



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

PROGRAMA DE DOCTORADO EN CIENCIAS BIOMÉDICAS

FACULTAD DE MEDICINA

**AISLAMIENTO Y CARACTERIZACIÓN DE FAGOS CON CAPACIDAD LÍTICA  
PARA *Vibrio cholerae* O1/O139**

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
**DOCTOR EN CIENCIAS**

PRESENTA:

**Biól. GUILLERMO ALEJANDRO SOLÍS SÁNCHEZ**

DIRECTOR DE TESIS:

**DR. CARLOS ALBERTO ESLAVA CAMPOS**  
FACULTAD DE MEDICINA

COMITÉ TUTOR:

**DRA. BERTHA GONZÁLEZ PEDRAJO**  
INSTITUTO DE FISIOLÓGÍA CELULAR

**DR. KARLEN GAZARIAN**

INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

MÉXICO, CDMX.

Agosto 2016



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.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada para la realización de este proyecto (204332).

La presente investigación se realizó en el laboratorio de Patogenicidad Bacteriana del Departamento de Salud Pública de la Facultad de Medicina de la UNAM, a cargo del Dr. Carlos Alberto Eslava Campos.

Laboratorio 325 Nte., del Departamento de Genética Molecular del Instituto de Fisiología Celular de la UNAM, a cargo de la Dra. Bertha González Pedrajo.

Laboratorio de Proteómica del Departamento de Bioquímica de la Facultad de Medicina, UNAM, a cargo del Dr. Guillermo Mendoza Hernández†.

Unidad de Imagenología del Instituto de Fisiología Celular de la UNAM, a cargo del Dr. Fernando García Hernández.

Laboratorio Central de Microscopía del Departamento de Investigación-SEPI de la Escuela Nacional de Ciencias Biológicas del Instituto Politécnico Nacional, a cargo del Dr. Edgar Oliver López Villegas.

A la Unidad Universitaria de Secuenciación Masiva de ADN del Instituto de Biotecnología de la UNAM, a cargo del Dr. Ricardo Alfredo Grande Cano.

## **Agradecimientos**

A todos los investigadores que apoyaron para la elaboración de este trabajo de tesis, por sus valiosas aportaciones:

Dr. Carlos Eslava Campos, Dra. Bertha González Pedrajo, Dr. Karlen Gazarian, Dr. Guillermo Mendoza Hernández†, Dr. Ulises Hernández Chiñas, Mtro. Armando Navarro Ocaña, Dr. Javier de la Mora, Dr. Fernando García Hernández, Mtra. Esther Sánchez Espíndola, Dr. Edgar Oliver López Villegas, Dr. Juan Enrique Morett, Dr. Ricardo Alfredo Grande Cano y a Leonardo Collado Torres.

### **Al jurado, integrado por:**

Dra. María de Lourdes Girard Cuesy.

Dra. Irma Rosas Pérez.

Dr. Gabriel Guarneros Peña.

Dr. Luis Kameyama Kawabe.

*“Lo que caracteriza al hombre de ciencia no es la posesión del conocimiento o de verdades irrefutables, sino la búsqueda desinteresada e incesante de la verdad”*

**-Karl Popper** (1902 - 1994)

*Dedicada a la memoria de mis amados padres:*

*Concepción Sánchez Garduño*

*Demetrio Solís Romero*

## Índice:

<b>Resumen</b> .....	<b>1</b>
<b>Abstract</b> .....	<b>2</b>
<b>1. Introducción</b> .....	<b>3</b>
<b>2. Antecedentes</b> .....	<b>8</b>
<b>2.1</b> Los bacteriófagos.....	<b>9</b>
<b>2.2</b> Taxonomía.....	<b>14</b>
<b>2.3</b> <i>Vibrio cholerae</i> .....	<b>17</b>
<b>2.4</b> Fagos y tipificación de <i>Vibrio cholerae</i> .....	<b>21</b>
<b>2.5</b> Los fagos en la epidemiología del cólera.....	<b>24</b>
<b>2.6</b> El cólera hoy en día, un problema no resuelto.....	<b>27</b>
<b>3. Justificación</b> .....	<b>29</b>
<b>4. Hipótesis</b> .....	<b>31</b>
<b>5. Objetivos</b> .....	<b>33</b>
<b>6. Material y métodos</b> .....	<b>35</b>
<b>6.1</b> Área de muestreo.....	<b>36</b>
<b>6.2</b> Aislamiento de fagos.....	<b>36</b>
<b>6.3</b> Purificación de fagos.....	<b>37</b>
<b>6.4</b> Identificación de bacteriana.....	<b>38</b>
<b>6.5</b> Análisis de rango de hospederos.....	<b>39</b>
<b>6.6</b> Microscopía electrónica.....	<b>39</b>
<b>6.7</b> Análisis del Genoma.....	<b>40</b>
<b>6.7.1</b> Extracción de DNA.....	<b>40</b>
<b>6.7.2</b> Análisis restricción de DNA.....	<b>40</b>
<b>6.7.3</b> Clonación de fragmentos de restricción.....	<b>41</b>
<b>6.7.4</b> Secuenciación.....	<b>41</b>

6.8 Análisis proteómico.....	42
6.9 Análisis filogenético.....	43
<b>7. Resultados.....</b>	<b>44</b>
7.1 Aislamiento de fagos y bacterias.....	45
7.2 Actividad lítica y especificidad de hospedero de los fagos aislados en la presa Endhó.....	45
7.3 Análisis morfológico de ØVC8.....	49
7.4 Caracterización del genoma de ØVC8.....	49
7.5 Secuencia del genoma de ØVC8.....	52
7.5.1 Organización funcional del genoma de ØVC8.....	56
7.5.2 Módulo de empaquetamiento.....	58
7.5.3 Módulo de morfogénesis de cápside y cola.....	59
7.5.4 Modulo de metabolismo.....	59
7.5.5 Módulo de replicación.....	60
7.6 Análisis comparativo del genoma de ØVC8 con el de VP2 y VP5.....	61
7.7 Relación filogenética de ØVC8 con fagos de la familia <i>Podoviridae</i> .....	61
7.8 Predicción de la estrategia de empaquetamiento del DNA de ØVC8.....	64
7.9 Análisis proteómico.....	66
<b>8. Discusión.....</b>	<b>68</b>
<b>9. Conclusiones.....</b>	<b>83</b>
<b>10. Perspectivas.....</b>	<b>85</b>
<b>11. Bibliografía.....</b>	<b>88</b>
<b>12. Artículo original.....</b>	<b>96</b>
<b>13. Propuesta de creación de un nuevo género de virus.....</b>	<b>111</b>



