerte efecto en la dinámica de la EOP y su trayectoria adaptativa, aunque el fago utilizado puede ser un factor crítico en la respuesta. En muchos casos, la propagación seriada promovió la optimización de la EOP en los hospederos ineficientes mientras se mantiene el rango de hospedero promiscuo, sugiriendo por lo tanto que esta es una característica evolutivamente bien conservada. En adición, una variación en la morfología de placa fue también observada en los linajes evolucionados que presentaron cambios en las EOPs (Figura S3). La propagación del fago Atoyac23 mostró los resultados más contrastantes en los experimentos. La replicación continua en *Pseudomonas* permitió a este fago alcanzar una eficiencia de plaqueo y el desarrollo de una morfología de placa similares a las observadas en *Aeromonas* (Figura 9, panel B1 y Figura S3, panel B1), el hospedero inicialmente más eficiente. En contraste, su propagación sucesiva en *Yersinia* no tuvo un impacto sustancial en la EOP en este hospedero. Tales respuestas disímiles podrían estar asociadas con la EOP inicial del fago en estos hospederos, es decir, Atoyac23 mostró una marcada menor eficiencia en *Pseudomonas* respecto a *Yersinia*, y por lo tanto un amplio margen para la optimización. Esta noción es consistente con lo observado en estudio reciente de evolución experimental con un fago de cianobacterias generalista, el cual adquirió un mayor número de cambios evolutivos con un efecto positivo en la eficiencia de infección cuando fue continuamente propagado en un hospedero subóptimo con respecto a un hospedero óptimo [55]. Alternativamente, cambios evolutivos con un efecto positivo en la EOP en *Yersinia* podrían requerir más pasos de propagación para surgir y fijarse en la población. La propagación serial del fago Atoyac1 en *Pseudomonas* mejoró su EOP en este hospedero así también como en el aislado de *Yersinia*, y viceversa. Mientras que el fago Atoyac23, en contraste, mostró una sustancial reducción en su virulencia hacia *Pseudomonas* como consecuencia de su propagación continua en el aislado de *Yersinia* (cambio reflejado también a través de la producción de placas casi imperceptibles; Figura S3, panel B2). La replicación del fago alternando hospederos impidió la aparición de este efecto, resaltando la importancia que tiene la propagación en múltiples hospederos para mantener la promiscuidad en algunos fagos. Nosotros especulamos que patrones de propagación similares podrían ocurrir en ciertos ambientes caracterizados por presentar una alta diversidad microbiana, conduciendo así a la preservación de fagos con un rango de hospedero promiscuo.

Conclusiones

La estrategia usada en nuestro estudio demuestra que es posible aislar fagos promiscuos a partir de fuentes ambientales como los ríos, haciendo uso de aislados bacterianos nativos de los sitios de estudio como hospederos potenciales.

El grupo de fagos promiscuos aislado a partir de muestras de ríos de diferentes regiones de México, mostró la capacidad de generar infecciones productivas en cepas diversas de los géneros *Aeromonas*, *Pseudomonas*, *Escherichia*, *Hafnia*, *Serratia* y *Yersinia*; infectándolos con diferentes eficiencias.

Los seis fagos aislados y nombrados Atoyac, presentaron una morfología de virión compartida, característica del orden Caudovirales y de la familia Podoviridae (virus de cola corta). Además, a nivel genómico mostraron estar altamente relacionados entre ellos, formando un nuevo grupo viral, con fagos homólogos remotos infectivos de los géneros *Yersinia* o *Aeromonas*.

Las principales regiones de variabilidad en los genomas de los fagos Atoyac se encontraron en genes asociados a proteínas de la cola, así también como en una serie de proteínas pequeñas de posible función accesoria y en regiones predichas como enriquecidas en promotores.

Los ensayos de evolución experimental realizados con 2 fagos representantes del grupo Atoyac y hospederos de diferentes géneros, revelaron que la eficiencia de infección tiende hacia la optimización en hospederos ineficientes, sugiriendo adicionalmente que la propagación entre diferentes hospederos podría ser clave para mantener un amplio rango de infección.

Perspectivas

Permanece dentro de nuestro interés el secuenciar los linajes de los fagos Atoyac 1 y 23 propagados a través de hospederos de diferentes géneros, particularmente aquellos que mostraron una optimización de la eficiencia de infección o reducción de la amplitud del rango de hospedero. Pensamos que el análisis de los genomas de estos linajes podría aportar datos valiosos que ayuden a comprender mejor la evolución del rango de hospedero de los fagos, incluyendo la identificación de las regiones genómicas y mutaciones implicadas en la adaptación hacia los hospederos. También nos interesa explorar el rol de este grupo de fagos en posibles eventos de transducción entre diferentes géneros de bacterias, así como evaluar su versatilidad en estrategias enfocadas hacia la optimización de la terapia con fagos. Además, estamos interesados en evaluar el impacto funcional de las proteínas de función desconocida de estos fagos promiscuos y la identificación de aquellas proteínas con un uso potencial biotecnológico.

Materiales y Métodos

Muestreo y conteo de coliformes

Una docena de muestras de ríos y aguas residuales provenientes del centro, norte y sur de México fueron recolectadas en frascos estériles (1 L) durante la temporada de estiaje (mayo) en los años 2015 y 2016. Parámetros fisicoquímicos como temperatura, pH, oxígeno disuelto, etc., fueron medidos in situ usando un instrumento multiparamétrico HANNA HI9828 (ver Tabla S2 en material suplementario). Los valores registrados fueron usados como una guía para explorar nuestras condiciones de crecimiento en el laboratorio. La estimación de los niveles de contaminación del agua en las diferentes muestras fue realizada mediante el conteo de coliformes termotolerantes en el medio mTEC modificado, basándonos en el protocolo de filtración por membrana previamente descrito [64]. Las colonias que produjeron un pigmento azul (mostrando actividad beta-glucoronidasa) fueron contadas, y el número de UFC por mililitro fue calculado basado en el factor de dilución de las muestras; estos experimentos fueron realizados en triplicado. Los sitios de muestreo con un número de coliformes $\geq 1 \times 10^3$ UFC/mL fueron clasificados como contaminados (Tabla S1).

Aislamientos bacterianos

Una colección diversa de aislados bacterianos fue ensamblada a partir de las muestras recolectadas. Las muestras fueron plaqueadas en diferentes medios de cultivo: LB agar (Lennox), BHI (brain heart infusion) agar, PIA (Pseudomonas isolation agar), PAF (Pseudomonas agar F), PAP (Pseudomonas agar P), NAA (nutrient-poor medium) agar, mTEC (membrane thermotolerant E. coli) agar, mFC (membrane fecal coliform) agar, TSI (triple sugar iron) agar y MC (MacConkey) agar, todos manufacturados por Difco y preparados siguiendo las instrucciones del proveedor [65], y luego incubadas a 30°C por 24 h. Después de la incubación, algunas colonias fueron seleccionadas, representando diferentes morfologías, tamaños y colores. Las colonias seleccionadas fueron purificadas mediante tres rondas de pases consecutivos de crecimiento en el mismo medio en que fueron aisladas. Los aislados axénicos fueron crecidos durante un overnight en caldo BHI a 30°C y luego almacenados a -72°C en glicerol al 30% (concentración final) para su respaldo. La colección final de aislados comprendió alrededor de 600 aislados provenientes de los 12 diferentes sitios de muestreo.

Aislamiento de fagos

Tapices bacterianos de los aislados de nuestra colección fueron preparados mezclando de 80 a 250 μ l de cultivos overnight con 3.5 ml de top agar (1% peptona, 0.5% NaCl, 0.7% Bacto agar, 10 mM MgSO₄), y vertiendo la mezcla en alguno de los siguientes medios: T Φ (1% peptona, 0.5% NaCl, 1.1% Bacto agar), LB, BHI, mFC, TSI, MC o MC sin sales [65] con el propósito de encontrar las mejores condiciones para la identificación de placas de fagos (los medios usados para el crecimiento óptimo y visualización de las placas en las cepas identificadas como susceptibles a los fagos Atoyac son enlistadas en el set de datos [Data Set S2](#), y puede ser consultado a través del link). Alícuotas de las muestras de agua usadas para aislar bacterias

fueron centrifugadas, filtradas (filtro de membrana MF-Millipore con tamaño de poro de 0.22- μm), diluidas serialmente y luego goteadas en los tapices bacterianos para identificar cepas que permitían la formación de placas (cepas indicadoras), indicativo de la susceptibilidad a la infección por fagos. Las placas con los tapices fueron incubadas overnight a 30°C. Para facilitar la visualización de placas líticas, los colorantes: ácido rosólico (0.01%), rojo fenol (0.0024%) o azul de bromofenol (0.0024%) fueron usados en alguno de los medios (T Φ , LB, BHI, TSI, MC y mFC). Las concentraciones finales de los colorantes usados fueron basadas en las reportadas ser usadas por los medios de la compañía Difco (tales como mFC o TSI, [65]).

Ensayos de infección cruzada fueron entonces usados para identificar candidatos a fagos promiscuos (es decir, fagos que podrían infectar una amplia diversidad de aislados). Para ello, placas individuales fueron picadas con un palillo estéril y estriadas sobre los tapices bacterianos de las cepas previamente identificadas como susceptibles a fagos. Para evaluar si la infección de una cepa determinada afectaba la infectividad de otros hospederos, los ensayos de infección cruzada se realizaron en dos diferentes maneras: 1) simultáneamente, es decir, una placa seleccionada fue transferida a otros aislados bacterianos en paralelo, 2) secuencialmente, es decir, una placa producida en uno de los aislados probados fue transferida al siguiente hospedero potencial (ver Figura 1). Los fagos capaces de infectar múltiples aislados fueron considerados candidatos a fagos promiscuos, pero únicamente aquellos con el rango de hospedero más amplio fueron seleccionados para una mayor caracterización. Estos fagos fueron entonces purificados mediante 4 pases de propagación secuencial en la cepa PIA_XB1_6 (identificada taxonómicamente como *Aeromonas* sp.), en la cual los fagos exhibieron algunos de los títulos más altos, placas de tamaño más grande y un periodo corto de generación de placas (~5 h) durante el primer screening para identificar cepas susceptibles a fagos; por lo tanto, se consideró un hospedero propagativo óptimo. Adicionales pases de purificación fueron evitados para reducir un posible efecto de domesticación de los fagos en este hospedero. Stocks de alto título fueron generados a partir del último pase de purificación siguiendo los pasos descritos previamente [56].

Identificación taxonómica de los aislados bacterianos a nivel de género

Los aislados bacterianos susceptibles a la infección por los fagos fueron crecidos en LB overnight a 30°C y luego el DNA genómico fue extraído a partir de los cultivos usando el protocolo previamente reportado por Chen y Kuo [66]. El DNA extraído fue usado como templado para amplificar mediante PCR los genes *gyrB* y *rpoD*. Nosotros favorecimos el uso de los genes housekeeping *rpoD* y *gyrB* sobre el uso del gen rRNA 16S como marcadores filogenético porque los primeros tienen una mayor tasa evolutiva y han mostrado mejores resultados para resolver grupos bacterianos complejos como *Pseudomonas*, *Enterobacteria* y *Aeromonas* a nivel intragenérico [67–69]. Los primers para amplificar *gyrB* (forward, 5'-GGCGGTAARTTYGAYGATAAC-3'; reverse, 5'-TAGCCTGGTTCTTACGGTT-3''; reportados en este estudio) fueron usados en aquellos aislados que mostraron características fenotípicas similares a las de la familia *Enterobacteriaceae* (por ejemplo, crecimiento en el medio MC). Mientras que los primers flanqueando *rpoD* (forward, 5'-ATYGAAATCGCCAARCG-3'; reverse, 5'-CGGTTGATKTCCTTGA-3' [70]) fueron usados con el resto de los aislados. Las amplificaciones por PCR fueron realizadas usando una Taq DNA polimerasa (Thermo Scientific)

con 0.2 μ M de cada primer y las siguientes condiciones: 95°C por 3 min; 30 ciclos, consintiendo cada ciclo en 95°C por 30 s, 52°C por 1 min para gyrB y 56°C por 1 min para rpoD, 72°C por 1 min; y 72°C por 10 min. Los productos de PCR fueron purificados usando placas MultiScreen PCR μ 96 y luego secuenciados mediante secuenciación Sanger en la empresa Macrogen. Los géneros bacterianos fueron determinados realizando una búsqueda tipo BLASTn de las secuencias obtenidas contra la base de datos del NCBI para identificar los géneros de los matches más cercanamente relacionados. Las secuencias homólogas (usando únicamente cepas tipo como referencia) fueron alineadas usando Clustal omega y los alineamientos usados para generar árboles neighbor-joining y evidenciar las relaciones taxonómicas [71].

Purificación de los fagos promiscuos

Los stocks de fago fueron tratados con DNasa I y RNasa (1 μ g/ml) e incubados por 30 min a 37°C. Las partículas de fago fueron precipitadas con polietilenglicol (PEG) 8000 al 16% (wt/vol) y 1.4 M NaCl overnight a 4°C y centrifugadas a 8,000 RCF (*relative centrifugal force*) durante 20 min. Los sobrenadantes fueron descartados y el pellet fue resuspendido en 1 ml de SM buffer (50 mM Tris HCl [pH 8], 10 mM MgSO₄, 100 mM NaCl, y 0.01% gelatina). En orden, para remover el PEG de los stocks, un volumen igual de cloroformo fue añadido a los tubos y centrifugado a 8,000 RCF por 5 min. La fase acuosa fue recuperada y purificada mediante ultracentrifugación en un gradiente discontinuo de CsCl como fue previamente descrito [72]. Finalmente, los stocks de fago fueron dializados para remover las sales y almacenados a 4°C.

Determinación del rango de hospedero y eficiencia de plaqueo

Ensayos de infección cruzada (ver arriba sección “Aislamiento de fagos”) fueron usados para caracterizar el rango de hospedero de los fagos promiscuos basado en su habilidad para producir placas individuales en los diferentes aislados bacterianos. Fagos capaces de infectar diferentes géneros bacterianos fueron clasificados como promiscuos y nombrados Atoyac, que significa río en Náhuatl (un lenguaje nativo de México). Para determinar la eficiencia de plaqueo (EOP) de los fagos Atoyac, diluciones seriadas de los stocks purificados por gradiente de CsCl fueron goteadas sobre los tapices de los aislados susceptibles de la colección de aislados bacterianos ensamblada a partir de las muestras de ríos. La EOP fue calculada como el promedio del título de fago en una cepa dada dividido entre el título obtenido en el aislado que fue usado inicialmente para propagar los fagos (PIA_XB1_6). Cada ensayo de plaqueo fue realizado con 5 réplicas.

Análisis estadístico

Con el fin de determinar si los datos de eficiencia de plaqueo presentaban una distribución normal, se graficaron histogramas, gráficos de cuantiles-cuantiles (QQ plots) y se corrió la prueba de Shapiro-Wilk, tomando como datos de entrada los valores de EOP de cada fago en los diferentes géneros bacterianos. Como los datos no presentaron una distribución normal debido a su naturaleza, el test no-paramétrico Mann-Whitney U fue usado y se aplicó una corrección de Bonferroni para múltiples comparaciones en la estima de los valores P. Los valores de $P < 0.05$ fueron considerados estadísticamente significativos (ver set de datos: [Data Set S2](#)). Para los datos de evolución experimental, se usó el análisis de varianza de una sola vía (ANOVA) para comparar las diferencia de la EOP para cada condición (los fagos en los diferentes esquemas de propagación) entre los diferentes puntos de tiempo (pases de propagación). Los valores de P menores de 0.05 fueron considerados estadísticamente significativos (ver [Data Set S4](#)).

Determinación de la morfología del virión

La morfología del virión fue determinada mediante tinción de contraste negativo y microscopía electrónica de transmisión como fue descrito previamente por Cazares *et al.* [56]. En breve, alícuotas de los stock de fagos purificados por CsCl fueron depositadas sobre una gradilla de cobre revestida con una película de carbono y teñida con acetato de uranilo al 2%. Las gradillas fueron examinadas bajo un microscopio electrónico de transmisión JEM-2000 a 80 kV.