<b>ÍNDICE DE FIGURAS:</b>	<b>Página</b>
Figura 1. Esquema de los ciclos de replicación lítico - lisogénico de fagos y su repercusión sobre poblaciones bacterianas.	<b>10</b>
Figura 2. Familias de fagos, los nombres están asignados de acuerdo con el tipo de ácidos nucleicos que poseen y con la morfología que presentan	<b>16</b>
Figura 3. Morfotipos de vibriofagos con cola.	<b>17</b>
Figura 4. Micrografía del fago filamentosos CTXØ de <i>Vibrio cholerae</i> , perteneciente a la familia <i>Inoviridae</i> .	<b>19</b>
Figura 5. Organización genética de CTXØ.	<b>20</b>
Figura 6. Modelo de la influencia de los fagos en el comportamiento estacional del cólera y dominancia de serogrupos específicos.	<b>26</b>
Figura 7. Casos de cólera reportados desde la llegada de cólera a México en 1991 hasta el último brote epidémico en 2013.	<b>28</b>
Figura 8. Localización de la presa Endhó y los puntos de muestreo en el estado de Hidalgo.	<b>36</b>
Figura 9. Ensayo de doble capa con agar suave, donde se observa la morfología de las placas de lisis producidas por ØVC8 sobre un cultivo bacteriano de <i>V. cholerae</i> O1 cepa E8021.	<b>46</b>
Figura 10. Micrográficas de microscopía electrónica del fago ØVC8.	<b>49</b>
Figura 11. Patrones de restricción de DNA de ØVC8 tratado con 14 enzimas de restricción.	<b>51</b>

Figura 12. Patrón electroforético del DNA de ØVC8 tratado con la enzima <i>Hpa</i> II.	<b>52</b>
Figura 13. Organización del genoma de ØVC8.	<b>55</b>
Figura 14. Modelos de la estructura secundaria de los terminadores Rho-independientes predichos en el genoma de ØVC8 usando el programa WebGeSTer.	<b>57</b>
Figura 15. Comparación de los genomas completos de ØVC8 contra VP2 y VP5 utilizando el programa EasyFig 2.1.	<b>62</b>
Figura 16. Árbol Neighbor-joining para comparar la secuencia de aminoácidos del ORF3 (subunidad mayor de la terminasa) de ØVC8 y de otros 54 fagos obtenidas de GenBank.	<b>63</b>
Figura 17. Árbol Neighbor-joining para comparar la secuencia de aminoácidos de la subunidad mayor de la terminasa del ORF3 del ØVC8 y de otros 88 fagos pertenecientes al orden <i>Caudovirales</i> .	<b>65</b>
Figura 18. Proteínas estructurales de ØVC8.	<b>67</b>

## ÍNDICE DE TABLAS:

Página

Tabla 1.	Factores de virulencia codificados de fagos involucrados en diversas etapas de la patogénesis bacteriana.	<b>13</b>
Tabla 2.	Fagos utilizados para la caracterización de <i>V. cholerae</i> .	<b>24</b>
Tabla 3.	Pandemias de cólera registradas desde 1817 a la fecha.	<b>25</b>
Tabla 4.	Perfiles de infección de los fagos aislados de la presa Endhó en cepas de <i>V. cholerae</i> .	<b>47</b>
Tabla 5.	Cepas utilizadas en el ensayo de perfil de infección y que no fueron susceptibles a la infección por fagos.	<b>48</b>
Tabla 6.	Marcos de lectura abierta (ORFs) predichos para ØVC8.	<b>53</b>
Tabla 7.	Regiones promotoras predichas en el genoma de ØVC8 usando el programa PromoterHunter.	<b>58</b>

## Resumen

El cólera es un padecimiento que ocasiona diarrea grave que conduce a la deshidratación del paciente, ocasionado por *Vibrio cholerae* de los serogrupos O1 y O139. En los últimos años, se ha propuesto que fagos líticos específicos para dichos serogrupos son un factor importante en la aparición cíclica de cólera en áreas endémicas de la enfermedad. Sin embargo, su participación en regiones no endémicas de cólera, con largos períodos interepidémicos, aún no se ha descrito. El objetivo de éste trabajo fue aislar y caracterizar fagos líticos específicos para *V. cholerae* O1 y O139. Se aislaron dieciséis fagos con actividad lítica y lisogénica, de estos solo uno, designado como ØVC8 mostró actividad lítica específica contra cepas de *V. cholerae* O1. Las imágenes de ØVC8 obtenidas por MET mostraron un fago de cola corta y cápside isométrica, características de la familia *Podoviridae*, en la cual se incluyen a los fagos VP2 y VP5 de *V. cholerae*. El análisis del genoma de ØVC8 mostró que se trata de DNA de doble cadena lineal con 39,422 pb y 50,8% G + C. De los 48 ORFs anotados, 16 mostraron homología con secuencias de función conocida y varios dominios conservados. El análisis bioinformático mostró múltiples dominios conservados incluido uno de tipo Ig, lo que sugiere que ØVC8 podría adherirse a sustratos mucosos como el epitelio intestinal humano. El análisis del genoma además sugiere que ØVC8 utiliza como estrategia de empaquetamiento del DNA de tipo "extremos cohesivos de cadena simple" descrita en el grupo lambda. Además, se aislaron y secuenciaron cuatro proteínas estructurales que corresponden a las proteínas: portal, cápside y de cola del fago. Este es el primer trabajo en México en el que se reporta la presencia de fagos específicos para *V. cholerae* O1, ØVC8 es un nuevo fago para *V. cholerae* O1 que junto con los fagos VP2 y VP5 forman un grupo específico de la familia *Podoviridae* y un nuevo género denominado *Vp5virus*.

## Abstract

Cholera is a severely dehydrating diarrhea caused by *Vibrio cholerae* O1 and O139 serogroups, which are responsible for epidemics and pandemics. In recent years, specific lytic phages for *V. cholerae* have been proposed to be an important factor in the cyclic occurrence of cholera in endemic areas. The role and potential participation of lytic phages during the long interepidemic periods of cholera in non-endemic regions have not yet been described. The purpose of this study was to isolate and characterize specific lytic phages against *V. cholerae* O1 and O139 strains. Sixteen phages with lytic and lysogenic activity were isolated, and only the ØVC8 phage showed specific lytic activity against *V. cholerae* O1 strains. The TEM images of ØVC8 showed a phage with a short tail and an isometric head, characteristics of the *Podoviridae* family. Additionally, ØVC8 is related with *V. cholerae* phages VP2 and VP5, which have been proposed as a new genus called *Vp5virus*. The ØVC8 genome consists of a linear double-stranded DNA with 39,422 bp and 50.8% of G + C. Of the 48 annotated ORFs, 16 showed homology with sequences of known function and several conserved domains. Bioinformatics analysis showed multiple conserved domains, including an Ig domain, suggesting that ØVC8 might to adhere to different mucus substrates, such as the human intestinal epithelium. Furthermore, the analyses suggest ØVC8 employ the “single-stranded cohesive ends” packaging strategy from the lambda-like group. Moreover, four structural proteins were isolated and sequenced, corresponding to portal, capsid, and tail proteins. This is the first work in Mexico, where is reported the presence specific phages for *V. cholerae* O1, ØVC8 is a new lytic phage for *V. cholerae* O1, which together with VP2 and VP5 vibriophages form a specific group of *Podoviridae* family what has been proposed to be included in a new genus named *Vp5virus*.

# 1. Introducción

Los bacteriófagos o simplemente fagos, son virus que infectan bacterias. Estos son las entidades biológicas más abundantes y diversas en el planeta, donde se estima que su densidad supera en un factor de 10 a 1 al de las bacterias, llegando a existir aproximadamente hasta  $10^{30}$  fagos en el planeta (Brüssow & Hendrix 2002; Rohwer 2003; Mann 2005; Suttle 2005). Su hábitat principal son los océanos, donde representan el mayor reservorio de diversidad genética del planeta (Paul *et al.*, 2002). Ecológicamente, los fagos tienen un impacto significativo en el planeta, tanto en los ciclos biogeoquímicos como en los eventos de diversificación de las comunidades bacterianas (Chibani-Chennoufi *et al.*, 2004). En la actualidad se les considera como uno de los factores más importantes en la evolución bacteriana (Miller 2001; Hendrix 1999).

El impacto que los fagos tienen sobre las comunidades bacterianas y en particular en los patógenos, ha motivado un creciente interés en su estudio durante las últimas dos décadas. Uno de los ejemplos mejor estudiados es el de *V. cholerae* y sus fagos, donde el principal factor de virulencia de ésta bacteria, la toxina de cólera (CT), está codificada en el genoma del fago filamentoso ØCTX, el cual contribuye a la patogenicidad y a la aparición de nuevas clonas toxigénicas de *V. cholerae* (Waldor & Mekalanos 1996). En áreas endémicas de cólera, se ha determinado que existe una estrecha relación entre las poblaciones de *V. cholerae* y la presencia de ciertos fagos, lo que ha llevado a proponer la hipótesis de que éstos constituyen un elemento importante en la regulación poblacional de la bacteria en el ambiente, por lo que se ha propuesto que pueden contribuir de manera importante en el comportamiento estacional del cólera. Con esta hipótesis, se intenta explicar el comportamiento cíclico de la enfermedad, observado tanto en las regiones donde el cólera es endémico, como

en aquellas en las que se presenta tras largos periodos interepidémicos (Faruque *et al.*, 2005).

El cólera es una enfermedad diarreica causada por los serogrupos O1 y O139 de *V. cholerae*, bacilo curvo Gram-negativo con un flagelo polar. Desde 1817 y hasta la fecha se han registrado 7 pandemias de cólera asociadas todas ellas a *V. cholerae* del serogrupo O1. Actualmente se considera que el cólera es una enfermedad no controlada que afecta principalmente a países en desarrollo (Verma *et al.*, 2012). Desde la aparición de la séptima pandemia de cólera en 1961, una de las principales estrategias para combatir la enfermedad ha sido predecir la aparición de nuevos brotes epidémicos y por lo tanto las investigaciones se han enfocado en identificar el o los factores involucrados con la supervivencia y la ocurrencia cíclica de *V. cholerae* en el medio ambiente (Borroto 1997). Algunos de éstos factores están relacionados con la bacteria como es el proceso de formación de biopelículas o la propiedad de adoptar un estado viable no cultivable (Alam *et al.*, 2007), otros son de índole ambiental como cambios en el pH, salinidad, temperatura, ausencia o incremento de nutrientes e incluso eventos de depredación por protistas y otras bacterias (Borroto 1998; Lipp *et al.*, 2002; Huq *et al.*, 2005). Sin embargo, y no obstante los avances en el conocimiento de *V. cholerae* aún no se ha logrado definir un modelo que explique de manera concluyente la epidemiología del cólera.

Las primeras seis pandemias de cólera se caracterizaron por la participación de cepas de *V. cholerae* del biotipo clásico y por su duración, que fue menor al mostrado por la séptima pandemia, ocasionada por cepas del biotipo El Tor. La séptima pandemia inició en Indonesia en 1961 y se extendió rápidamente por el resto del mundo, llegó a India en 1962, África en 1970, sureste de Europa en 1970 y 30 años



después al continente americano. En América latina el cólera inició en Perú en enero de 1991, de donde se propagó rápidamente a casi todos los países de Latinoamérica, con excepción de las islas del caribe y Uruguay (Sepúlveda *et al.*, 1992). El cólera llegó a México en junio de 1991, afectando gran parte del país y ocasionando más de 98,000 casos en un periodo de 10 años (Borroto & Martínez-Piedra 2000). Posteriormente tras un periodo interepidémico de 8 años (2002-2009), en el 2010 el Sistema Nacional de Vigilancia Epidemiológica (SiNaVE) identificó un caso nuevo en Navolato, Sinaloa con el cual inició una nueva epidemia de cólera en el país que se ha mantenido hasta el 2014 (<http://colera.sinave.gob.mx>).

Con la llegada del cólera a México en 1991, se realizó el monitoreo de *V. cholerae* en algunas zonas del país afectadas por la enfermedad (Isaac-Márquez *et al.*, 1998), los resultados mostraron la presencia de la bacteria en diferentes ambientes acuáticos, aunque, mayormente pertenecientes a los serogrupos No-O1/No-O139. Así mismo, desde que se reportó el último caso de cólera en 2001, las cepas del serogrupo O1 solo fueron aisladas esporádicamente y nunca asociadas a casos de cólera.

En octubre de 2010 después de 100 años de ausencia se presentó un brote de cólera en Haití (Chin *et al.*, 2011; Orata *et al.*, 2014), tal hecho sugería el resurgimiento del padecimiento en el continente y la posible emergencia de una nueva epidemia en la región. A pesar de las medidas sanitarias tomadas para contener la enfermedad, se presentaron brotes en República Dominicana, Cuba y México (WHO 2013). En nuestro país, solo se reportaron casos aislados de 2010 hasta 2012, y fue en 2013 en el que se reportó un aumento considerable en el número de casos, con más de 180 reportados. Sin embargo, en 2014 hasta el mes de septiembre, el número de casos identificado fue de solo 14 (<http://colera.sinave.gob.mx>).