Extracción del DNA de los fagos y secuenciación genómica

El DNA de los fagos Atoyac fue extraído a partir de 500 μ l de los stocks purificados por CsCl siguiendo el protocolo previamente descrito [72], excluyendo el paso con proteinasa K. Los valores de pureza y concentración fueron medidos mediante espectrometría en un Nanodrop 2000. Alícuotas de 500 ng de DNA fueron usadas para estimar el tamaño del genoma de los fagos mediante reacciones de digestión-restricción. Finalmente, el DNA para los diferentes fagos fue secuenciado usando la tecnología Illumina (paired-end) en la Unidad Universitaria de Secuenciación Masiva y Bioinformática (UUSMB)-Universidad Nacional Autónoma de México (UNAM).

Análisis de los genomas de los fagos

Las lecturas obtenidas del servicio de secuenciación fueron recortadas para remover los adaptadores y filtradas por calidad usando Trimmomatic v.0.36 [73]. Las lecturas filtradas fueron entonces ensambladas usando Velvet v.1.2.10 [74] y mapeadas contra los ensamblados de novo usando BWA v.0.7.17 [75]. Los archivos con las lecturas mapeadas fueron entonces usados para corregir los ensamblados usando Pilon v.1.22 [76]. Los genomas fueron preliminarmente anotados con Prokka v.1.13 [76, 77] y reanotados manualmente basado en las búsquedas de homología de proteínas y dominios conservados usando InterProScan 5.30-69.0, PSI-BLAST y CD Search [78–80].

Los genomas anotados fueron usados para análisis comparativos de los genomas de los fagos Atoyac y sus fagos homólogos (remotos): 4vB_YenP_ISAO8 (GenBank accession no. [NC_028850.1](#)) y phiAS7 ([JN651747.1](#)). Las regiones de secuencia codificantes (CDS) del fago Atoyac1 fueron usadas como referencia para comparar los fagos usando el servidor GView (<https://server.gview.ca/> [81]) y generar un mapa circular comparativo (BLAST Atlas). Similarmente, los genomas de los fagos Atoyac fueron alineados usando Mauve v.2.4.0 [82] con los parámetros default para crear un mapa lineal comparativo y visualizar con mayor detalle las principales regiones variables entre los genomas. Finalmente, los genomas de los fagos fueron globalmente alineados con Clustal omega [71] y un árbol BioNeighbor-Joining basado en todo el genoma fue construido.

Predicción in silico de promotores

La búsqueda in silico de promotores para la región intergénica fue realizada usando la secuencia del fago Atoyac1 (del nucleótido 21377 al 23783) como referencia del grupo y las herramientas de predicción de promotores (en ambas cadenas): Neural Network Promoter Prediction (NNPP), PromoterHunter y BPROM a través de sus servidores web (www.fruitfly.org/seq_tools/promoter.html, www.phisite.org/main/index.php?nav=tools&nav_sel=hunter, y www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb [83–85].

Búsqueda de los genomas de los fagos Atoyac en datos metagenómicos

El genoma del fago Atoyac1 fue usado como referencia para búsqueda de homólogos en datos metagenómicos depositados en el Archivo de lecturas de secuencia (SRA) del NCBI. Metagenomas shotgun completos disponibles a partir del SRA y previamente curados con PARTIE [86] fueron mapeados al genoma de referencia usando bowtie2 [87] con los parámetros de default a través del Search SRA Gateway (<https://www.searchsra.org/>) [50, 88, 89]. Únicamente alineamientos de lecturas > 50 bp fueron considerados. El número de lecturas mapeadas y la cobertura de los alineamientos fueron determinados a partir de archivos BAM ordenados e indexados usando Samtools v.1.7 [50, 82, 90].

Evolución experimental de la eficiencia de plaqueo

Nuestra estrategia de evolución experimental consistió en la propagación secuencial de los fagos Atoyac1 y Atoyac23 mediante ensayos de plaqueo usando dos diferentes esquemas de propagación. En el primer esquema, los fagos fueron propagados en un solo hospedero, usando ya sea la cepa de *Pseudomonas* SR91 o la de *Yersinia* XS2-2 (ver Figura 8, esquemas A y B), en las cuales ambos fagos bajas EOP respecto a la cepa de *Aeromonas* XB1-6 usada como la referencia para calcular la eficiencias. En el segundo esquema, los fagos fueron propagados en múltiples hospederos cíclicamente, alternando el uso de las cepas XS2-2, SR91 y XB1-6 (ver Figura 8, esquema C). Los fagos Atoyac1 y Atoyac23 fueron seleccionados como los representantes más divergentes de los dos subgrupos identificados dentro del grupo Atoyac (excluyendo el fago Atoyac15 que representa un outlier [ver Figura 7]) y porque estos exhibieron distintos perfiles de EOP en los hospederos bacterianos elegidos para los experimentos. En

ambos esquemas, los fagos fueron propagados durante 10 pases a partir de tres placas aisladas (linajes, R1-R3) obtenidas mediante el plaqueo del stock ancestral de los fagos en el tapiz bacteriano de los hospederos seleccionados. En cada paso de transferencia, cultivos overnight de las versiones ancestrales de los hospederos crecidos a partir de sus respectivos stocks en glicerol, fueron usados para preparar los tapices bacterianos en los ensayos de plaqueo; por lo tanto, únicamente la evolución de los virus fue permitida en los experimentos. Las EOPs de los fagos fueron calculadas en los pasos de propagación 0, 5 y 10 usando stocks de alto título generados en estas etapas que luego fueron goteados en diluciones sobre los tapices de las versiones ancestrales de las cepas de *Pseudomonas*, *Yersinia* y *Aeromonas*. El título registrado en *Aeromonas* en el correspondiente paso de propagación fue usado como referencia para determinar la EOP, la cual fue calculada a partir de 5 réplicas técnicas (ver Figura 8 para la representación del diseño del experimento).

Disponibilidad de los datos y material biológico

Los genomas secuenciados con Illumina y luego anotados de los fagos Atoyac1, Atoyac10, Atoyac13, Atoyac14, Atoyac15 y Atoyac23 fueron depositados en el GenBank bajo los siguientes números de acceso: a partir de [MT682386](https://doi.org/10.1093/nucleic-acids/gtaa001) hasta [MT682391](https://doi.org/10.1093/nucleic-acids/gtaa001). Los seis fagos Atoyac y tres de sus hospederos (*Aeromonas* PIA_XB1_6, *Yersinia* PAF_XS2_2 y *Pseudomonas* mFC_SR91) fueron depositados en el Centro de Referencia para Virus de Bacterias Félix d'Hérelle (Félix d'Hérelle Reference Center for Bacterial Viruses <https://www.phage.ulaval.ca/en/home/>) ubicado en la Universidad de Laval (Université Laval), Quebec, Canadá. Las claves HER asignadas para los seis fagos son: 633 (Atoyac1), 634 (Atoyac10), 635 (Atoyac13), 636 (Atoyac14), 637 (Atoyac15) y 638 (Atoyac23). Mientras que para los tres hospederos son las siguientes: 1633 (*Aeromonas* PIA_XB1_6), 1634 (*Yersinia* PAF_XS2_2) y 1635 (*Pseudomonas* mFC_SR91).

Publicación de los resultados de este estudio

Los hallazgos de este estudio fueron publicados preliminarmente en el servidor para preprints bioRxiv ([doi: https://doi.org/10.1101/2020.08.07.242396](https://doi.org/10.1101/2020.08.07.242396)) y luego en su versión final en el journal mSystem de la Sociedad Americana de Microbiología (ASM) ([DOI: https://doi.org/10.1128/mSystems.00773-20](https://doi.org/10.1128/mSystems.00773-20) (ver sección de Anexos para consultar esta versión).

Material suplementario

Figuras suplementarias



Figura S1. Comparación genómica de los fagos Atoyac y sus principales regiones variables. Una visualización lineal de la comparación de los genomas de los fagos Atoyac obtenida con Mauve (48) es mostrada. Cada genoma de fago es representado por un gráfico de similitud en color rojo (los nombres respectivos son mostrados a la izquierda) y las regiones codificantes anotadas son mostradas como cajas blancas. La altura o relleno de los gráficos es proporcional al nivel de identidad de secuencia en la región del genoma que está siendo observada, con espacios vacíos representando regiones con mayor variabilidad (menor nivel de similitud) y secuencias que son específicas para algunos de los fagos. Una vista ampliada de las dos principales regiones variables en los fagos Atoyac es mostrada en el panel

inferior. Una de las regiones comprende dos ORFs estructurales asociados con las colas de los fagos (a la izquierda), y la otra corresponde a una región no codificante (enriquecida en promotores putativos, [ver material descargable: [Table S1](#)]) y a una serie de pequeños ORFs putativos de función desconocida (a la derecha). Los bloques pequeños de un mismo color corresponden a porciones de DNA compartidas entre los fagos en esta región variable.

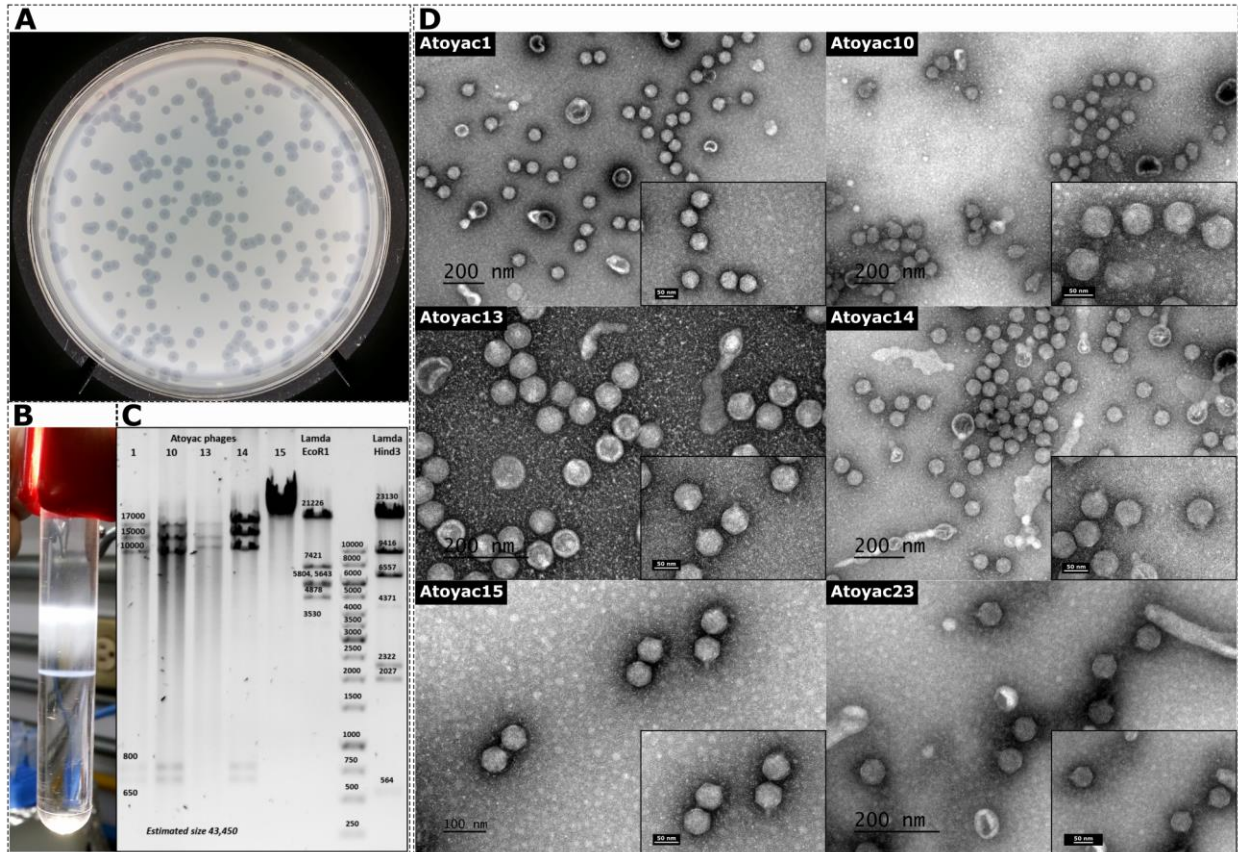


Figura S2. Evidencia de la pureza de los fagos Atoyac. Los fagos Atoyac fueron aislados siguiendo múltiples pasos de control para asegurar su pureza (ver Figura 1 y detalles en Materiales y Métodos). **(A)** Los stocks purificados mediante ultracentrifugación en gradiente de CsCl siempre mostraron una consistente morfología de placa en el hospedero propagativo y otras cepas susceptibles (Figura 2). **(B)** Una sola banda definida (conteniendo los fagos puros) fue observada después de la purificación en el gradiente de CsCl. **(C)** La estimación de los tamaños de los genomas obtenida a partir de la digestión de sus DNAs genómicos (43,450 pb, enzima NdeI) fue muy similar a el tamaño del genoma del único ensamblado obtenido a partir de la secuenciación genómica (41,7 pb en promedio). **(D)** Una sola morfología de virión (correspondiendo al morfotipo podovirus) fue observada mediante microscopía electrónica de transmisión para los stocks puros de los fagos Atoyac.

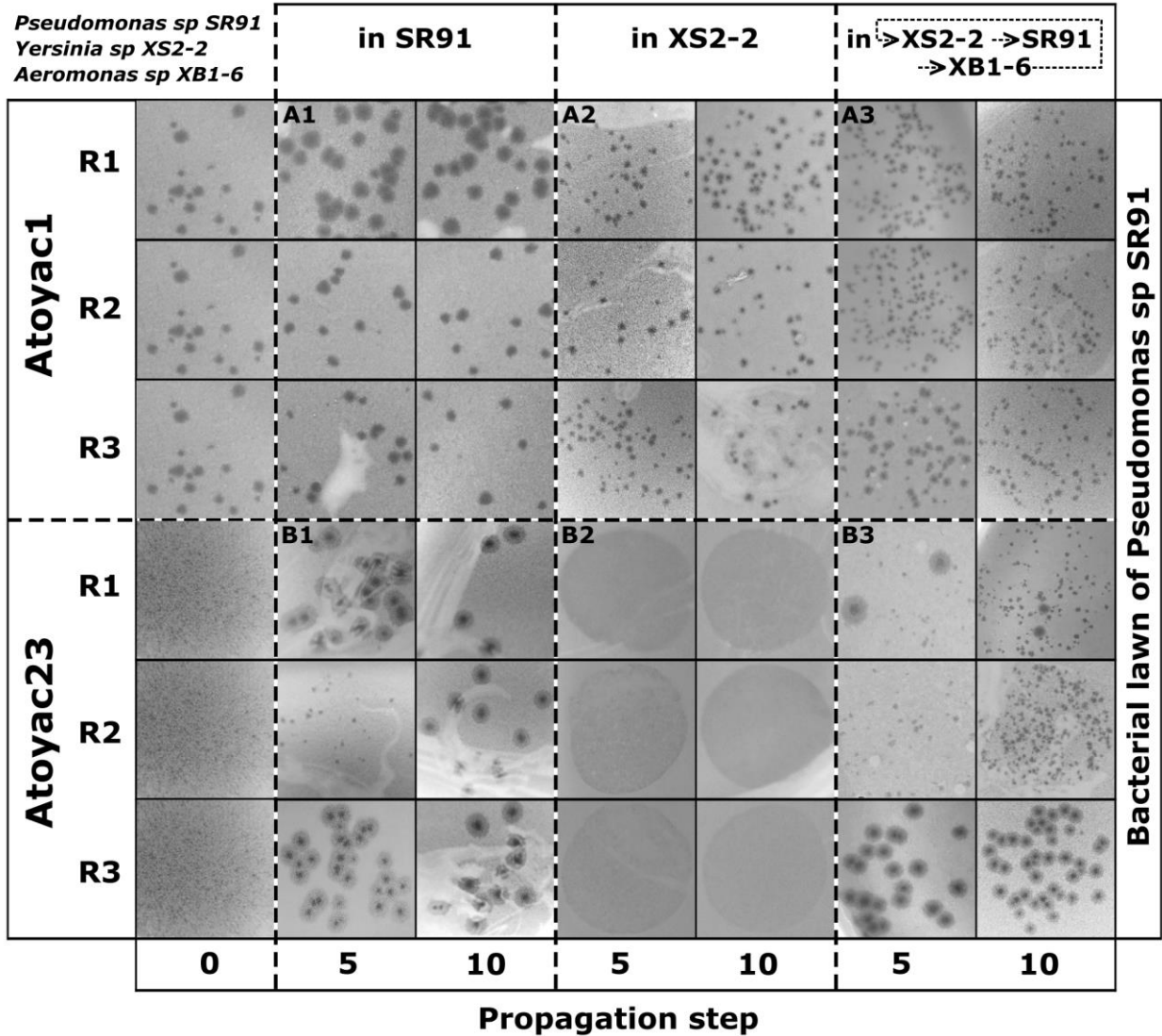


Figura S3. Comparación de la morfología de placa de los fagos Atoyac1 y Atoyac23 a través de los experimentos de evolución experimental. Tres réplicas biológicas (linajes, R1-R3) de los fagos Atoyac1 y Atoyac23 fueron propagadas secuencialmente durante 10 pases en la cepa de *Pseudomonas* SR91 (paneles A1 y B1), en la cepa de *Yersinia* XS2-2 (paneles A2 y B2) y alternando hospederos entre las cepas XS2-2, SR91 y *Aeromonas* XB1-6 (paneles A3 y B3). Las resultantes morfologías de placa de los pasos de propagación 5 y 10 son comparadas respecto a la morfología inicial (paso 0) en la cepa de *Pseudomonas* SR91.