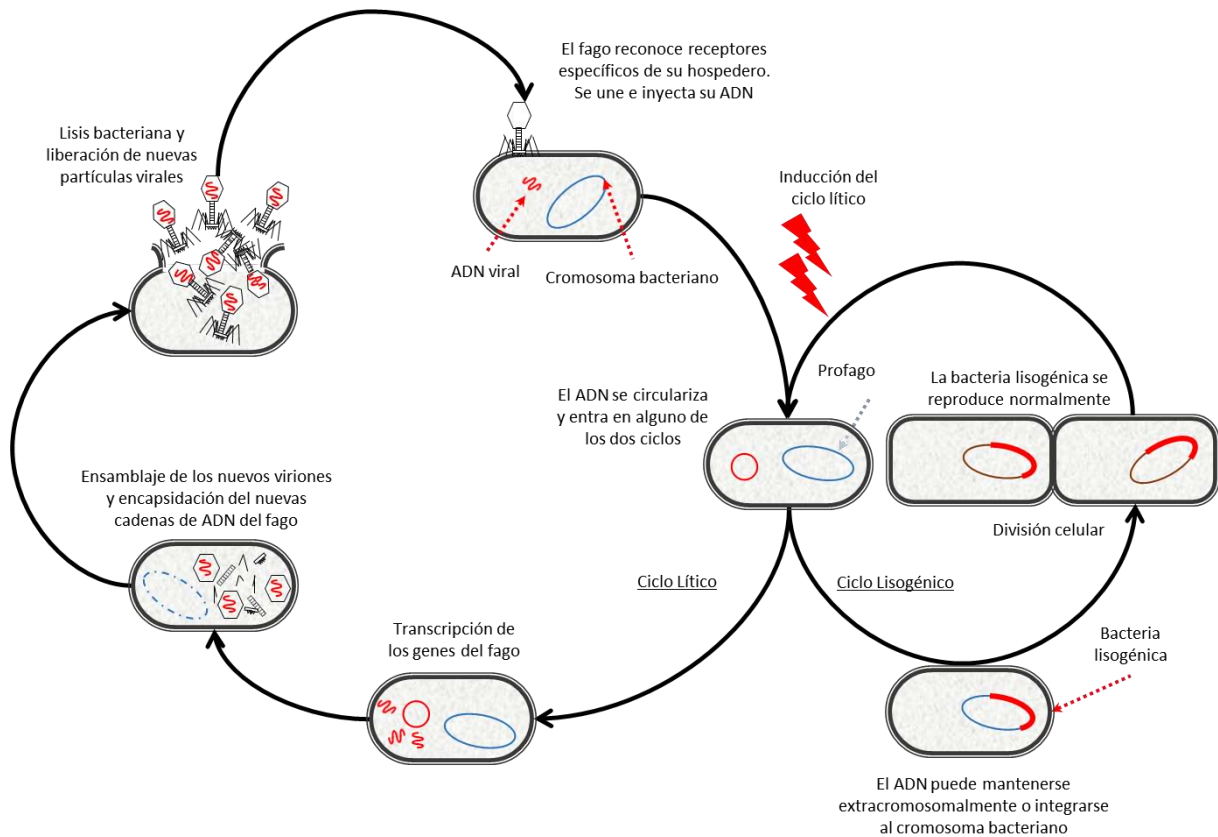
Hasta el momento, los estudios realizados en áreas endémicas del cólera, principalmente en el delta del río Ganges en la frontera entre India y Bangladesh, indican que los fagos pueden ser un factor importante en la epidemiología del cólera. Sin embargo, en áreas afectadas por la enfermedad, pero en las que los periodos interepidémicos han sido de hasta 100 años, como es el caso de México y Haití, no se han realizado estudios que permitan conocer la interacción entre fagos y la bacteria causante del cólera.

## **2. Antecedentes**

## 2.1 Los bacteriófagos

Los bacteriófagos o simplemente fagos, fueron descubiertos de manera independiente por Frederick W. Twort en 1915 y por Felix d'Herelle en 1917 (Weinbauer 2004). Twort describió la formación de parches vítreos, mucoides y acuosos en cultivos de *Micrococcus* que habían crecido sobre la placa de agar donde intentaba cultivar el virus vaccinia. Dichas observaciones fueron atribuidas a la presencia de una estructura ultramicroscópica y cuya organización era más simple que la de una bacteria o una ameba. Por su parte, Felix d'Herelle, señaló la presencia de un microbio “antagonista” para las bacterias y que daba como resultado la lisis de éstas en cultivo líquido, de igual forma observó parches o placas de lisis sobre la superficie de agar sembrado con las bacterias. D'Herelle concibió a estos microbios invisibles como ultravirus, los cuales invadían a las bacterias y se multiplicaban a expensas de ellas y a las cuales les llamó bacteriófagos, literalmente comedores de bacterias (Summers 2011).

Los fagos son parásitos absolutos, algunos pueden poseer genes involucrados en la biosíntesis de ácidos nucleicos y en la replicación, no poseen metabolismo propio, por lo que necesitan de un hospedero para poder reproducirse. La replicación o multiplicación de un fago se lleva a cabo en varias etapas (Figura 1), comunes a todos los virus e incluyen: adsorción, inyección del material genético, expresión y replicación de los genes virales, ensamblaje de viriones, y liberación de los nuevos virus (Guttman *et al.* 2005).



**Figura 1.** Representación esquemática del ciclo replicativo de los fagos.

El ciclo replicativo de un fago inicia con el reconocimiento de receptores específicos presentes en componentes de la superficie de la bacteria como pilis, fimbrias, LPS, OMPs, etc. En algunos fagos la unión al hospedero puede involucrar dos etapas y dos receptores distintos, como es el caso del fago T4 que en una primera etapa reconoce al OmpC como receptor primario al cual se unen al menos tres de las seis fibras de la cola, uniéndose de manera, lo que desencadena modificaciones estructurales en las proteínas de la base y dando lugar a un segundo paso en el cual se da la vinculación irreversible con un segundo receptor, que en el caso del fago T4 son residuos de heptosa en el core interno del LPS bacteriano (Guttman *et al.*, 2005). Posteriormente, la pared celular bacteriana es lisada mediante enzimas que se encuentran en la base de cola o en la cápside del fago y el material genético del fago

es transferido a la célula hospedera (Molineux 2001). En el caso de fagos con genomas lineales, el DNA generalmente se circulariza por sus extremos tras ser transferido a la bacteria hospedante y posteriormente sus genes se transcriben. Los productos de dichos genes denominados tempranos, producen proteínas que protegen al genoma del fago contra los mecanismos de defensa del hospedero, inactivando proteasas y/o bloqueando enzimas de restricción de la bacteria (Guttman *et al.*, 2005). Otro grupo de genes conocidos como intermedios, participan en la replicación del DNA del fago y finalmente los genes tardíos codifican para los componentes de las nuevas partículas virales (Weigel & Seitz 2006).

Al final del ciclo replicativo se lleva a cabo la lisis de la célula y la liberación de las nuevas partículas virales. Con excepción de los fagos filamentosos, los cuales, debido a su morfología y morfogénesis, son los únicos que pueden salir mediante extrusión sin ocasionar consecuencias letales para el microorganismo, todos los demás lisan a la bacteria (Smith 1988). El proceso de lisis de la pared bacteriana se lleva a cabo mediante un sistema doble en el que participan peptidoglicano hidrolasas (endolisinas) y holinas. Las holinas son proteínas hidrofóbicas que se insertan en la membrana citoplasmática de la bacteria lo que permite a las endolisinas llegar al peptidoglicano, en particular a residuos de mureína de la pared celular (Young 2002; Loessner 2005).

Los fagos pueden presentar los llamados ciclos lítico, lisogénico, pseudolisogénico e infecciones crónicas, cada uno de ellos tiene un efecto importante sobre las poblaciones bacterianas. En el ciclo lítico, los fagos infectan y reprograman el metabolismo celular de su hospedante para producir nuevos fagos los cuales son liberados mediante la lisis de la bacteria. Por el contrario, en el ciclo lisogénico el

genoma del fago se mantiene en la célula hospedera (profago), permaneciendo estable durante varias generaciones hasta que el ciclo lítico es inducido. Como consecuencia, el genoma del fago puede alterar el fenotipo de la bacteria hospedera mediante la expresión de genes que no existían antes de la infección, tales como factores de colonización, adhesión, invasión, resistencia a suero y producción de exotoxinas, este evento se conoce como conversión lisogénica (Tabla 1). Un ejemplo clásico de este proceso es el de *V. cholerae*, cuyo principal factor de virulencia, la toxina de cólera (CT) codificada por el bacteriófago filamentosos CTXØ, actúa de manera coordinada con otros factores de virulencia relacionados con la presencia de otros fagos (Saunders *et al.*, 2001; Boyd & Brüssow 2002; Wagner & Waldor 2002; Abedon & Lejeune 2005).