Tablas suplementarias

Sampling site	State	Region	Coordinates	Origin **	Coliforms (cfu/mL)		Classification ***	Promiscuous phages isolated
					Average	SD		
Paraíso *	Morelos	Central	18°49'14"N 99°12'29"W	Sewage	6.47E+05	6.03E+04	Contaminated	Atoyac1, Atoyac10
Río de los Remedios	Ciudad de México	Central	19°31'04"N 99°08'04"W	River/Sewage	3.50E+05	9.00E+04	Contaminated	Atoyac14
Temixco *	Morelos	Central	18°51'13"N 99°13'20"W	River/Sewage	4.30E+04	9.64E+03	Contaminated	Atoyac13
Paseo de Ovejas	Veracruz	South-east	19°16'50"N 96°26'15"W	River/Sewage	1.32E+03	1.41E+02	Contaminated	Atoyac23
Río Tijuana *	Baja California	North-west	32°32'28"N 117°02'13"W	River/Sewage	ND	ND	Contaminated	Atoyac15
Zacatepec	Morelos	Central	18°38'26"N 99°11'40"W	River	9.00E+02	2.00E+02	Intermediate	No
Xochitepec	Morelos	Central	18°46'19"N 99°14'06"W	River	1.37E+02	3.21E+01	Intermediate	No
Río Yauteppec	Morelos	Central	18°43'23"N 99°06'53"W	River	1.23E+02	4.16E+01	Intermediate	No
Civac	Morelos	Central	18°54'14"N 99°09'41"W	Sewage (treated)	4.33E+01	5.77E+00	Clean	No
Sauces	Morelos	Central	18°41'49"N 99°07'45"W	River	9.67E+00	1.53E+00	Clean	No
Estacas	Morelos	Central	18°43'57"N 99°06'47"W	River	1.67E+00	5.77E-01	Clean	No
Chapultepec	Morelos	Central	18°55'08"N 99°12'33"W	River	0.00E+00	0.00E+00	Clean	No

Tabla S1. Sitios de muestreo y lugares de aislamiento de los fagos Atoyac. La tabla indica los sitios muestreados en este estudio, su ubicación geográfica, niveles de contaminación asociados y los fagos promiscuos aislados a partir de ellos. Notas: ND, no determinado, * En estos sitios el fago crAss fue detectado por PCR, ** Origen de los sitios de las muestras: Río = Efluente natural con escasa perturbación humana por residuos provenientes de las ciudades, Río/Aguas residuales = Efluentes naturales altamente perturbados con residuos de las ciudades, Aguas residuales = Efluente artificial designado para el transporte de residuos de las ciudades. *** Categorías que reflejan el número de coliformes detectado: Contaminado => 10E3/mL, Intermedio => 10E2, Limpio = < 10E2.

Sampling site	Environmental Parameters															
	Temperature (°C)		pH		Salinity (per mille)		DO (ppm)		DO (perc)		ORP,mv		TDS,ppm		mbar	
	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD
Paraíso	23.54	0.06	7.58	0.03	0.28	0.01	2.49	0.31	33.93	4.29	30.76	21	289.7	12.44	883.27	0.15
Río de los Remedios	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
Temixco	21.8	0.01	7.52	0.05	0.17	5.59E-17	4.26	0.16	56.14	2.22	13.92	16.65	180.02	1.36	881	0.12
Paseo de Ovejas	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
Río Tijuana	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
Zacatepec	27.68	0.13	7.79	0.02	0.42	0.004	5.48	0.17	78	2.65	-9.94	2.49	430.33	4.8	911.45	0.18
Xochitepec	31.09	0.17	8.47	0.04	0.24	0	11.77	0.28	180.24	4.27	-4.48	3.2	256.21	1.8	898.52	0.18
Río Yauteppec	22.97	0.01	7.32	0.03	0.53	0	6.6	0.12	86.49	1.61	-9.04	2.17	535.55	5.23	907.6	0.01
Civac	25.25	0.05	7.76	0.07	0.13	2.80E-17	6.6	5.77	92.36	1.77	21.46	20.38	134.97	0.16	885.81	0.16
Sauces	20.11	0.16	7.41	0.07	0.27	2.81E-16	7.62	0.11	93.21	1.57	9.8	2.23	280.86	0.75	918.16	0.07
Estacas	22.64	0.03	6.93	0.03	0.63	0.005	2.43	0.09	31.65	1.25	-8.29	4.36	633.84	3.85	910.24	0.05
Chapultepec	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

Tabla S2. Parámetros ambientales de los sitios de muestreo. La tabla muestra los valores de diversos parámetros ambientales registrados en los sitios muestreo de este estudio. Los valores promedio mostrados provienen de las mediciones obtenidas con un instrumento HANNA multi-paramétrico HI9828, operado en modo de mediciones continuas durante un minuto a lo largo de un transecto de 10 m para cada sitio muestreado (las muestras y sus parámetros ambientales fueron tomadas entre el día Mayo 13 y Mayo 16 del año 2016, considerada temporada de estiaje). Notas: ND = no determinado, DO = oxígeno disuelto, ORP = potencial óxido-reducción, TDS = sólidos disueltos totales.

Table S1 : Predicción in silico de promotores en la región intergénica de los fagos Atoyac (archivo descargable nombrado msystems.00773-20-st001.xlsx).

Sets de datos suplementarios descargables

[Data Set S2](#) (archivo [msystems.00773-20-sd002.xlsx](#)): Medios y condiciones usadas en los ensayos de plaqueo. Títulos y EOPs de los fagos Atoyac en los aislados susceptibles. Test de Mann-Whitney U de las diferencias de las EOPs entre grupos taxonómicos y subgrupos para cada fago.

[Data Set S3](#) (archivo [msystems.00773-20-sd003.xlsx](#)): Mapeo de lecturas de metagenomas a los genomas del fago crAss (crAssphage) y del fago T4.

[Data Set S4](#) (archivo [msystems.00773-20-sd004.xlsx](#)): EOPs de los ensayos de evolución experimental y test de ANOVA.

Notas

Diversos archivos del material suplementario pueden ser consultados solo a través de los links asociados para su descarga ([Table S1](#); [Data Set S2](#); [Data Set S3](#); [Data Set S4](#)) debido a la gran cantidad de datos que comprenden.

Referencias

1. Suttle C. The virosphere: the greatest biological diversity on Earth and driver of global processes. *Environmental Microbiology* 2005; **7**: 481–482.
2. Suttle CA. Marine viruses — major players in the global ecosystem. *Nat Rev Microbiol* 2007; **5**: 801–812.
3. Suttle CA. Marine viruses — major players in the global ecosystem. *Nat Rev Microbiol* 2007; **5**: 801–812.
4. Wommack KE, Colwell RR. Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev* 2000; **64**: 69–114.
5. Pratama AA, van Elsas JD. The ‘Neglected’ Soil Virome - Potential Role and Impact. *Trends Microbiol* 2018; **26**: 649–662.
6. Virgin HW. The virome in mammalian physiology and disease. *Cell* 2014; **157**: 142–150.
7. Srinivasiah S, Bhavsar J, Thapar K, Liles M, Schoenfeld T, Wommack KE. Phages across the biosphere: contrasts of viruses in soil and aquatic environments. *Res Microbiol* 2008; **159**: 349–357.
8. Suttle CA. Viruses in the sea. *Nature* 2005; **437**: 356–361.
9. Hatfull GF, Hendrix RW. Bacteriophages and their genomes. *Curr Opin Virol* 2011; **1**: 298–303.
10. Canchaya C, Proux C, Fournous G, Bruttin A, Brüßow H. Prophage genomics. *Microbiol Mol Biol Rev* 2003; **67**: 238–76, table of contents.
11. Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall AM, Mead D, et al. Genomic analysis of uncultured marine viral communities. *Proc Natl Acad Sci U S A* 2002; **99**: 14250–14255.
12. Abedon ST. Chapter 1 Phage Evolution and Ecology. *Advances in Applied Microbiology*. 2009. Academic Press, pp 1–45.

13. Salmond GPC, Fineran PC. A century of the phage: past, present and future. *Nat Rev Microbiol* 2015; **13**: 777–786.
14. Suttle CA. The significance of viruses to mortality in aquatic microbial communities. *Microb Ecol* 1994; **28**: 237–243.
15. Ohnishi M, Kurokawa K, Hayashi T. Diversification of *Escherichia coli* genomes: are bacteriophages the major contributors? *Trends Microbiol* 2001; **9**: 481–485.
16. Menouni R, Hutinet G, Petit M-A, Ansaldi M. Bacterial genome remodeling through bacteriophage recombination. *FEMS Microbiol Lett* 2015; **362**: 1–10.
17. Chiang YN, Penadés JR, Chen J. Genetic transduction by phages and chromosomal islands: The new and noncanonical. *PLoS Pathog* 2019; **15**: e1007878.
18. Brüßow H, Canchaya C, Hardt W-D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* 2004; **68**: 560–602, table of contents.
19. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* 2005; **3**: 722–732.
20. Fillol-Salom A, Alsaadi A, Sousa JAM de, Zhong L, Foster KR, Rocha EPC, et al. Bacteriophages benefit from generalized transduction. *PLoS Pathog* 2019; **15**: e1007888.
21. Thierauf A, Perez G, Maloy AS. Generalized transduction. *Methods Mol Biol* 2009; **501**: 267–286.
22. Saye DJ, Ogunseitan O, Sayler GS, Miller RV. Potential for transduction of plasmids in a natural freshwater environment: effect of plasmid donor concentration and a natural microbial community on transduction in *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 1987; **53**: 987–995.
23. Sander M, Schmieger H. Method for host-independent detection of generalized transducing bacteriophages in natural habitats. *Appl Environ Microbiol* 2001; **67**: 1490–1493.
24. Waldor MK, Mekalanos JJ. Lysogenic conversion by a filamentous phage

encoding cholera toxin. *Science* . 1996. , **272**: 1910–1914

25. Perna NT, Plunkett G 3rd, Burland V, Mau B, Glasner JD, Rose DJ, et al. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 2001; **409**: 529–533.
26. Casas V, Maloy S. Role of bacteriophage-encoded exotoxins in the evolution of bacterial pathogens. *Future Microbiol* 2011; **6**: 1461–1473.
27. Rolain JM, Fancello L, Desnues C, Raoult D. Bacteriophages as vehicles of the resistome in cystic fibrosis. *J Antimicrob Chemother* 2011; **66**: 2444–2447.
28. Muniesa M, Colomer-Lluch M, Jofre J. Potential impact of environmental bacteriophages in spreading antibiotic resistance genes. *Future Microbiol* 2013; **8**: 739–751.
29. Balcazar JL. Bacteriophages as vehicles for antibiotic resistance genes in the environment. *PLoS Pathog* 2014; **10**: e1004219.
30. Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F, Chisholm SW. Transfer of photosynthesis genes to and from Prochlorococcus viruses. *Proc Natl Acad Sci U S A* 2004; **101**: 11013–11018.
31. Hargreaves KR, Kropinski AM, Clokie MRJ. What does the talking?: quorum sensing signalling genes discovered in a bacteriophage genome. *PLoS One* 2014; **9**: e85131.
32. Kåhrström CT. Microbiome: with a little help from my phage friends. *Nat Rev Microbiol* 2013; **11**: 507.
33. Modi SR, Lee HH, Spina CS, Collins JJ. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 2013; **499**: 219–222.
34. Labrie SJ, Samson JE, Moineau S. Bacteriophage resistance mechanisms. *Nat Rev Microbiol* 2010; **8**: 317–327.
35. (Ph)ighting Phages: How Bacteria Resist Their Parasites. *Cell Host Microbe* 2019; **25**: 184–194.
36. de Jonge PA, Nobrega FL, Brouns SJJ, Dutilh BE. Molecular and Evolutionary

- Determinants of Bacteriophage Host Range. *Trends Microbiol* 2019; **27**: 51–63.
37. Drulis-Kawa Z, Majkowska-Skrobek G, Maciejewska B, Delattre A-S, Lavigne R. Learning from bacteriophages - advantages and limitations of phage and phage-encoded protein applications. *Curr Protein Pept Sci* 2012; **13**: 699–722.
38. Hyman P, Abedon ST. Bacteriophage Host Range and Bacterial Resistance. 2010. Elsevier, pp 217–248.
39. Ackermann H-W. Viruses of Prokaryotes: General Properties of Bacteriophages. 1987.
40. Matsuzaki S, Rashel M, Uchiyama J, Sakurai S, Ujihara T, Kuroda M, et al. Bacteriophage therapy: a revitalized therapy against bacterial infectious diseases. *J Infect Chemother* 2005; **11**: 211–219.
41. Koskella B, Meaden S. Understanding bacteriophage specificity in natural microbial communities. *Viruses* 2013; **5**: 806–823.
42. Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. Computational approaches to predict bacteriophage-host relationships. *FEMS Microbiol Rev* 2016; **40**: 258–272.
43. Paez-Espino D, Eloë-Fadrosch EA, Pavlopoulos GA, Thomas AD, Huntemann M, Mikhailova N, et al. Uncovering Earth's virome. *Nature* 2016; **536**: 425–430.
44. Jensen EC, Schrader HS, Rieland B, Thompson TL, Lee KW, Nickerson KW, et al. Prevalence of broad-host-range lytic bacteriophages of *Sphaerotilus natans*, *Escherichia coli*, and *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 1998; **64**: 575–580.
45. Yu P, Mathieu J, Li M, Dai Z, Alvarez PJJ. Isolation of Polyvalent Bacteriophages by Sequential Multiple-Host Approaches. *Appl Environ Microbiol* 2016; **82**: 808–815.
46. Ballesté E, Pascual-Benito M, Martín-Díaz J, Blanch AR, Lucena F, Muniesa M, et al. Dynamics of crAssphage as a human source tracking marker in potentially faecally polluted environments. *Water Res* 2019; **155**: 233–244.
47. Kutter E. Phage host range and efficiency of plating. *Methods Mol Biol* 2009; **501**:

141–149.

48. Kim JH, Son JS, Choresca CH, Shin SP, Han JE, Jun JW, et al. Complete genome sequence of bacteriophage phiAS7, a T7-like virus that infects *Aeromonas salmonicida* subsp. *salmonicida*. *J Virol* 2012; **86**: 2894–2895.

49. Kawasaki T, Shimizu M, Satsuma H, Fujiwara A, Fujie M, Usami S, et al. Genomic characterization of *Ralstonia solanacearum* phage phiRSB1, a T7-like wide-host-range phage. *J Bacteriol* 2009; **191**: 422–427.

50. Levi K, Rynge M, Abeysinghe E, Edwards RA. Searching the Sequence Read Archive using Jetstream and Wrangler. *Proceedings of the Practice and Experience on Advanced Research Computing*. 2018. Association for Computing Machinery, New York, NY, USA, pp 1–7.

51. Krishnapillai V. Molecular genetic analysis of bacterial plasmid promiscuity. *FEMS Microbiol Rev* 1988; **4**: 223–237.

52. Luhtanen A-M, Eronen-Rasimus E, Oksanen HM, Tison J-L, Delille B, Dieckmann GS, et al. The first known virus isolates from Antarctic sea ice have complex infection patterns. *FEMS Microbiol Ecol* 2018; **94**.

53. Heineman RH, Springman R, Bull JJ. Optimal foraging by bacteriophages through host avoidance. *Am Nat* 2008; **171**: E149–57.

54. Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR, et al. Antagonistic coevolution accelerates molecular evolution. *Nature* 2010; **464**: 275–278.

55. Enav H, Kirzner S, Lindell D, Mandel-Gutfreund Y, Béjà O. Adaptation to sub-optimal hosts is a driver of viral diversification in the ocean. *Nat Commun* 2018; **9**: 4698.

56. Cazares A, Mendoza-Hernández G, Guarneros G. Core and accessory genome architecture in a group of *Pseudomonas aeruginosa* Mu-like phages. *BMC Genomics* 2014; **15**: 1146.

57. Lorén JG, Farfán M, Miñana-Galbis D, Fusté MC. Prediction of whole-genome

- DNA G+C content within the genus *Aeromonas* based on housekeeping gene sequences. *Syst Appl Microbiol* 2010; **33**: 237–242.
58. Janda JM, Abbott SL. The genus *Aeromonas*: taxonomy, pathogenicity, and infection. *Clin Microbiol Rev* 2010; **23**: 35–73.
59. Leclerc H, Schwartzbrod L, Dei-Cas E. Microbial agents associated with waterborne diseases. *Crit Rev Microbiol* 2002; **28**: 371–409.
60. Ochoa-Sánchez LE, Vinuesa P. Evolutionary Genetic Analysis Uncovers Multiple Species with Distinct Habitat Preferences and Antibiotic Resistance Phenotypes in the Complex. *Front Microbiol* 2017; **8**: 1548.
61. Breton-Deval L, Sanchez-Flores A, Juárez K, Vera-Estrella R. Integrative study of microbial community dynamics and water quality along The Apatlaco River. *Environ Pollut* 2019; **255**: 113158.
62. Datta N, Hedges RW, Shaw EJ, Sykes RB, Richmond MH. Properties of an R factor from *Pseudomonas aeruginosa*. *J Bacteriol* 1971; **108**: 1244–1249.
63. Häuser R, Blasche S, Dokland T, Haggård-Ljungquist E, von Brunn A, Salas M, et al. Bacteriophage protein-protein interactions. *Adv Virus Res* 2012; **83**: 219–298.
64. Ciebin BW, Brodsky MH, Eddington R, Horsnell G, Choney A, Palmateer G, et al. Comparative evaluation of modified m-FC and m-TEC media for membrane filter enumeration of *Escherichia coli* in water. *Appl Environ Microbiol* 1995; **61**: 3940–3942.
65. Difco & BBL Manual: Manual of Microbiological Culture Media. 2009. Becton Dickinson and Company, Sparks, MD.
66. Chen WP, Kuo TT. A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. *Nucleic Acids Res* 1993; **21**: 2260.
67. Yamamoto S, Kasai H, Arnold DL, Jackson RW, Vivian A, Harayama S. Phylogeny of the genus *Pseudomonas*: intrageneric structure reconstructed from the nucleotide sequences of *gyrB* and *rpoD* genes. *Microbiology* 2000; **146 (Pt 10)**: 2385–2394.

68. Fukushima M, Kakinuma K, Kawaguchi R. Phylogenetic analysis of *Salmonella*, *Shigella*, and *Escherichia coli* strains on the basis of the *gyrB* gene sequence. *J Clin Microbiol* 2002; **40**: 2779–2785.
69. Navarro A, Martínez-Murcia A. Phylogenetic analyses of the genus *Aeromonas* based on housekeeping gene sequencing and its influence on systematics. *J Appl Microbiol* 2018; **125**: 622–631.
70. Mulet M, Bennasar A, Lalucat J, García-Valdés E. An *rpoD*-based PCR procedure for the identification of *Pseudomonas* species and for their detection in environmental samples. *Mol Cell Probes* 2009; **23**: 140–147.
71. Sievers F, Higgins DG. Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol Biol* 2014; **1079**: 105–116.
72. Sambrook J, Russell DW. *Molecular Cloning: A Laboratory Manual*. 2001. CSHL Press.
73. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30**: 2114–2120.
74. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008; **18**: 821–829.
75. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**: 1754–1760.
76. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014; **9**: e112963.
77. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–2069.
78. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 2014; **30**: 1236–1240.

79. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**: 3389–3402.
80. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 2017; **45**: D200–D203.
81. Petkau A, Stuart-Edwards M, Stothard P, Van Domselaar G. Interactive microbial genome visualization with GView. *Bioinformatics* 2010; **26**: 3125–3126.
82. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 2010; **5**: e11147.
83. Reese MG. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput Chem* 2001; **26**: 51–56.
84. Li RW. Metagenomics and Its Applications in Agriculture, Biomedicine and Environmental Studies. 2011. Nova Biomedical Books.
85. Klucar L, Stano M, Hajduk M. phiSITE: database of gene regulation in bacteriophages. *Nucleic Acids Res* 2010; **38**: D366–70.
86. Torres PJ, Edwards RA, McNair KA. PARTIE: a partition engine to separate metagenomic and amplicon projects in the Sequence Read Archive. *Bioinformatics* 2017; **33**: 2389–2391.
87. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; **9**: 357–359.
88. Towns J, Cockerill T, Dahan M, Foster I, Gaither K, Grimshaw A, et al. XSEDE: Accelerating scientific discovery. *Comput Sci Eng* 2014; **16**: 62–74.
89. Stewart CA, Cockerill TM, Foster I, Hancock D, Merchant N, Skidmore E, et al. Jetstream: a self-provisioned, scalable science and engineering cloud environment. *Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by*


Enhanced Cyberinfrastructure. 2015. Association for Computing Machinery, New York, NY, USA, pp 1–8.

90. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.

Anexos



A Novel Group of Promiscuous Podophages Infecting Diverse Gammaproteobacteria from River Communities Exhibits Dynamic Intergenous Host Adaptation

 Daniel Cazares,^a Adrian Cazares,^{b,c} Wendy Figueroa,^d Gabriel Guarneros,^e Robert A. Edwards,^f Pablo Vinuesa^g

^aCentro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

^bEMBL's European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge, United Kingdom

^cWellcome Sanger Institute, Wellcome Genome Campus, Cambridge, United Kingdom

^dDepartment of Biochemistry, University of Cambridge, Cambridge, United Kingdom

^eDepartamento de Genética y Biología Molecular, CINVESTAV IPN, Mexico City, Mexico

^fFAME, Flinders University, Adelaide, SA, Australia

Adrian Cazares and Wendy Figueroa contributed equally to this work.

ABSTRACT Phages are generally described as species specific or even strain specific, implying an inherent limitation for some to be maintained and spread in diverse bacterial communities. Moreover, phage isolation and host range determination rarely consider the phage ecological context, likely biasing our notion on phage specificity. Here we isolated and characterized a novel group of six promiscuous phages, named Atoyac, existing in rivers and sewage by using a diverse collection of over 600 bacteria retrieved from the same environments as potential hosts. These podophages isolated from different regions in Mexico display a remarkably broad host range, infecting bacteria from six genera: *Aeromonas*, *Pseudomonas*, *Yersinia*, *Hafnia*, *Escherichia*, and *Serratia*. Atoyac phage genomes are ~42 kb long and highly similar to each other, but not to those currently available in genome and metagenome public databases. Detailed comparison of the phages' efficiency of plating (EOP) revealed variation among bacterial genera, implying a cost associated with infection of distant hosts, and between phages, despite their sequence similarity. We show, through experimental evolution in single or alternate hosts of different genera, that efficiency of plaque production is highly dynamic and tends toward optimization in hosts rendering low plaque formation. However, adaptation to distinct hosts differed between similar phages; whereas one phage optimized its EOP in all tested hosts, the other reduced plaque production in one host, suggesting that propagation in multiple bacteria may be key to maintain promiscuity in some viruses. Our study expands our knowledge of the virosphere and uncovers bacterium-phage interactions overlooked in natural systems.

IMPORTANCE In natural environments, phages coexist and interact with a broad variety of bacteria, posing a conundrum for narrow-host-range phage maintenance in diverse communities. This context is rarely considered in the study of host-phage interactions, typically focused on narrow-host-range viruses and their infectivity in target bacteria isolated from sources distinct to where the phages were retrieved from. By studying phage-host interactions in bacteria and viruses isolated from river microbial communities, we show that novel phages with promiscuous host range encompassing multiple bacterial genera can be found in the environment. Assessment of hundreds of interactions in diverse hosts revealed that similar phages exhibit different infection efficiency and adaptation patterns. Understanding host range is fundamental in our knowledge of bacterium-phage interactions and their impact on microbial communities. The dynamic nature of phage promiscuity revealed in our study has

Citation Cazares D, Cazares A, Figueroa W, Guarneros G, Edwards RA, Vinuesa P. 2021. A novel group of promiscuous podophages infecting diverse gammaproteobacteria from river communities exhibits dynamic intergenous host adaptation. *mSystems* 6:e00773-20. <https://doi.org/10.1128/mSystems.00773-20>.

Editor Seth Bordenstein, Vanderbilt University

Ad Hoc Peer Reviewer Antonet Svircev, Agriculture and Agri-food Canada

The review history of this article can be read [here](#).

Copyright © 2021 Cazares et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

Address correspondence to Daniel Cazares, cazares@comunidad.unam.mx, or Pablo Vinuesa, vinuesa@ccg.unam.mx.

Received 6 August 2020

Accepted 13 January 2021

Published 2 February 2021

implications in different aspects of phage research such as horizontal gene transfer or phage therapy.

KEYWORDS bacteriophages, environmental microbiology, experimental evolution, microbial ecology, phage-bacterium interactions

Bacteriophages, or phages for short, have been systematically described as species- or even strain-specific viruses (1–3). Since phage infection is initially conditioned to chance encounters with a receptive host cell, the premise of the narrow host range of the phages implies a severe limitation for them to be maintained and spread within environments featuring diverse microbial communities (4), which could promote the existence of phages with a wider range of potential hosts. Although several phages capable of infecting bacteria beyond the taxonomic rank of species or genus have been described, they are still few in number compared to the substantial volume of reports on genus- or species-specific phages (1–3).

Characterization of the host range is a paramount step toward assessing phage specificity and identifying specimens with a broad scope of host targets. Since the host range of a bacteriophage is described as the taxonomic diversity of hosts it can successfully infect (5), the strategies used to interpret a “successful infection” are key to characterize this trait. In order to infect, phages must attach to the host cell, transport their genomes into the cell, overcome various defense mechanisms, and hijack the host’s molecular machinery to replicate, assemble, and release the new viral progeny (5, 6). Infection assays in liquid cultures, known as lysis curves, and spotting or plaque assays resulting in the formation of lysis spots or individual lytic plaques represent examples of commonly used methods to infer phage infection (2). As factors other than phages (e.g., pyocins) and mechanisms independent of phage production (e.g., lysis from without) can also lead to cell death by lysis, the isolation of individual lytic plaques, known as phage plaquing, is regarded as a reliable approximation to identify productive infections (i.e., an infection cycle yielding new phage progeny). Still, it is worth noting that plaque assays are susceptible to false-negative results in cases where phage infection do not lead to plaque formation (e.g., with reduced infection vigor or chronic infections) (2).

In spite of being labeled as infrequent, the isolation and study of phages with the ability to infect multiple species or genera can significantly contribute to advancing our understanding of the host range evolution and phage ecology. Recent metagenomics studies have provided valuable insights into the distribution of broad-host-range phages in different environments (5, 7, 8); however, the nature of this approach restricts further experimental characterization of these viruses which could target unculturable bacteria or strains that have not been isolated yet. On the other hand, some pioneering studies have focused on the targeted isolation of broad-host-range phages through enrichment procedures by either coculture (9) or consecutive monoculture (10) with potential hosts of different taxonomic origin. Such approaches led to the isolation of interesting unrelated specimens infecting bacteria from multiple species, including *Pseudomonas aeruginosa* and *Escherichia coli*.

Here we addressed the identification of broad-host-range phages by assembling a large collection of bacterial and phage isolates present in freshwater and wastewater samples. We hypothesized that using a taxonomically diverse panel of bacteria native to these environments as potential hosts would facilitate the isolation of phages infecting multiple species from such microbial communities. This approach led to the identification of a group of closely related podophages, named Atoyac (Nahuatl for river), capable of productively infecting bacteria from six different genera and three orders within the *Gammaproteobacteria* class. These phages, here referred to as “promiscuous” comprise a novel viral group with undetectable presence in available metagenomic data. The Atoyac phages display different efficiencies of plating among the genera they infect; however, such efficiency rapidly changes over consecutive propagation

in a single or multiple genomic backgrounds, indicative of dynamic host adaptation. The adaptation trajectories are phage dependent and lead to different outcomes despite the high similarity existing between the phages. Our findings contribute to expanding the virosphere and provide insights into short-term host range evolution and virus-host interactions in nature.

RESULTS

Isolation of promiscuous phages. We searched for promiscuous phages, i.e., phages capable of infecting hosts of distinct taxonomic origin, by assembling a large and diverse collection of over 600 bacterial isolates and free phages existing in river and wastewater water samples from 12 different sites in Mexico (see Materials and Methods and Fig. S1 in the supplemental material). Serially diluted aliquots from most of the samples formed clearing zones and lytic plaques when spotted onto lawns of the panel of bacterial isolates retrieved from the water samples. The use of standard media such as T Φ or LB did not always allow the identification of phage plaques; however, utilization of diverse media and different dyes (see Materials and Methods) led to plaque visualization. A cross-infection screening among the susceptible isolates led to the identification of phages producing lytic plaques in multiple strains (see Materials and Methods and Fig. S1). Remarkably, six phages retrieved from five sampling sites were able to infect 54 bacterial isolates in the panel. Only these phages, which we named Atoyac, were selected for further analyses in this study. Susceptible bacteria included members of the genera *Aeromonas*, *Pseudomonas*, *Hafnia*, *Escherichia*, *Serratia*, and *Yersinia*, confirming the promiscuous nature of the Atoyac phage host range, which crosses multiple taxonomic ranks within the *Gammaproteobacteria* class (Fig. 1 and Fig. S2). The Atoyac phages were isolated from samples obtained from Central (Atoyac1, Atoyac10, Atoyac13, and Atoyac14), Southeast (Atoyac23), and Northwest (Atoyac15) regions in Mexico (see Data Set S1 in the supplemental material). Assessment of fecal contamination by either coliform count or presence of crAssphage (used as a marker to detect human fecal pollution in water [11]), revealed that the Atoyac phages were prevalent in contaminated sites, which represented most of our samples (Data Set S1). Two Atoyac phages (1 and 10) were isolated from sewage, and the other four were retrieved from rivers contaminated with sewage discharge. Contaminated samples were also the source of the highest phage titers determined in the plaquing assays.

Efficiency of plating of the Atoyac phages. The six Atoyac phages displayed similar host ranges. Differences in host range detected among the phages corresponded to the ability to infect some of the *Aeromonas*, *Hafnia*, and *Yersinia* isolates (Fig. 2A). Phage Atoyac15, for instance, was the only member of the group capable of infecting all the *Aeromonas* isolates in the panel but could not infect most of the *Yersinia* strains or some representatives of the *Hafnia* group.

We determined the efficiency of plating (EOP) (12) across the panel of susceptible strains to gain further insights into the infection capacity of the Atoyac phages and compared the EOP on each strain to the number of plaques formed on the reference strain *Aeromonas* sp. strain PIA_XB1_6 (isolated in this study and used as “reference” in estimating EOP) (Fig. 2). The results revealed complex EOP profiles among the susceptible isolates, although a clear pattern emerged. We found that clustering the bacterial isolates based on the phages’ EOPs largely reflects their taxonomic relatedness at the genus level (Fig. 2A), suggesting that the efficiency of the Atoyac phages to produce PFU is influenced by the taxonomic origin of the host and that crossing the genus barrier imposes a cost. In general, Atoyac phages generated significantly more lytic plaques on *Aeromonas* isolates, followed by strains from the *Hafnia* and *Yersinia* genera, respectively (Fig. 2 and Data Set S2). No statistically significant differences in EOP per phage were detected between *Aeromonas* subclusters (Mann-Whitney U test and *P* value of 0.05), indicating that the major EOP differences occur between genera (Fig. 2B and Data Set S2). Some *Aeromonas* isolates featuring EOP values lower than those recorded in strains of other genera were also detected.