La pseudolisogenia es un estado en el que el fago no se integra en el genoma del hospedero, ni entra en el ciclo lítico, por lo que no está claro si este estado representa un ciclo de vida *per se* o se trata de una "pausa" que requiere el fago como consecuencia de la reducción de la actividad de la célula debido a estados de inanición. En cualquier caso, el evento tiene implicaciones importantes en las interacciones bacteria-fago en la naturaleza. Si los nutrientes son limitados transitoriamente en un entorno determinado, el fago no se puede reproducir por la falta de crecimiento de las bacterias, pero puede continuar su reproducción en la fase donde se había quedado una vez que se reanude el crecimiento celular (Weinbauer 2004; Clokie *et al.*, 2011).

Por su parte, en la infección crónica o estado portador (carrier-state), una vez que la bacteria es infectada, el genoma del fago no se integra al cromosoma bacteriano, permaneciendo en forma episomal y liberando progenie constantemente de la célula hospedera mediante extrusión. Los fagos formados se desprenden

lentamente de la célula durante un periodo largo de tiempo sin ocasionar la muerte de la bacteria hospedera (Khemayan *et al.*, 2006; Cenens *et al.*, 2013).

**Tabla 1.** Factores de virulencia involucrados en diversas etapas de la patogénesis bacteriana codificados por fagos (Adaptado de: Saunders *et al.*, 2001; Boyd & Brüssow 2002; Wagner & Waldor 2002; Abedon & Lejeune 2005).

Bacteria	Factor de virulencia (gen asociado)	Fago
<b>Resistencia a suero/fagocitosis</b>		
<i>Escherichia coli</i>	OMP ( <i>bor</i> )	$\lambda$
<i>Escherichia coli</i>	Superoxido dismutasa ( <i>sodC</i> )	Sp4
<i>Pseudomonas aeruginosa</i>	alteracion de antígeno O	D3
<i>Salmonella enterica</i>	Superoxido dismutasa ( <i>sodCI</i> )	Gifsy-2
	alteracion de antígeno O ( <i>rfb</i> )	Epsilon34
<b>Producción de exotoxinas</b>		
<i>Bordetella avium</i>	Toxina de pertusis	Ba1
<i>Clostridium botulinum</i>	Toxina botulinica ( <i>c1</i> )	C1
<i>Corynebacterium diphtheriae</i>	Toxina difterica ( <i>tox</i> )	$\beta$ -Phage
<i>Escherichia coli</i>	Toxinas de shiga ( <i>stx</i> )	H-19B
<i>Pseudomonas aeruginosa</i>	Citotoxinas	$\emptyset$ CTX
	Toxina exfoliativa A ( <i>eta</i> )	phiETA
<i>Staphylococcus aureus</i>	Toxina de síndrome de choque tóxico	SapI
	Isla de patogenicidad	80alfa
	Enterotoxina ( <i>see</i> )	T12
<i>Streptococcus pyogenes</i>	Enterotoxina P ( <i>sep</i> )	CS112
	Enterotoxina A ( <i>entA</i> )	phi13
	Enterotoxina A ( <i>sea</i> )	phiMu50A
	Stafilocinasa ( <i>sak</i> )lwof	phi13
<i>Vibrio cholerae</i>	Toxina oclusora de la zonula ( <i>zot</i> )	CTXphi
	Enterotoxina accesoria de colera ( <i>ace</i> )	
<i>Escherichia coli</i>	Enterohemolisina ( <i>hly</i> )	$\emptyset$ FC3208



Continúa **Tabla 1.**

Bacteria	Factor de virulencia (gen)	Fago
<b>Colonización y adhesión</b>		
<i>Escherichia coli</i>	Adhesión a epitelio bucal ( <i>lom</i> )	lamnda
<i>Mycoplasma arthritidis</i>	Adhesión ( <i>vir</i> )	MAV1
<i>Pseudomonas aeruginosa</i>	Adhesión a epitelio bucal	FIZ15
<i>Streptococcus mitis</i>	Adhesión a plaquetas ( <i>pblA</i> , <i>pblB</i> )	Sm1
<i>Vibrio cholerae</i>	Pilli corregulado por toxina ( <i>tcp</i> )	VPIphi
<b>Invasión</b>		
<i>Salmonella enterica</i>	Sistema de secreción tipo III ( <i>sopE</i> )	SopEphi
	Sistema de secreción tipo III ( <i>gogB</i> )	Gifsy-1
	Secuencia IS-like ( <i>gipA</i> )	Gifsy-1
	Sistema de secreción tipo III ( <i>gtgB</i> )	Gifsy-2
	Sistema de secreción tipo III ( <i>sspH1</i> )	Gifsy-3
<i>Streptococcus pyogenes</i>	Hialuronidasa ( <i>hylP</i> )	H4489A
<i>Staphylococcus aureus</i>	Fibrinolisisina	S phi-C

## 2.2 Taxonomía

A pesar de que los fagos fueron identificados desde principios del siglo XX, se desconocía cuál era su estructura. Con el desarrollo del microscopio electrónico realizado por los hermanos Ernst y Helmut Ruska en 1930, fue posible identificar por primera vez la ultraestructura de los virus. Las primeras micrografías realizadas por microscopía electrónica se publicaron en 1940 en Alemania y los primeros fagos descritos fueron T1, T4, T5, T7, 7-11, Vi1 de *E. coli* y PB1 de *Pseudomonas* spp., mostrando que poseían una cola y eran altamente heterogéneos (Ackermann 2011a; Ackermann 2011b).

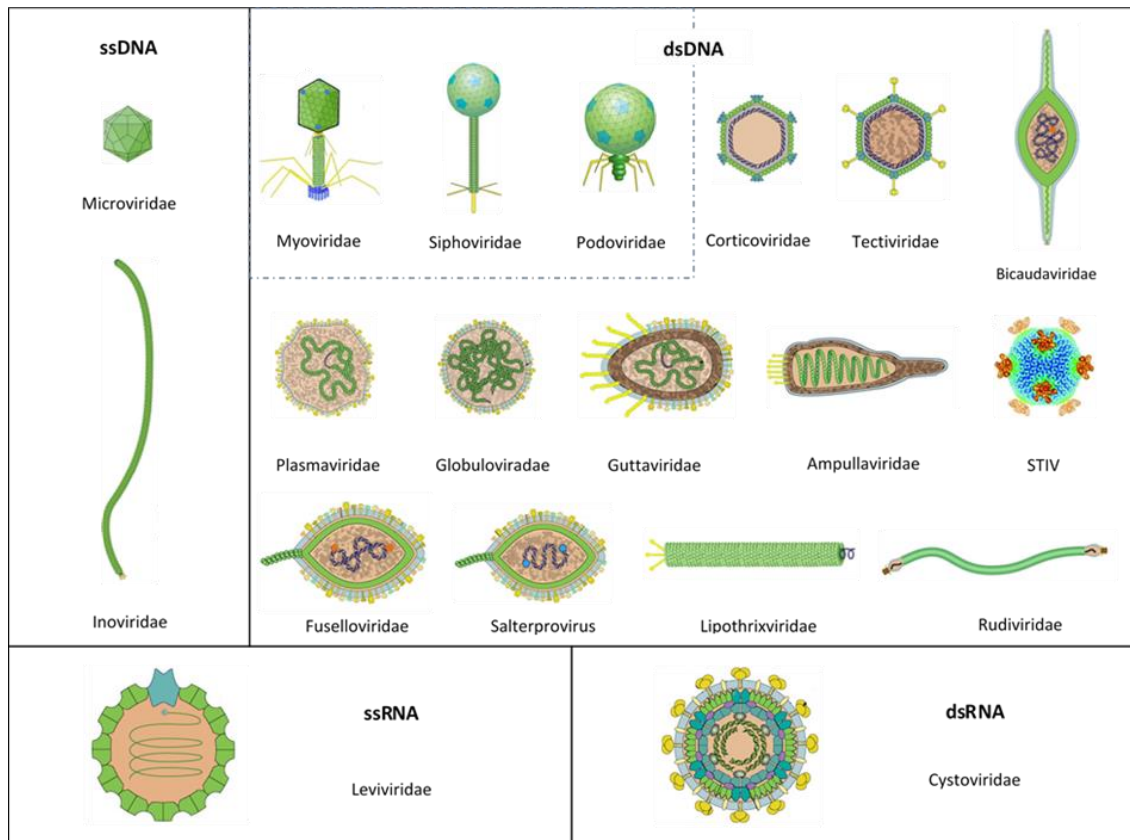
Frank Macfarlane Burnet demostró en 1937 que los fagos diferían en tamaño y resistencia contra agentes fisicoquímicos, posteriormente en 1939, Helmut Ruska demostró que los fagos eran morfológicamente diversos, fueron los pioneros en

plantear procedimientos para clasificar los fagos (Kruger *et al.* 2000). En 1943, Ruska propone una clasificación de los virus basada en su morfología y en 1948, Francis O. Holmes clasifica los virus en tres familias, agrupando a los fagos en el orden *Virales*, suborden *Phagineae* el cual contiene una sola familia *Phagaceae* y un solo género, *Phagus* (Ackermann 2011a).

En 1962 André Lwoff, Robert Horne y Paul Tournier propusieron que la clasificación de los virus debía basarse en algunas de sus propiedades como son el tipo de ácido nucleico (DNA o RNA), forma de la cápside, presencia de envoltura y número de capsómeros; este esquema fue conocido como el sistema LHT. Posteriormente en 1965, este sistema fue adoptado por el Comité Provisional para la Nomenclatura de Virus (PCNV), que después se convertiría en el Comité Internacional de Taxonomía de Virus (ICTV). En 1971, el ICTV publicó su primer informe en el que se incluían seis géneros de fagos los cuales correspondían a los fagos-T y otros grupos ejemplificados por los fagos  $\lambda$ , PM2, ØX174, MS2 y fd (Ackermann 2011a). Actualmente el ICTV es el único organismo especializado en la taxonomía de virus, de él se derivan diversos subcomités especializados en virus de vertebrados, invertebrados, bacterias, plantas, protozoos y hongos. De acuerdo con el último reporte del ICTV, para clasificar a los virus se han establecido 7 órdenes, 104 familias, 23 subfamilias, 505 géneros y 3,186 especies.

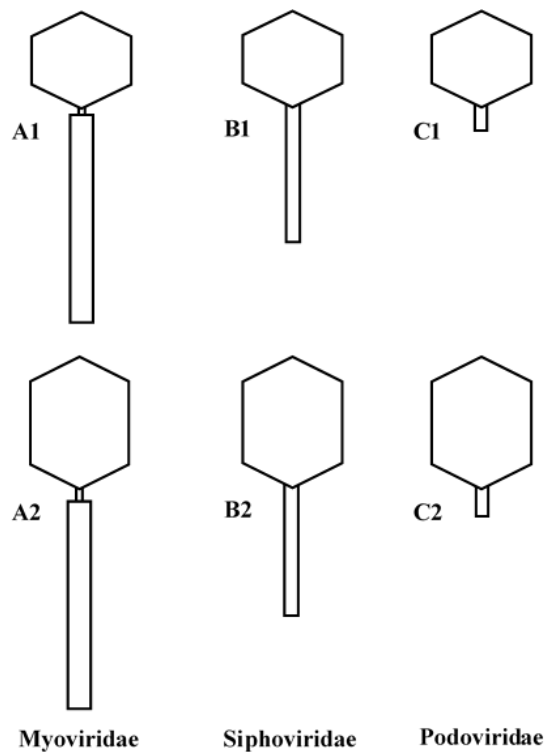
Los fagos son el grupo de virus más grande conocido, son además un grupo extremadamente heterogéneo en cuanto a sus propiedades estructurales, fisicoquímicas y biológicas (Guttman *et al.*, 2005) y a pesar de su gran número y distribución, es poco lo que se sabe acerca de su biodiversidad, biogeografía y filogenia. Hasta el momento se han descrito más de 6,300 fagos mediante microscopía

electrónica, la mayoría de los cuales (96.3%) pertenecen al orden *Caudovirales* [Figura 2]. En este orden se agrupan las familias *Myoviridae* (del tipo T4), *Siphoviridae* (del tipo  $\lambda$ ) y *Podoviridae* (del tipo T7). Mientras que el resto de los fagos descritos (3.6 %) corresponde a los fagos filamentosos, poliédricos y pleomorfos (Ackermann 2012).



**Figura 2.** Familias de fagos, los nombres están asignados de acuerdo con el tipo de ácido nucleico que poseen y con la morfología que presentan. El recuadro punteado muestra a las familias correspondientes al orden *Caudovirales*, las demás familias no poseen un orden asignado aún. Modificado de Ackermann, 2007; <http://viralzone.expasy.org/>.

Uno de los primeros trabajos para la clasificación de los fagos de bacterias del género *Vibrio* lo realizaron Ackermann y cols. (1984), tomando como referencia una colección de 85 fagos, 84 de los cuales son fagos con cola y uno filamentoso. Ackermann dividió a los fagos con cola de acuerdo a su morfología, además de propiedades fisicoquímicas del virión, tipo de ácidos nucleicos y relaciones serológicas en seis morfotipos básicos, lo que permitió determinar las similitudes morfológicas con fagos de otros géneros bacterianos (Figura 3).