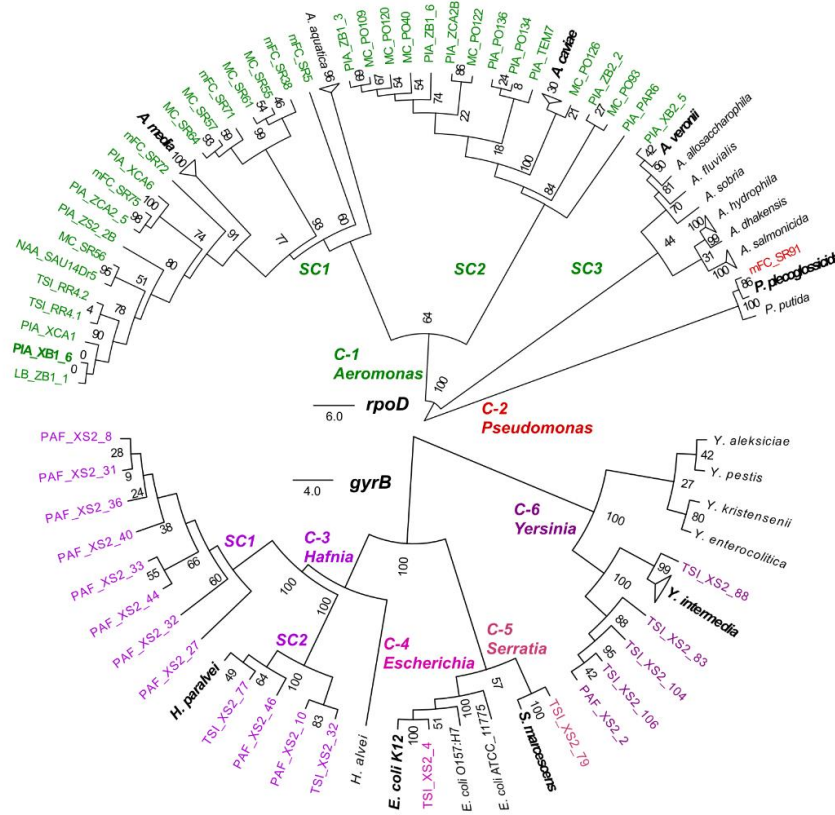


FIG 1 Environmental bacterial isolates susceptible to infection by the promiscuous phages Atoyac. Neighbor-joining trees of the bacterial isolates infected by phages of the Atoyac group were reconstructed from the alignment of nucleotide sequences of the genes *rpoD* (top) or *gyrB* (bottom) under the J-C substitution model. Strains were considered susceptible only when Atoyac phages formed individual lytic plaques on them (see Materials and Methods). The 54 isolates from our collection are indicated in color labels corresponding to the bacterial genus they belong to. Reference strains of different bacterial species included in the trees are indicated in black labels, and the reference strains clustering with isolates from our collection are indicated in bold type. Clusters and subclusters of the different bacterial genera identified in the tree are indicated by the “C-” and “SC-” prefixes, respectively. Branch support values from 100 replica bootstrap tests are shown in the trees. The scale bar represents the expected number of substitutions per site under this model.

The EOP values measured in the collection of susceptible strains ranged from $-1,000$ to 2.5 times the EOP on the reference strain *Aeromonas* sp. PIA_XB1_6 (original propagative strain, see Materials and Methods). For example, the stock of the Atoyac13 phage produced 6.72×10^8 PFU/ml on a bacterial lawn of *Yersinia* sp. strain TSL_XS2_88, 2.68×10^{11} PFU/ml on *Aeromonas* sp. strain PIA_XCA1, and 1.04×10^{11} PFU/ml on the reference strain *Aeromonas* sp. PIA_XB1_6 (Data Set S2).

Virion morphology and genome comparison. We characterized the Atoyac phages at the virion morphology and genome levels to investigate the similarities between them and to other phages. Electron microscopy analysis of the CsCl gradient-purified viral particles revealed a shared morphology typical of phages of the *Podoviridae* family, featuring an ~ 50 -nm icosahedral head attached to a very short tail (Fig. 3 and Fig. S5).

The Atoyac phage genomes average 41.7 kb in size and 59% GC content. A

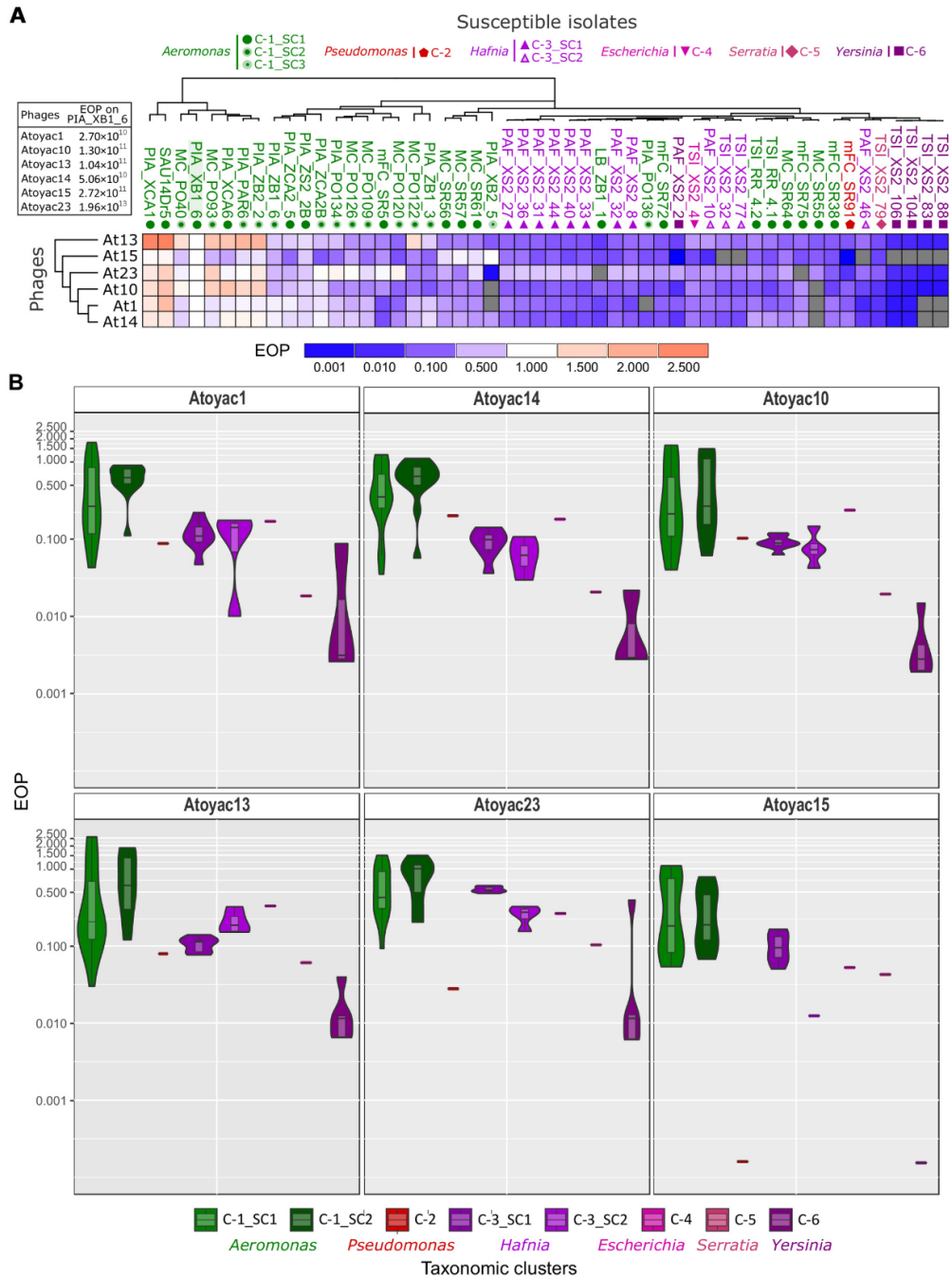


FIG 2 Host range and efficiency of plating of the Atoyac phages. (A) The heatmap illustrates the efficiency of plating (EOP) recorded for six Atoyac (At) phages (rows) on a panel of 54 environmental strains isolated in this work (columns). Names of the bacterial isolates are color coded (Continued on next page)

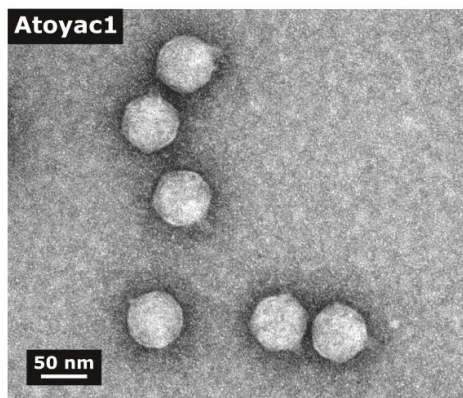


FIG 3 Electron micrograph of virions of the Atoyac phage 1. The morphology of phage Atoyac1 is representative of those observed for all members of the Atoyac group. The CsCl-purified viral particles were negatively stained with 2% uranyl acetate and visualized at 150,000-fold magnification.

comparative analysis unveiled that they comprise a monophyletic group sharing from 85.9 to 98.4% overall nucleotide sequence identity (Fig. 4). We identified two subgroups of phages (group 1, Atoyac phages 1, 14, and 10; group 2, Atoyac phages 13 and 23) that were more closely related to each other in the whole-genome tree and displayed very similar infection profiles. Phage Atoyac15 was the most divergent in the genome comparison and also exhibited one of the most dissimilar infection profiles. The major source of diversity among the Atoyac phage genomes corresponded to a large noncoding region featuring GC content lower than the global average and an adjacent cluster of open reading frames (ORFs) encoding mostly small (<100-amino-acids [aa]) hypothetical proteins of unknown function (Fig. 4 and Fig. S3). Marked sequence variation was also detected in contiguous ORFs putatively encoding a tail fiber and a glycosidase superfamily protein.

Of the 46 to 50 ORFs identified in the Atoyac genomes, 23 were assigned a function (46% to 50%) based on sequence homology and the presence of conserved domains. These ORFs are organized in two major functional modules encoding proteins related to (i) virion morphogenesis and release such as capsid, tail, and lysis proteins and (ii) transcription and replication, including RNA and DNA polymerases, a DNA primase, helicase, ligase, exonuclease, and endonuclease with a potential role in recombination (Fig. 4). Genes associated with a temperate lifestyle (e.g., repressor or integrase) were not identified in the genomes, implying that the Atoyac phages are virulent. This observation is consistent with our unsuccessful attempts to isolate lysogens of Atoyac phages in the *Aeromonas* and *Yersinia* backgrounds.

Homology searches revealed that the Atoyac phages constitute a novel group. Database searching revealed two groups of phages distantly related to the Atoyac clade. The most similar representatives from these groups correspond to the podophages 4vB_YenP_ISA08 infecting *Yersinia enterocolitica* (hereafter referred to as

FIG 2 Legend (Continued)

according to the taxonomic genus they belong to as indicated at the top of the figure. Symbols below the isolate names indicate the bacterial cluster or subcluster inferred through sequence alignment of a marker gene (see Fig. 1). The titer determined for each phage on the propagating strain PIA_XB1_6 (isolated in this study), indicated in the left inset, was used as the reference value to calculate the EOP. The scale at the bottom of the heatmap depicts the deviation from the EOP recorded in the reference isolate (EOP = 1.0). Gray cells correspond to interactions that did not generate detectable lytic plaques, thus indicating nonsusceptible strains. Both bacterial isolates and phages are hierarchically clustered based on the EOP values using the Euclidean distance method. (B) The EOP values recorded for each Atoyac phage grouped by the taxonomic cluster (genus) or subcluster of their hosts are plotted. Clusters are color coded as in panel A and Fig. 1. The EOP values were calculated from five independent biological replicates, and the averages are plotted. Values in the violin plots are in log scale.

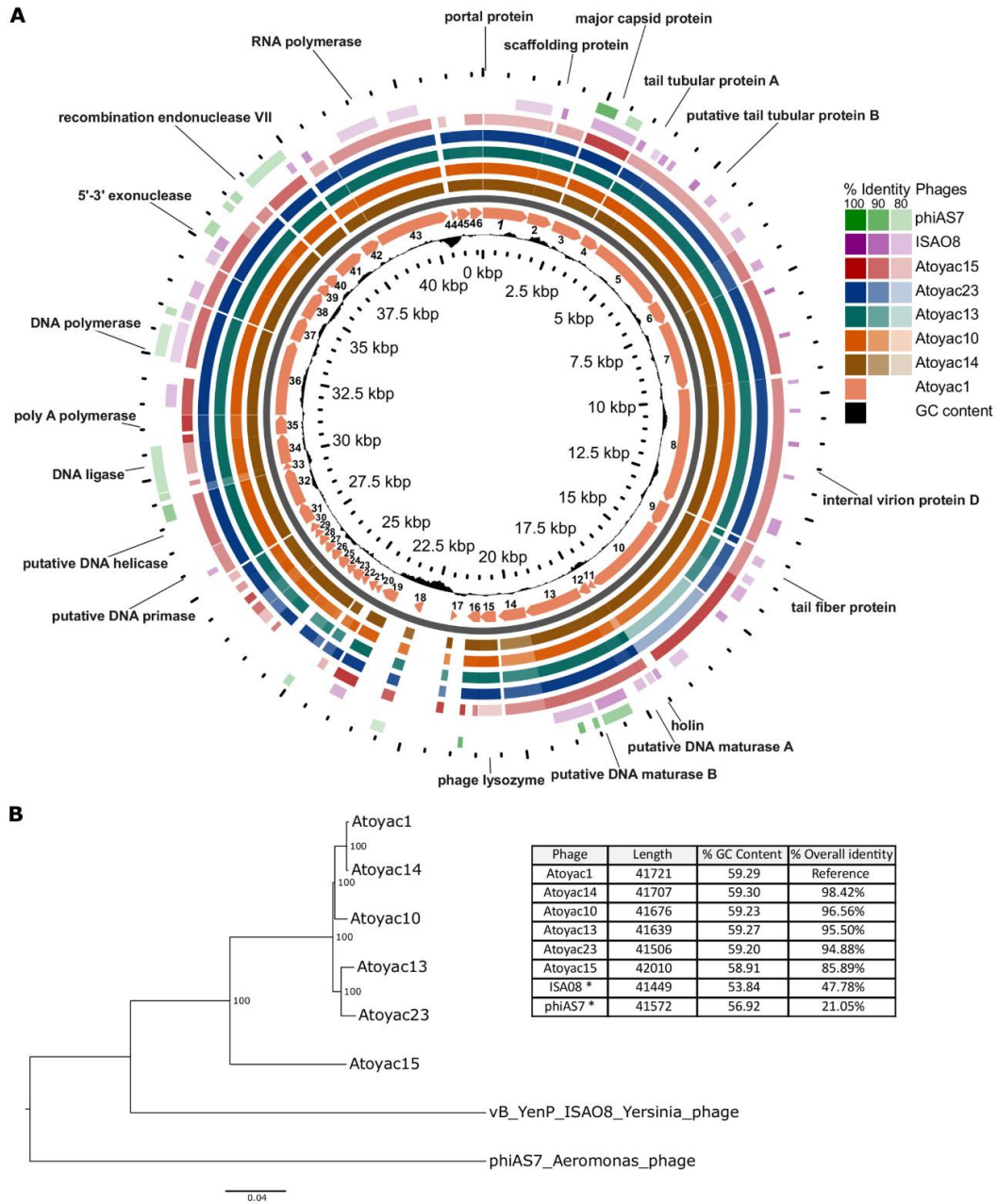


FIG 4 Genomic comparative analysis and phylogenetics of the Atoyac phage group. (A) The genomes of six Atoyac phages and two distant homologues are represented as color rings and indicated to the right of the figure. The genome of phage Atoyac1, represented by a map in the innermost color ring, was used as a reference to compare the ORF sequences of the other phages at the nucleotide level. Arrows in the map correspond to ORFs pointing toward the direction of their transcription. Functions identified in the genome are indicated outside the rings pointing to the corresponding Atoyac1 ORFs. Coordinates of the reference genome and the distribution of its GC content regarding the average (59%) are indicated inside the corresponding ring. The (Continued on next page)

ISAO8, GenBank accession no. [NC_028850.1](#)) and phiAS7 infecting *Aeromonas salmonicida* ([JN651747.1](#) [13]), which share 47.78 and 21.05% overall sequence identity with the genome of the phage Atoyac1, respectively (Fig. 4). We also compared the Atoyac phage genomes to that of the *E. coli* phage T7, a virulent archetype of the *Podoviridae* family. Despite the lack of extensive sequence similarity between the genomes at the nucleotide or protein level, the presence of a large nonhomologous and noncoding region was recognized in similar relative genome positions. A similar region has been identified in genomes of T7-like phages of different bacterial species and reported to contain multiple transcriptional promoters (14). In line with this observation, we identified a series of putative promoters in both strands of the large noncoding region of the Atoyac genomes (Fig. 4 and Table S1).

Search of Atoyac-like phages in metagenomics data. Atoyac phages were identified in different geographical regions in Mexico separated by ~3.9 to 2,525 km from each other, thus implying a broad distribution for this novel group of promiscuous viruses (Data Set S1). Since no other members of the group could be detected through homology searches in public genome databases, we decided to further investigate the distribution of the Atoyac phages by searching for similar sequences in available metagenomics data.

We sought to recover sequencing reads homologous to the Atoyac1 genome by analyzing metagenome data sets deposited in the NCBI Sequence Read Archive (SRA). Surprisingly, our analysis of more than 65,000 data sets using a recently reported searching strategy (see Materials and Methods) (15) did not identify metagenomes with sequencing reads matching the Atoyac reference. This result contrasts with those obtained when genomes of other two unrelated phages, crAssphage and T4, were used as reference and control for the search. Thousands of reads covering up to 100 and 86.5% of the crAssphage and T4 genomes were recruited in multiple samples (Data Set S3), suggesting absence or low abundance of Atoyac-like phages in the analyzed data sets. Alternatively, underexploration of environments inhabited by Atoyac-like phages (e.g., fresh waters) through metagenomics approaches could account for this outcome (8).

Experimental evolution of efficiency of plating. Since the Atoyac phages displayed higher EOPs in isolates of the *Aeromonas* group than in strains of the other susceptible genera, we took an experimental evolution approach to investigate the feasibility of reversing this trend in hosts with low plaque production efficiency. To diversify the set of hosts in the experiment, we chose *Yersinia* as representative of the *Enterobacteriaceae* family and *Pseudomonas* as the only non-*Enterobacteriaceae* genus identified as susceptible, besides *Aeromonas*. Isolates from the two selected genera recorded some of the lowest EOP values in the collection (Fig. 2A). Phages 1 and 23, displaying distinct EOPs in the bacterial isolates chosen for the experiment (*Pseudomonas* sp. strain mFC_SR91 and *Yersinia* sp. strain PAF_XS2_2; Fig. 2A), were selected as representatives of two subclusters identified within the Atoyac group (Fig. 4). Importantly, we allowed the phages to evolve by using ancestral bacteria at all times in the evolution experiments (see Materials and Methods).

Three individual plaques of Atoyac1 and Atoyac23, each representing a lineage, were serially propagated for 10 passages in a single host, *Pseudomonas* or *Yersinia* (see experimental design in Fig. S4, schemes A and B). To assess changes in plaque production over the course of the experiment, the EOP in *Pseudomonas* and *Yersinia* was calculated before the experiment and at propagation steps 5 and 10, using *Aeromonas* as

FIG 4 Legend (Continued)

level of sequence identity detected in the genomes (>79%) respecting the Atoyac1 ORFs is color coded and indicated in the figure. The two outermost rings correspond to representatives of the phage groups displaying the highest sequence similarity to genomes of the Atoyac group. (B) The genomic tree of the Atoyac phages and two distant homologues (neighbor joining) was constructed from the alignment of the genomes at the nucleotide level. The bootstrap support value (100 replicates) of the branches is shown in each node. The overall sequence identity percentage (i.e., over the entire genome length) of Atoyac phages and their distant homologues (marked with an asterisk regarding phage Atoyac1 (reference) is shown in the table, as well as the length and average GC content of each phage.

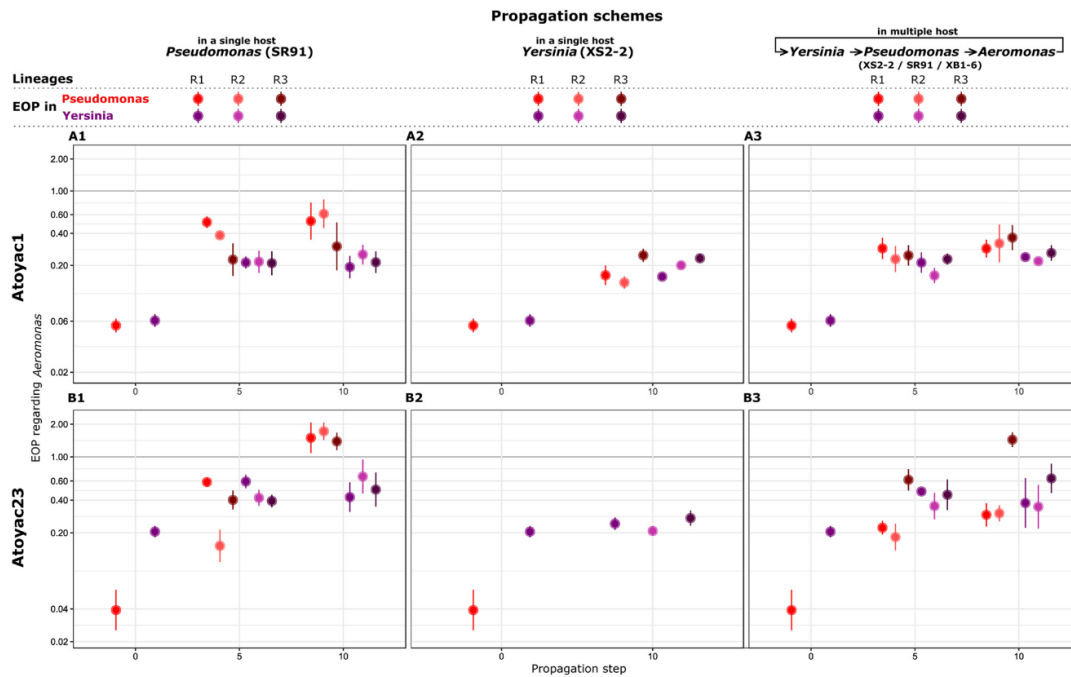


FIG 5 Experimental evolution of the efficiency of plating of two Atoyac phages. The charts represent the trajectories of the dynamics of efficiency of plating recorded during the experimental evolution of phages Atoyac1 and Atoyac23 under two different propagation schemes (see Fig. S4 in the supplemental material). Note that in these experiments the bacteria used in every transfer corresponds to the ancestral host (see Materials and Methods); hence, only the virus was allowed to evolve. The top row (panels A1, A2, and A3) shows the outcomes recorded for phage Atoyac1. The bottom row (panels B1, B2, and B3) illustrates the results for phage Atoyac23. In the first column (A1 and B1), the charts show the EOPs calculated in *Pseudomonas* SR91 (red) or *Yersinia* XS2-2 (purple), using *Aeromonas* XB1-6 as the reference, when the phages were serially propagated in *Pseudomonas* for 0, 5, and 10 passages. Values plotted in propagation step 0 correspond to the EOPs of the ancestral phage stocks. The second column (B2 and A2) indicates the EOP trajectories recorded, when the phages were sequentially propagated in *Yersinia*. No EOP values for *Pseudomonas* are shown in panel B2 at propagation step 10 because we could no longer count plaques of the phage 23 on this host. Finally, the third column (A3 and B3) displays the results obtained when the phages were propagated alternating hosts between *Pseudomonas* SR91, *Yersinia* XS2-2, and *Aeromonas* XB1-6, at every transfer. R1, R2, and R3 represent independent lineages from the same ancestral phage stock. EOPs for every lineage were calculated from five technical replicates each. The filled circles show the averages, and the bars represent the standard deviations for the lineages. The y axes in all the charts are displayed in a log scale. All the data plotted in the charts are presented in Table S1 in the supplemental material.

the reference. As expected, the ancestral stock of both phages produced more plaques in the *Aeromonas* strain than in the isolates of the other two genera (Fig. 5). Nevertheless, this difference in efficiency of plating narrowed down in both *Pseudomonas* and *Yersinia* when the phages were consecutively replicated in the *Pseudomonas* isolate, suggestive of collateral evolution of expanded host range (Fig. 5, panels A1 and B1). A similar outcome was observed when Atoyac1 was serially propagated in the *Yersinia* isolate (Fig. 5, panel A2); however, Atoyac23 did not show a substantial change in EOP in the *Yersinia* strain and considerably reduced its virulence toward the *Pseudomonas* isolate (decrease in plaque number and size) under the same propagation conditions (Fig. 5, panel B2).

In parallel to the serial replication in a single strain, we also propagated the Atoyac phages cyclically alternating their host between *Yersinia*, *Pseudomonas*, and *Aeromonas* to explore whether this condition alters the dynamics in EOP resulting from continuous replication using a single host (Fig. S4, scheme C). We found that changes in EOP for both phages over the course of the experiment resembled those recorded when the phages are exclusively propagated in the *Pseudomonas* strain or when Atoyac1 is continuously replicated in *Yersinia*, i.e., a marked reduction in the EOP gap compared to the reference *Aeromonas* isolate (Fig. 5, panels A3 and B3). Interestingly, under this propagation scheme, Atoyac23 did not lose infectivity toward the *Pseudomonas* strain despite being replicated

in the *Yersinia* isolate in 4 of the 10 propagation steps, thus implying that host alternation may be key to maintain promiscuity.

DISCUSSION

Phages are considered the most abundant and diverse biological entities on earth (16). Although numerous efforts exist to isolate and study new types of phages, a majority of them are targeted/biased toward isolating phages with a narrow host range (4), thus limiting our knowledge on phages with infection capabilities crossing between different taxonomic levels. These viruses are usually referred to as broad-host-range phages; however, we favor the use of “promiscuous” following the term applied to plasmids with mobilization and replication capabilities crossing multiple taxonomic levels (17). Moreover, the use of the term promiscuous allows distinction between viruses with hosts of different taxonomic origin and those infecting several strains of the same species, commonly referred to as broad range (5, 14). Promiscuous phages represent powerful models to further our understanding of phage-bacterium interactions and the evolution of host range; however, it is not yet clear how prevalent they can be (5).

Recent studies predicting phage-host relationships using various strategies have inferred interactions beyond the level of species or genus, suggesting that the abundance of promiscuous phages may be higher than currently postulated (4, 5, 7, 8). Previous attempts to isolate promiscuous phages primarily relied on enrichment methods using multiple bacteria of different species and mainly differ in either simultaneous or sequential addition of the potential hosts. These approaches led to the identification of phages infecting strains of two different genera or multiple species within a genus (9, 10). In these reports, either ATCC or other well-known type strains (e.g., *P. aeruginosa* PAO1 and *E. coli* K-12) were used as hosts (9, 10), which, although proven effective, limit the isolation to phages infecting strains that may not be representative of those existing in different environments. In contrast, our isolation strategy explored the use of a taxonomically diverse collection of bacterial isolates native to the sampled sites as potential hosts (see Fig. S1 in the supplemental material), under the rationale that isolating promiscuous phages is more likely when multiple bacterial strains used as indicators come from the same microbial communities they occur in.

Our isolation strategy allowed the identification of six phages, named Atoyac, infecting bacteria of the *Aeromonas*, *Hafnia*, *Yersinia*, *Escherichia*, *Serratia*, and *Pseudomonas* genera, which are part of the *Gammaproteobacteria* class. Interestingly, all the phages were retrieved from contaminated rivers. We speculate that in these environments there is a greater microbial diversity in high-density populations due to the presence of abundant resources available because of the constant human discharges, which would therefore favor the presence of phages capable of infecting different potential hosts. Nevertheless, we do not rule out that promiscuous phages could also be recovered from cleaner environments. We think our growth conditions could be applied to samples obtained from similar ecosystems. We encourage researchers to explore growth conditions that mimic the natural environment of the samples in order to retrieve bacterial isolates representative of the microbial diversity to maximize the possibility to find promiscuous phages. Following a similar approach regarding the use of native bacteria as potential hosts, other authors were able to isolate a phage from the Antarctic sea infecting strains of two genera from different proteobacteria classes (18).

Detection of phages infecting different species can be the result of a mix of multiple phages that have not been separated properly, leading to misidentification of a promiscuous host range (2). Host range assessment based on bactericidal effect (e.g., lysis spot) rather than productive infection can also lead to its overestimation (2). Here we addressed these issues through comprehensive evaluation of phage purity (Fig. S5) and determination of host range based on plaque production only. As previously recommended (2), our phage stock preparation included more than three rounds of plaque purification; additionally, the stocks were purified by cesium chloride (CsCl) density gradient prior to final host range determination (Fig. S1). Our experimental evolution

assays also involved 10 sequential passages in either single or multiple hosts of different genera (Fig. S4), conclusively demonstrating intergeneric propagation. In our experiments, we observed consistent evidence of phage stock purity (Fig. S5), including the presence of a well-defined band in the CsCl gradients, similar DNA restriction patterns and genome assembly sizes between phages, uniformity in plaque morphology, and homogeneity in virion morphology within and between samples.

Our results show that the six Atoyac phages have similar host ranges, consistent with their genomic similarity, and indicate that even the most divergent member of the group (Atoyac phage 15, displaying 86% overall sequence similarity to phage 1) is able to infect six bacterial genera. Interestingly, we found a good correlation between the clustering of the Atoyac phages based on the whole-genome comparison (Fig. 4) and that inferred from the EOP and host range profiles (Fig. 2A), suggesting that sequence variability largely accounts for the differential infection capability displayed by the phages. We identified a pair of putative structural genes encoding the tail fiber and a glycosidase superfamily protein, and two adjacent regions, one intergenic featuring putative promoters and one encoding a cluster of small ORFs of unknown function, that could account for the differences observed in the infection phenotypes given their sequence variation within the group. Mutations in virion structural proteins, particularly those that are components of the tail, have been reported as responsible for shifts in the host range of several phages, including in experimental evolution studies (19–21). Variable regions encoding accessory ORFs have also been documented in phage genomes (22); however, these have not been extensively characterized in terms of their functional impact on the phage biology.

Although the Atoyac phages infect six bacterial genera, the EOP analysis showed they produce more lytic plaques on *Aeromonas* isolates than in the other susceptible bacteria, suggesting that this phage group is better adapted to infect the *Aeromonas* genus (Fig. 2). The highly similar GC content between the Atoyac phages (59%) and *Aeromonas* (58 to 62% [23]) genomes and larger lytic plaques formed in the *Aeromonas* strains further support this notion (Fig. S2). An *Aeromonas* strain (PIA_XB1_6) was selected as the propagation host for the Atoyac phages because it was identified as highly sensitive to phage infection, along with other *Aeromonas*, in our initial screening to detect susceptible strains from the water samples. It is worth noting that propagation in PIA_XB1_6 prior to the final determination of host range could have also enhanced the EOP in this genus as a consequence of a potential domestication effect. Although we limited the number of passages to avoid this effect, it represents a potential artifact inherent to standard protocols to obtain pure phage stocks, a key requirement before assessing host range (2). Since the *Aeromonas* genus is known to be abundant in various aquatic environments like rivers (24), we hypothesize it serves as the primary host of the Atoyac phages, whereas isolates from the other susceptible genera, also present in rivers (25, 26), act as secondary hosts. We speculate that the presence of promiscuous phages in rivers may be related to the dynamics of this type of environment which is commonly influenced by strong seasonality (marked seasons of drought and heavy rain) that impact the composition of microbial communities diversifying it regularly (26, 27). Within this context, a promiscuous host range would be advantageous for the phage, as it would allow the use of alternative hosts to spread and remain in the community even if the primary host is not within its reach.