**Figura 3.** Morfotipos de vibriofagos con cola.

### 2.3 *Vibrio cholerae*

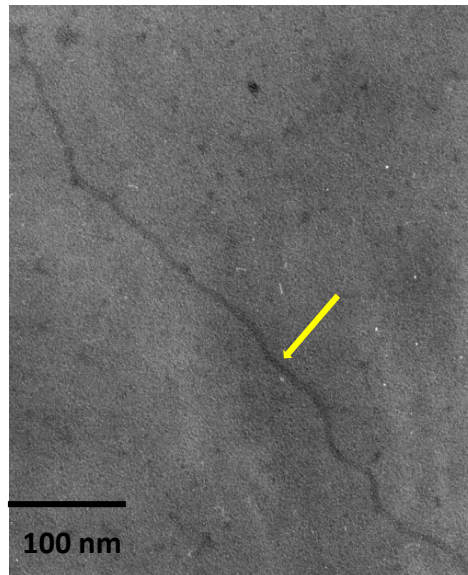
Cepas de *Vibrio cholerae* son los agentes responsables de la enfermedad diarreica conocida como cólera. *V. cholerae* es además una bacteria de vida libre y muchas de sus cepas no son patógenas y son importantes para los ecosistemas. Esta bacteria

fue descrita por primera vez en 1854 por Filippo Pacini, es un bacilo curvo Gram-negativo con un flagelo polar del cual se han identificado más de 200 serogrupos. Sin embargo, de éstos solo los serogrupos O1 (biotipos Clásico y El Tor) y O139 (Bengal), son responsables de ocasionar el cólera (Waldor & Mekalanos 1994). Los otros serogrupos que se conocen como No-O1/No-O139, constituyen un grupo heterogéneo que se ha relacionado con casos esporádicos de diarrea e infecciones en pacientes inmunodeprimidos. Sin embargo, algunos serogrupos como es el caso de O37, han sido relacionados con brotes epidémicos de diarrea en Checoslovaquia en 1965 y en Sudan en 1968 (Li 2002).

Cepas de *V. cholerae* producen una variedad de productos extracelulares que tienen un efecto dañino en células eucariotas, la diarrea es ocasionada por la enterotoxina de cólera (CT), la cual induce un efecto de ADP ribosilación que activa la adenilato ciclasa, incrementando la secreción de iones  $\text{ClO}^-$  y conduciendo a la salida de líquidos dando como resultado la pérdida de más de 20 litros de fluidos por día. Esta bacteria posee, además, toxinas adicionales que participan en la producción de una diarrea moderada. Adicionalmente, la habilidad de la bacteria para colonizar al hospedero depende de la producción de un pili tipo IV conocido como TCP (Toxin Corregulated Pilus). La presencia de TCP en la superficie de la bacteria es necesaria para permitir su adherencia y colonización al epitelio intestinal. La expresión de CT y TCP está bajo control de ToxT que activa la expresión de dichos genes, a su vez, ToxT está regulado por la cascada de virulencia conocida como regulón ToxR el cual responde a estímulos ambientales (Kaper *et al.*, 1995; Sack *et al.*, 2004).

En 1996, Waldor y Mekalanos identificaron que CT, el principal factor de virulencia de *V. cholerae*, está codificada en el genoma del fago lisogénico CTXØ, el

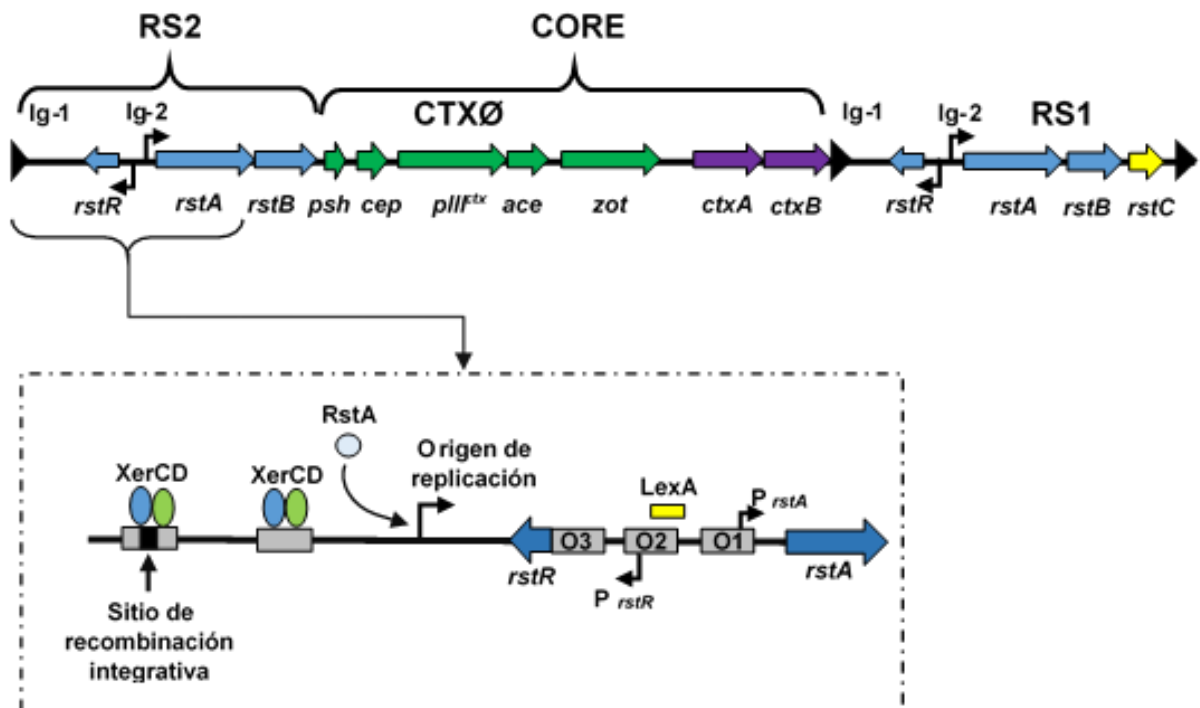
cual es un fago filamentososo perteneciente a la familia *Inoviridae* (Figura 4). El genoma de dicho fago (Figura 5), porta los genes para las subunidades A y B de CT, además de los correspondientes de la toxina oclusora de la zónula (ZOT) y de la enterotoxina accesoria de cólera (ACE). Además de CTXØ, se han descrito otros fagos filamentosos y satélites que contribuyen en la morfogénesis de CTXØ o están estrechamente relacionados a éste. Entre los cuales se encuentran los fagos KSF-1, VGJ, VSK, VSKK, 493, fs1 y fs2. Algunos de estos fagos están involucrados en la transferencia horizontal genética como KSF-1 y VGJ, los cuales pueden proveer componentes de la cápside, ensamble y empaquetamiento del fago RS-1 y posiblemente de otros fagos satélite (Wang *et al.*, 2013).