We observed statistically significant differences in the Atoyac phages' EOPs between *Aeromonas* and other susceptible genera but not between *Aeromonas* subclusters (see Data Set S2 in the supplemental material), suggesting that infection of distantly related bacteria involves a biological cost in terms of lytic plaque production. This notion is consistent with observations on other broad-host-range phages that exhibit substantial variation in EOPs when infecting distantly related hosts (9, 10). Variation in the efficiency of transfer or replication between distant hosts has also been reported for promiscuous plasmids (28), implying that the ability to spread among multiple taxonomic backgrounds does not come without a cost and that such trade-off is not exclusive of a certain type of mobile genetic element. However, despite the variation in the efficiency of plating, the

Atoyac phages did not show a considerable reduction in the EOPs in most of their hosts, even among the different genera (~85% of the isolates showed less than 100-fold EOP variations; Fig. 2). Bacteria often have intracellular defense strategies that allow them to limit infection by phages and other mobile genetic elements, the restriction-modification system and the CRISPR-cas system being some of the most widespread antiphage defenses in bacterial genomes. These molecular mechanisms have been shown to dramatically impact the EOP by several orders of magnitude (2, 12). Therefore, it is possible that one or multiple putative ORFs of unknown function that were found during the genomic annotation and comparison in the Atoyac phages could encode mechanisms to counteract antiphage defense systems. A previous study by Jensen et al. (9) showed that broad-host-range phages that do not display significant decreases in their EOPs in different hosts (such as *Sphaerotilus natans* and *P. aeruginosa*) are insensitive to type I and II restriction enzymes. Similarly, the broad-host-range podophage T7 encodes the Orc protein that inhibits type I restriction enzymes as a mechanism to counterattack bacterial defense (29). Hence, we believe that by studying this type of phages, interesting proteins with biotechnology applications can be found, and it remains in our interest to study more in detail what the ORFs of unknown function in the Atoyac phages encode.

The Atoyac phages were isolated from river water samples obtained in different geographic regions across Mexico, suggesting a wide distribution of the group in similar environmental samples. Surprisingly, our search in metagenomics data did not identify sequencing reads mapping to the genome of the phage Atoyac1. While this outcome suggests the absence of the query genome in the explored data sets, it may also be indicative of the low abundance of Atoyac-like phages in the analyzed samples. The searching strategy used in this work involves a data set subsampling step to make the search through the SRA metagenomics data more efficient (15) with the cost of being susceptible to some false-negative results, particularly when the sequencing reads of certain taxa are in low abundance. Our study shows that culture-based approaches are still a good alternative to gain new insights on the virosphere diversity with the additional advantage of leading to the isolation of specimens that can be further investigated in the lab.

The comparative genome analysis identified the Atoyac phages as a novel viral group. Interestingly, the most similar genomes detected outside the group correspond to distantly related phages infecting either *Yersinia* or *Aeromonas* species, two of the bacterial groups identified as susceptible to infection by Atoyac phages in this work. Similarity matches to ORFs of phages with hosts of different genera (e.g., *Pseudomonas*, *Serratia*, *Escherichia*) can be detected by comparing sequences of conserved proteins encoded in the structural or replication modules of the Atoyac genomes; however, these matches feature low levels of sequence similarity, again indicating distant relationships (data not shown). Although we hypothesize *Aeromonas* as the primary host of the Atoyac phages, they display higher sequence similarity to the genome of the *Yersinia enterocolitica* phage ISAO8 (~48%) than to that of the *Aeromonas salmonicida* phage phiAS7 (~21%). It is unclear whether the phages ISAO8 and phiAS7 are specialized to infect their reported hosts or have a promiscuous host range as the phages of the Atoyac group. Further characterization of ISAO8 and phiAS7 could provide valuable insights into the evolution of the Atoyac phages' promiscuous host range in the future. Nevertheless, our study highlights the importance of comprehensively characterizing the host range of a phage by including taxonomically diverse hosts in the panel of tested strains. In the case presented here, the Atoyac phages could have been reported as *Aeromonas* "specific," given the large number of host strains identified within this genus, thus veiling the broader infection scope of the group.

Isolating promiscuous phages represents an unique opportunity to explore host range evolution among phylogenetically distant hosts; hence, we exploited this model to investigate the dynamics of infection of two Atoyac phages on nonoptimal hosts (*Pseudomonas* and *Yersinia*) using different propagation schemes. Our results show that successive propagation on either single or alternate hosts has a strong effect on

the EOP dynamics and its evolutionary trajectory, albeit the phage used is a critical factor in the outcome. In most cases, the serial propagation promoted EOP optimization in the inefficient hosts while maintaining the promiscuous host range, thus suggesting this is a well-conserved feature. The propagation of Atoyac23 exhibited the most contrasting results in the experiment. Continuous replication in *Pseudomonas* allowed the phage to reach an EOP and develop a plaque morphology, similar to those observed in *Aeromonas*, the initial most efficient host. In contrast, successive propagation on *Yersinia* did not have a substantial impact on the EOP in this host. Such dissimilar outcomes could be associated with the phage initial EOP in both hosts, i.e., Atoyac23 marked lower efficiency in *Pseudomonas* regarding *Yersinia*, and hence its margin for optimization. This notion is consistent with a recent experimental evolution study of a generalist cyanophage which acquired a larger number of evolutionary changes with a positive effect on infection efficiency when continuously propagated on a suboptimal host rather than an optimal host (21). Alternatively, evolutionary changes improving EOP in *Yersinia* might require more propagation steps to emerge and fix in the population.

Serial propagation of Atoyac1 in *Pseudomonas* improved its EOP in this host as well as in the *Yersinia* isolate, and vice versa. Phage Atoyac23, in contrast, displayed a substantial reduction in virulence toward *Pseudomonas* as a consequence of its continuous propagation in the *Yersinia* isolate. Replication in alternate hosts prevented the emergence of this effect, highlighting the importance of propagation in multiple hosts to maintain promiscuity in some phages. We speculate that similar propagation patterns could occur in certain environments featuring high microbial diversity, thus leading to preservation of phages with a promiscuous host range. Our results show that the adaptation outcome from continuous propagation on single or alternate hosts is largely dependent on the phage studied, even between highly similar phages, making the evolutionary trajectories difficult to predict.

MATERIALS AND METHODS

Sample collection and coliform count. A dozen samples were collected in 1-liter sterile flasks from different rivers and sewage in central, northern, and southern Mexico. Physicochemical parameters such as temperature, pH, dissolved oxygen, etc., were measured *in situ* using a HANNA multiparametric HI9828 instrument (see Data Set S1 in the supplemental material). The recorded values were used as a guide for exploring our growth conditions in the laboratory. Estimation of water pollution in the different samples was performed by thermotolerant coliform counts on modified mTEC media, following the membrane filtration protocol previously described (30). Colonies producing a blue pigment (showing beta-glucuronidase activity) were counted, and the numbers of CFU per milliliter were calculated based on the dilution of the samples; these experiments were performed in triplicate. Sampling sites with coliform counts of $\geq 1 \times 10^3$ CFU/ml were classified as contaminated (Data Set S1).

Bacterial isolation. A diverse collection of bacterial isolates was assembled from the collected samples. Samples were plated onto different media: LB agar (Lennox), BHI (brain heart infusion) agar, PIA (*Pseudomonas* isolation agar), PAF (*Pseudomonas* agar F), PAP (*Pseudomonas* agar P), NAA (nutrient-poor medium) agar, mTEC (membrane thermotolerant *E. coli*) agar, mFC (membrane fecal coliform) agar, TSI (triple sugar iron) agar, and MC (MacConkey) agar, all manufactured by Difco and prepared following the manufacturer's instructions (31) and then incubated at 30°C for 24 h. After incubation, some colonies were picked to represent a variety of morphologies, sizes, and colors. The selected colonies were then purified by three consecutive passages on the same medium they were isolated from. Pure isolates were grown overnight in BHI broth at 30°C and stored at -80°C in 30% glycerol. The final collection was composed of around 600 isolates coming from the 12 different sampling sites.

Phage isolation. Bacterial lawns of the isolates in our collection were prepared by mixing 80 to 250 μ l of overnight cultures with 3.5 ml of top agar (1% peptone, 0.5% NaCl, 0.7% Bacto agar, 10 mM MgSO₄), and plating the mixture on T Φ (1% peptone, 0.5% NaCl, 1.1% Bacto agar), LB, BHI, mFC, TSI, MC, and MC without salts media (31) to find the best conditions for identification of phage plaques (the media used for optimal growth and visualization of plaques in the strains identified as susceptible to the Atoyac phages are listed in Data Set S2). Aliquots of the water samples used to isolate bacteria were centrifuged, filtered (MF-Millipore membrane filter, 0.22- μ m pore size), serially diluted, and then spotted on the bacterial lawns to identify strains allowing the formation of isolated lytic plaques, indicative of susceptibility to phage infection. Plates were incubated overnight at 30°C. To aid the visualization of the lytic plaques, rosolic acid (0.01%), phenol red (0.0024%), or bromophenol blue (0.0024%) were used in some of the media (T Φ , LB, BHI, TSI, MC, and mFC). The concentrations of dye used are based on those reported for Difco's media such as mFC and TSI (31).

Cross-infection assays were then used to identify candidate promiscuous phages (i.e., phages that

could infect a wide diversity of isolates). Briefly, individual plaques were picked with a sterile toothpick and streaked onto bacterial lawns of the strains previously identified as phage susceptible. To assess whether infection of a given strain affected infectivity in other hosts, the cross-infection assays were performed in two different ways: simultaneously, i.e., one selected plaque was transferred to the other bacterial isolates in parallel, and sequentially, i.e., a plaque produced in one tested bacterial isolate was transferred to the next potential host (see Fig. S1 in the supplemental material). Phages infecting multiple isolates were considered candidates to promiscuous phages, but only those with the broadest host range were selected for further characterization. These phages were then purified by four sequential passages in the strain PIA_XB1_6 (identified taxonomically as *Aeromonas* sp.), in which the phages exhibited some of the highest titers, largest plaque size, and a short plaque generation period (~5 h) during the first screening to identify phage-susceptible strains; hence, it was considered the optimal propagation host. Additional purification passages were avoided to reduce a possible domestication effect of phages in this host. High-titer stocks were generated from the last purification pass as described previously (22).

Taxonomic identification of bacterial isolates to genus level. Bacterial isolates that were susceptible to infection by phages were grown on LB overnight at 30°C. Genomic DNA was extracted from the cultures by following the protocol previously reported by Chen and Kuo (32). The extracted DNA was used as the template for PCR amplifying the genes *gyrB* and *rpoD*. We favored the use of the housekeeping genes *rpoD* and *gyrB* over the 16S rRNA gene as markers because the former have a higher evolutionary rate and have shown better results to resolve complex groups such as *Pseudomonas*, enterobacteria, and *Aeromonas* to the intragenetic level (33–35).

Oligonucleotides to amplify the *gyrB* (forward, 5'-GGCGGTAARTTYGAYGATAAC-3'; reverse, 5'-TAGCCTGGTCTTACGGTT-3''; reported here) were used in those isolates that displayed phenotypic traits similar to those of the *Enterobacteriaceae* family (e.g., growth in MC plates). Primers flanking *rpoD* (forward, 5'-ATYGAAATCGCCAARCG-3'; reverse, 5'-CGGTGATKCTCCTGA-3' [36]) were used with the rest of the bacterial isolates. PCR assays were performed using Thermo Scientific *Taq* DNA polymerase with 0.2 μ M each primer and the following cycling conditions: 95°C for 3 min; 30 cycles with 1 cycle consisting of 95°C for 30 s, 52°C for 1 min for *gyrB* and 56°C for 1 min for *rpoD*, 72°C for 1 min; and 72°C for 10 min. PCR products were purified using MultiScreen PCR μ 96 and sequenced by Sanger sequencing in Macrogen, Inc. The bacterial genera were determined by performing a BLASTn search of the obtained sequences against the NCBI database in order to identify the genera of the closest matches. The homologous sequences (using only type strains as the reference) were aligned using Clustal omega and used to generate neighbor-joining trees (37).

Purification of promiscuous phages. Phage stocks were treated with DNase I and RNase I (1 μ g/ml) and incubated for 30 min at 37°C. Phage particles were then precipitated with 16% (wt/vol) polyethylene glycol (PEG) 8000 and 1.4 M NaCl overnight at 4°C and centrifuged at 8,000 relative centrifugal force (RCF) for 20 min. The supernatants were discarded, and the pellet was resuspended in 1 ml of SM buffer (50 mM Tris HCl [pH 8], 10 mM MgSO₄, 100 mM NaCl, and 0.01% gelatin). In order to remove the PEG from the stocks, an equal volume of chloroform was added to the tubes and centrifuged at 8,000 RCF for 5 min. The aqueous phase was recovered and purified by ultracentrifugation in a discontinuous CsCl gradient as previously described (38). Finally, the pure phage stocks were dialyzed to remove the salts and stored at 4°C.

Determination of phage host range and efficiency of plating. Cross-infection assays (see "Phage isolation" section above) were used to characterize the host range of the promiscuous phages based on their ability to produce isolated plaques in the different bacterial isolates. Phages capable of infecting different taxonomic genera were classified as promiscuous and named Atoyac, which means river in Náhuatl (an indigenous language of Mexico). To determine the EOP of the Atoyac phages, serial dilutions of the CsCl-purified stocks were spotted onto the susceptible isolates of the bacterial collection. The EOP was calculated as the average of the titer of the phage in a given strain divided by the titer in the isolate that was used to initially propagate the phage (PIA_XB1_6). Each plaque assay was performed with five replicates.

Statistical analysis. In order to determine whether the data of efficiency of plating was normally distributed, histograms, quantile-quantile (QQ) plots, and Shapiro-Wilk tests were used taking as input the EOP values of each phage on the different bacterial genera. As the data were not normally distributed due to the nature of the samples, the nonparametric Mann-Whitney U test was used and *P* values were corrected for multiple comparisons using the Bonferroni correction. Corrected *P* values lower than 0.05 were considered statistically significant (Data Set S2). For the experimental evolution data, one-way analysis of variance (ANOVA) tests were used to compare the difference of EOP for each condition among the different time points. *P* values lower than 0.05 were considered statistically significant (Data Set S4).

Virion morphology. The virion morphology was determined by negative-contrast stain transmission electron microscopy as previously described by Cazares et al. (22). Briefly, aliquots of the pure phage stocks were deposited on a carbon-coated copper grid and stained with 2% uranyl acetate. Grids were examined under a JEM-2000 transmission electron microscope at 80 kV.

Phage DNA extraction and genome sequencing. DNA from the six Atoyac phages was extracted from 500 μ l of CsCl-purified phage stocks following the protocol previously described (38), excluding the step with proteinase K. Purity and concentration values were measured by spectrometry (Nanodrop 2000). Aliquots of 500 ng of DNA were used to determine the size of the genome by restriction digestion. Finally, the DNA for the different phages was sequenced by Illumina (paired-end) technology in the Unidad Universitaria de Secuenciación Masiva y Bioinformática (UUSMB)-Universidad Nacional Autónoma de México (UNAM).

Phage genome analysis. The reads were adapter trimmed and quality filtered using Trimmomatic v.0.36 (39). The filtered reads were then assembled using Velvet v.1.2.10 (40) and mapped against the *de novo* assemblies using BWA v.0.7.17 (41). The files with the mapped reads were then used to correct the assemblies using Pilon v.1.22 (42). The genomes were preliminarily annotated with Prokka v.1.13 (43) and reannotated manually based on protein homology searches and conserved domains using InterProScan 5.30-69.0, PSI-BLAST, and CD Search (44–46).

The genomes were used for comparative genomics analysis of the Atoyac phages and their homologous, 4vB_YenP_JSAO8 (GenBank accession no. [NC_028850.1](#)) and phiAS7 ([JN651747.1](#)) phages. The coding sequence (CDS) regions of Atoyac1 phage were used as reference to compare with the other phages using the GView Server (<https://server.gview.ca/> [47]) and generate a comparative circular map (BLAST Atlas). Similarly, Atoyac phages were aligned using Mauve v.2.4.0 with the default parameters to create a comparative linear map and closely visualize the main variable regions of the genomes (48). Finally, the phages' genomes were globally aligned with Clustal omega and a BioNeighbor-Joining tree was constructed based on the alignment (37).

Promoter prediction. The *in silico* promoter search for the intergenic region was performed using the phage Atoyac1 sequence (from nucleotide 21377 to 23783), as the reference of the group, with the promoter prediction tools (for both strands): Neural Network Promoter Prediction (NNPP), PromoterHunter, and BPROM through their web servers (www.fruitfly.org/seq_tools/promoter.html, www.phisite.org/main/index.php?nav=tools&nav_sel=hunter, and www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb [49–51]).

Search of Atoyac genomes in metagenomics data. The genome of phage Atoyac1 was used as a reference to search for homologues in metagenomics data deposited in the NCBI Sequence Read Archive (SRA). Whole shotgun metagenomes available from SRA and previously curated with PARTIE (52) were mapped to the reference genome using bowtie2 (53) with default parameters through the Search SRA Gateway (<https://www.searchsra.org/>) (15, 54, 55). Only alignments of reads > 50 bp were considered. The number of mapped reads and coverage of the alignments were determined from the sorted and indexed BAM files using Samtools v.1.7 (15, 48, 56).

Experimental evolution of efficiency of plating. Our experimental evolution strategy consisted of the sequential propagation of the phages Atoyac1 and Atoyac23 by plaque assays under two different schemes (Fig. S4). In the first scheme, phages were propagated in a single host, using either the *Pseudomonas* strain SR91 or *Yersinia* XS2-2, in which both phages recorded low EOP regarding the *Aeromonas* strain XB1-6 used as the reference to calculate the efficiencies. In the second scheme, phages were propagated in multiple hosts, cyclically alternating the use of the strains XS2-2, SR91, and XB1-6. Atoyac1 and -23 were selected as the most divergent representatives of two subgroups identified within the Atoyac group (excluding Atoyac15 which represents the outlier [Fig. 4]) and because they exhibited distinct EOP profiles in the bacterial hosts chosen for the experiments. In both evolutionary schemes, the phages were propagated for 10 passages starting from three isolated plaques (lineages) obtained by plating the ancestral phage stock on bacterial lawns of the selected hosts. In every transfer, overnight cultures of the ancestral bacteria, grown from frozen glycerol stocks, were used to prepare the lawns utilized in the plaque assays; therefore, only the virus was allowed to evolve in the experiments. EOPs of the phages were calculated at propagation steps 0, 5, and 10, using high-titer phage stocks generated in these stages and plating them onto lawns of the ancestral *Pseudomonas*, *Yersinia*, and *Aeromonas* strains. The titer recorded in *Aeromonas* in the corresponding propagation step was used as reference to determine the EOP, which was calculated from five technical replicates (see Fig. S4 for experimental design).

Data availability. The Illumina-sequenced genomes and annotations of the Atoyac1, Atoyac10, Atoyac13, Atoyac14, Atoyac15, and Atoyac23 phages were deposited in GenBank under accession numbers [MT682386](#) to [MT682391](#). The six Atoyac phages and three of their hosts (*Aeromonas* PIA_XB1_6, *Yersinia* PAF_XS2_2, and *Pseudomonas* mFC_SR91) were deposited in the Félix d'Hérelle Reference Center for Bacterial Viruses.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

DATA SET S1, XLSX file, 0.02 MB.

DATA SET S2, XLSX file, 0.1 MB.

DATA SET S3, XLSX file, 0.1 MB.

DATA SET S4, XLSX file, 0.04 MB.

FIG S1, TIF file, 0.3 MB.

FIG S2, TIF file, 0.8 MB.

FIG S3, TIF file, 0.5 MB.

FIG S4, TIF file, 1 MB.

FIG S5, TIF file, 2.5 MB.

TABLE S1, XLSX file, 0.01 MB.

ACKNOWLEDGMENTS

D.C. gratefully acknowledges the Programa de Doctorado en Ciencias Biomédicas, UNAM, and the Ph.D. scholarship 586079 from Consejo Nacional de Ciencia y Tecnología (CONACYT, México). A.C. has been supported by the EMBL-EBI/Wellcome Trust Sanger Institute Join Post-Doctoral Fellowship Program (ESPOD). P.V. acknowledges funding from DGAPA-PAPIIT/UNAM IN206318 and CONACYT 23_A1-S-11242. G.G. acknowledges funding from CONACYT CB 255255. W.F. acknowledges funding from Cambridge Trust-CONACYT 706017.

We thank María de Lourdes Rojas-Morales from Microscopía Electrónica, CINVESTAV, for technical assistance in our electron microscopy studies and Javier Rivera Campos from CCG, UNAM, for technical support with lab experiments.

We declare that we have no competing interests.

D.C., A.C., W.F., and P.V. conceptualized the study. D.C., A.C., and W.F. drafted the manuscript with input from the other authors. D.C. isolated the phages and carried out the characterization experiments and data analysis and visualization. A.C. contributed to the genome and metagenome data analysis, virion purification, and characterization by electron microscopy. W.F. performed the statistical analyses and contributed to virion purification, phage DNA extraction, and characterization by electron microscopy and EOP. G.G. contributed by supervising the work. R.A.E. contributed to the metagenomics data analysis and revision of the manuscript. P.V. was responsible for funding acquisition and contributed to the phylogenomic analysis.

REFERENCES

- Ackermann H-W, DuBow MS. 1987. Viruses of prokaryotes, vol 1. General properties of bacteriophages, p 49–85. CRC Press, Boca Raton, FL.
- Hyman P, Abedon ST. 2010. Bacteriophage host range and bacterial resistance. *Adv Appl Microbiol* 70:217–248. [https://doi.org/10.1016/S0065-2164\(10\)70007-1](https://doi.org/10.1016/S0065-2164(10)70007-1).
- Matsuzaki S, Rashed M, Uchiyama J, Sakurai S, Ujihara T, Kuroda M, Ikeuchi M, Tani T, Fujieda M, Wakiguchi H, Imai S. 2005. Bacteriophage therapy: a revitalized therapy against bacterial infectious diseases. *J Infect Chemother* 11:211–219. <https://doi.org/10.1007/s10156-005-0408-9>.
- Koskella B, Meaden S. 2013. Understanding bacteriophage specificity in natural microbial communities. *Viruses* 5:806–823. <https://doi.org/10.3390/v5030806>.
- de Jonge PA, Nobrega FL, Brouns SJJ, Dutilh BE. 2019. Molecular and evolutionary determinants of bacteriophage host range. *Trends Microbiol* 27:51–63. <https://doi.org/10.1016/j.tim.2018.08.006>.
- Drulis-Kawa Z, Majkowska-Skrobek G, Maciejewska B, Delattre A-S, Lavigne R. 2012. Learning from bacteriophages – advantages and limitations of phage and phage-encoded protein applications. *Curr Protein Pept Sci* 13:699–722. <https://doi.org/10.2174/138920312804871193>.
- Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. 2016. Computational approaches to predict bacteriophage–host relationships. *FEMS Microbiol Rev* 40:258–272. <https://doi.org/10.1093/femsre/fuv048>.
- Paez-Espino D, Eloe-Fadros EA, Pavlopoulos GA, Thomas AD, Huntemann M, Mikhailova N, Rubin E, Ivanova NN, Kyrpides NC. 2016. Uncovering earth's virome. *Nature* 536:425–430. <https://doi.org/10.1038/nature19094>.
- Jensen EC, Schrader HS, Rieland B, Thompson TL, Lee KW, Nickerson KW, Kokjohn TA. 1998. Prevalence of broad-host-range lytic bacteriophages of *Sphaerotilus natans*, *Escherichia coli*, and *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 64:575–580. <https://doi.org/10.1128/AEM.64.2.575-580.1998>.
- Yu P, Mathieu J, Li M, Dai Z, Alvarez PJJ. 2016. Isolation of polyvalent bacteriophages by sequential multiple-host approaches. *Appl Environ Microbiol* 82:808–815. <https://doi.org/10.1128/AEM.02382-15>.
- Ballesté E, Pascual-Benito M, Martín-Díaz J, Blanch AR, Lucena F, Muniesa M, Jofre J, García-Aljaro C. 2019. Dynamics of crAssphage as a human source tracking marker in potentially faecally polluted environments. *Water Res* 155:233–244. <https://doi.org/10.1016/j.watres.2019.02.042>.
- Kutter E. 2009. Phage host range and efficiency of plating. *Methods Mol Biol* 501:141–149. https://doi.org/10.1007/978-1-60327-164-6_14.
- Kim JH, Son JS, Choresca CH, Shin SP, Han JE, Jun JW, Kang D-H, Oh C, Heo S-J, Park SC. 2012. Complete genome sequence of bacteriophage phiAS7, a T7-like virus that infects *Aeromonas salmonicida* subsp. *salmonicida*. *J Virol* 86:2894–2895. <https://doi.org/10.1128/JVI.07131-11>.
- Kawasaki T, Shimizu M, Satsuma H, Fujiwara A, Fujie M, Usami S, Yamada T. 2009. Genomic characterization of *Ralstonia solanacearum* phage phiRSB1, a T7-like wide-host-range phage. *J Bacteriol* 191:422–427. <https://doi.org/10.1128/JB.01263-08>.
- Levi K, Rynga M, Abeysinghe E, Edwards RA. 2018. Searching the Sequence Read Archive using Jetstream and Wrangler, p 1–7. *In* Proceedings of the Practice and Experience in Advanced Research Computing. Association for Computing Machinery, New York, NY.
- Suttle CA. 2007. Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol* 5:801–812. <https://doi.org/10.1038/nrmicro1750>.
- Krishnapillai V. 1988. Molecular genetic analysis of bacterial plasmid promiscuity. *FEMS Microbiol Rev* 4:223–237. <https://doi.org/10.1111/j.1574-6968.1988.tb02744.x>.
- Luhtanen A-M, Eronen-Rasimus E, Oksanen HM, Tison J-L, Delille B, Diekmann GS, Rintala J-M, Bamford DH. 2018. The first known virus isolates from Antarctic sea ice have complex infection patterns. *FEMS Microbiol Ecol* <https://doi.org/10.1093/femsec/fiy028>.
- Heineman RH, Springman R, Bull JJ. 2008. Optimal foraging by bacteriophages through host avoidance. *Am Nat* 171:E149–E157. <https://doi.org/10.1086/528962>.
- Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR, Quail M, Smith F, Walker D, Libberton B, Fenton A, Hall N, Brockhurst MA. 2010. Antagonistic coevolution accelerates molecular evolution. *Nature* 464:275–278. <https://doi.org/10.1038/nature08798>.
- Enav H, Kirzner S, Lindell D, Mandel-Gutfreund Y, Béjà O. 2018. Adaptation to sub-optimal hosts is a driver of viral diversification in the ocean. *Nat Commun* 9:4698. <https://doi.org/10.1038/s41467-018-07164-3>.
- Cazares A, Mendoza-Hernández G, Guarneros G. 2014. Core and accessory genome architecture in a group of *Pseudomonas aeruginosa* Mu-like phages. *BMC Genomics* 15:1146. <https://doi.org/10.1186/1471-2164-15-1146>.
- Lorén JG, Farfán M, Miñana-Galbis D, Fusté MC. 2010. Prediction of whole-genome DNA G+C content within the genus *Aeromonas* based on housekeeping gene sequences. *Syst Appl Microbiol* 2010.03.007. <https://doi.org/10.1016/j.syapm.2010.03.007>.
- Janda JM, Abbott SL. 2010. The genus *Aeromonas*: taxonomy, pathogenicity, and infection. *Clin Microbiol Rev* 23:35–73. <https://doi.org/10.1128/CMR.00039-09>.
- Leclerc H, Schwartzbrod L, Dei-Cas E. 2002. Microbial agents associated with waterborne diseases. *Crit Rev Microbiol* 28:371–409. <https://doi.org/10.1080/1040-840291046768>.

26. Ochoa-Sánchez LE, Vinuesa P. 2017. Evolutionary genetic analysis uncovers multiple species with distinct habitat preferences and antibiotic resistance phenotypes in the complex. *Front Microbiol* 8:1548. <https://doi.org/10.3389/fmicb.2017.01548>.
27. Breton-Deval L, Sanchez-Flores A, Juárez K, Vera-Estrella R. 2019. Integrative study of microbial community dynamics and water quality along the Apatlaco River. *Environ Pollut* 255:113158. <https://doi.org/10.1016/j.envpol.2019.113158>.
28. Datta N, Hedges RW, Shaw EJ, Sykes RB, Richmond MH. 1971. Properties of an R factor from *Pseudomonas aeruginosa*. *J Bacteriol* 108:1244–1249. <https://doi.org/10.1128/JB.108.3.1244-1249.1971>.
29. Häuser R, Blasche S, Dokland T, Haggård-Ljungquist E, von Brunn A, Salas M, Casjens S, Molineux I, Uetz P. 2012. Bacteriophage protein–protein interactions. *Adv Virus Res* 83:219–298. <https://doi.org/10.1016/B978-0-12-394438-2.00006-2>.
30. Ciebin BW, Brodsky MH, Eddington R, Horsnell G, Choney A, Palmateer G, Ley A, Joshi R, Shears G. 1995. Comparative evaluation of modified m-FC and m-TEC media for membrane filter enumeration of *Escherichia coli* in water. *Appl Environ Microbiol* 61:3940–3942. <https://doi.org/10.1128/AEM.61.11.3940-3942.1995>.
31. Becton Dickinson and Company. 2009. Difco & BBL manual: manual of microbiological culture media. Becton Dickinson and Company, Sparks, MD.
32. Chen WP, Kuo TT. 1993. A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. *Nucleic Acids Res* 21:2260. <https://doi.org/10.1093/nar/21.9.2260>.
33. Yamamoto S, Kasai H, Arnold DL, Jackson RW, Vivian A, Harayama S. 2000. Phylogeny of the genus *Pseudomonas*: intragenetic structure reconstructed from the nucleotide sequences of *gyrB* and *rpoD* genes. *Microbiology* 146:2385–2394. <https://doi.org/10.1099/00221287-146-10-2385>.
34. Fukushima M, Kakinuma K, Kawaguchi R. 2002. Phylogenetic analysis of *Salmonella*, *Shigella*, and *Escherichia coli* strains on the basis of the *gyrB* gene sequence. *J Clin Microbiol* 40:2779–2785. <https://doi.org/10.1128/jcm.40.8.2779-2785.2002>.
35. Navarro A, Martínez-Murcia A. 2018. Phylogenetic analyses of the genus *Aeromonas* based on housekeeping gene sequencing and its influence on systematics. *J Appl Microbiol* 125:622–631. <https://doi.org/10.1111/jam.13887>.
36. Mulet M, Benassar A, Lalucet J, García-Valdés E. 2009. An *rpoD*-based PCR procedure for the identification of *Pseudomonas* species and for their detection in environmental samples. *Mol Cell Probes* 23:140–147. <https://doi.org/10.1016/j.mcp.2009.02.001>.
37. Sievers F, Higgins DG. 2014. Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol Biol* 1079:105–116. https://doi.org/10.1007/978-1-62703-646-7_6.
38. Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
39. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
40. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
41. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
42. Walker BJ, Abeel T, Shea T, Priest M, Boueiliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
43. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
44. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>.
45. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
46. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45:D200–D203. <https://doi.org/10.1093/nar/gkw1129>.
47. Petkau A, Stuart-Edwards M, Stothard P, Van Domselaar G. 2010. Interactive microbial genome visualization with GViz. *Bioinformatics* 26:3125–3126. <https://doi.org/10.1093/bioinformatics/btq588>.
48. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. <https://doi.org/10.1371/journal.pone.0011147>.
49. Reese MG. 2001. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput Chem* 26:51–56. [https://doi.org/10.1016/S0097-8485\(01\)00099-7](https://doi.org/10.1016/S0097-8485(01)00099-7).
50. Li RW (ed). 2011. Metagenomics and its applications in agriculture, biomedicine and environmental studies. Nova Science Publishers, Hauppauge, NY.
51. Klucar L, Stano M, Hajduk M. 2010. phiSITE: database of gene regulation in bacteriophages. *Nucleic Acids Res* 38:D366–D370. <https://doi.org/10.1093/nar/gkp911>.
52. Torres PJ, Edwards RA, McNair KA. 2017. PARTIE: a partition engine to separate metagenomic and amplicon projects in the Sequence Read Archive. *Bioinformatics* 33:2389–2391. <https://doi.org/10.1093/bioinformatics/btx184>.
53. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
54. Towns J, Cockerill T, Dahan M, Foster I, Gaither K, Grimshaw A, Hazlewood V, Lathrop S, Lifka D, Peterson GD, Roskies R, Scott JR, Wilkins-Diehr N. 2014. XSEDE: accelerating scientific discovery. *Comput Sci Eng* 16:62–74. <https://doi.org/10.1109/MCSE.2014.80>.
55. Stewart CA, Turner G, Vaughn M, Gaffney NI, Cockerill TM, Foster I, Hancock D, Merchant N, Skidmore E, Stanzione D, Taylor J, Tuecke S. 2015. Jetstream: a self-provisioned, scalable science and engineering cloud environment, p 1–8. Proceedings of the 2015 XSEDE Conference on Scientific Advancements Enabled by Enhanced Cyberinfrastructure - XSEDE '15. Association for Computing Machinery, New York, NY.
56. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.