**Figura 4.** Micrografía del fago filamentososo CTXØ de *V. cholerae* señalado con la flecha amarilla.

A pesar de que CTXØ se encuentra principalmente en los serogrupos O1 y O139 de *V. cholerae*, se han reportado casos esporádicos de *V. cholerae* No-O1/No-O139 que poseen el profago CTX. La baja distribución de CTXØ entre los serogrupos

No-O1/No-O139 probablemente se deba a la escasa presencia de TCP entre dichos serogrupos componente que es el receptor principal del fago. La presencia de CTXØ se ha reportado en cepas de *Vibrio mimicus* en las cuales también se ha identificado la presencia de TCP, lo que sugiere una estrategia de infección del fago similar a la de *V. cholerae*. Adicionalmente, reportes preliminares señalan la presencia de secuencias homólogas de CTXØ y TCP en el genoma de *Vibrio fischeri*, hecho que cambia la perspectiva epidemiológica del cólera, ya que resalta el impacto de los fagos en la movilidad de factores de patogenicidad entre serogrupos patogénicos y no patogénicos e incluso entre distintas especies. Esto explica la existencia de cepas de *V. cholerae* O1 no toxigénicas, así como la posibilidad del surgimiento de nuevos serogrupos No-O1 productoras de CT y con la capacidad para producir cólera (Karaolis *et al.*, 1999; Brüßow *et al.*, 2004).



**Figura 5.** (Página anterior) Representación esquemática de la organización genética del profago CTXØ. Las flechas de color señalan los marcos de lectura abiertos y los triángulos negros en los extremos representan las secuencias repetidas que flanquean al profago. Las flechas curvadas indican los promotores para *rstR*, *rstA* y *ctxAB*. Los genes de color verde están relacionados con la encapsidación y secreción del DNA del fago. Los genes que son casi idénticos entre CTX y RS1 se muestran en color azul y están implicados en la regulación de la transcripción (*rstR*), replicación (*rstA*) e integración (*rstB*). El gen *rstC* (en color amarillo) se encuentra en RS1 y no en CTX, codifica para un anti represor de *rstR*. Los genes en púrpura (*ctxAB*) codifican para las subunidades A y B de la toxina del cólera. En el recuadro de la parte inferior de la figura, se presenta una vista detallada de las regiones intergénica 1 (Ig-1) e intergénica 2 (Ig-2). Ig-1 contiene el sitio de unión para el fago (*attP*) donde se produce la recombinación con el cromosoma de *V. cholerae* para formar el profago integrado. La región *attP*, consta de dos sitios de unión para el dímero formado por XerC y XerD (óvalos azules y verdes, respectivamente). La región Ig-1, contiene también el origen de replicación que es el sitio probable de acción para la proteína RstA (círculo azul claro). La región Ig-2, contiene las regiones reguladoras para la transcripción de *rstR* y *rstA* (*PrstR* y *PrstA*). Esta región contiene también tres operadores RstR (rectángulos grises etiquetados O1, O2 y O3), además de un sitio de unión para LexA (rectángulo amarillo), que se superpone al operador O2 (Las figuras no están a escala; modificado de McLeod *et al.*, 2005).

## 2.4 Fagos y la tipificación de *Vibrio cholerae*

La identificación y caracterización de los microorganismos se realiza por diferentes procedimientos, uno de estos es el análisis de sus características metabólicas, otro mediante la detección de algunos de sus componentes antigénicos con sueros y uno más es el empleo de fagos que por su especificidad permiten identificar y tipificar a estas bacterias (Shattock, 1955; Noguerola & Blanch 2008; Schmelcher & Loessner 2008). La tipificación por fagos es un procedimiento ampliamente utilizado y de gran importancia para la identificación y diferenciación de bacterias patógenas, incluido *Vibrio cholerae* (Miller 2003).

Sin embargo, la alta especificidad de los fagos ha dado la pauta para proponer su empleo para el control y tratamiento de enfermedades ocasionadas por bacterias



en humanos, plantas o animales, (Rees & Loessner 2005; Schmelcher & Loessner 2008). Las diferencias en la capacidad infectiva de los fagos es reflejo de la presencia a nivel celular (bacteria), de receptores específicos que permiten que solo un tipo particular de fago infecte una cepa y no a otra, aunque pertenezcan al mismo género y/o especie (Chibani-Chennoufi *et al.*, 2004). Diversos estudios han demostrado la eficacia para discriminar entre diferentes cepas bacterianas con un mismo cuadro clínico, entre los que se encuentran *Escherichia coli* O157:H7, *Campylobacter jejuni*, *Streptococcus* del grupo D, *Salmonella enterica* subsp. *enterica* serotipos Enteritidis y Typhimurium, *Clostridium difficile* y *Listeria* spp, entre otros (Caprioli *et al.*, 1975; Anderson & Ward 1977; Sell *et al.*, 1983; Ward *et al.*, 1987; Khakhria *et al.*, 1990; Khakhria & Lior 1992; Hagens & Loessner 2014).

Los primeros fagos específicos para *V. cholerae* fueron descritos en 1926 por Felix d'Herelle y desde entonces se han identificado un gran número de éstos. En 1952 un comité de la organización mundial de la salud, sugirió el estudio de *Vibrio cholerae* desde el punto de vista de su significancia epidemiológica, usando el método de fago-tipificación. Durante el estudio inicial más de 600 fagos fueron probados, los cuales se dividieron en cuatro grupos antigénicos, se desarrolló entonces un esquema con cinco fagos para la identificación de *Vibrio cholerae* O1 Clásico (Mukerjee & Phil 1963). Posteriormente, con el surgimiento en 1960 de una nueva pandemia ocasionada por el biotipo El Tor, dio lugar a que se creara un esquema ampliado (Basu & Mukerjee 1968), el cual incluía cinco fagos (I, II, III, IV y V) por el cual *Vibrio cholerae* podía ser diferenciado en 6 grupos, esquema que posteriormente se siguió de rutina (Tabla 2). Con el tiempo el esquema fue restringido a los fagos, tipo I y tipo II. Con el aumento en el número de cepas no tipificables se tuvieron que adicionar nuevos fagos. En 1993

Chattopadhyay y colaboradores presentaron un nuevo esquema de tipificación con cinco nuevos fagos además de los cinco utilizados por Basu & Mukerjee (1968), con éste se probaron 1000 cepas de *Vibrio cholerae* O1 biotipo El Tor, y se identificaron 146 tipos diferentes y se logró tipificar el 99.6% de las cepas probadas.

Hasta 1992, *V. cholerae* O1 era considerado el único serogrupo relacionado con la aparición de brotes de cólera, en octubre de ese año el escenario del cólera cambió debido a la aparición de un nuevo serogrupo que ocasionó un brote de cólera en la ciudad de Madrás al sur de India y que posteriormente fue conocido como serogrupo O139 Bengal. Esta cepa surgió abruptamente y se extendió rápidamente por los países vecinos donde el cólera es endémico (Bik 1995). El nuevo agente causal de cólera dio lugar a que se propusiera un esquema más eficiente de tipificación para el organismo. Para estos Chakrabarty & cols., (2000), propusieron un nuevo esquema de tipificación basado en la utilización de 5 fagos líticos específicos para *V. cholerae* O139. Estos se probaron sobre un total de 500 cepas de *V. cholerae* O139, todas fueron tipificables y se encontró que estas se podían dividir en 10 grupos. Este esquema se sumó al de O1 El Tor como una herramienta más en el estudio de la epidemiología del cólera.

Hasta el momento no se ha registrado ningún cambio en los esquemas de tipificación de *V. cholerae* O1 y O139, tampoco se tiene reportada la aparición de otro serogrupo relacionado con el cólera. El empleo de estos fagos puede ser útil para la predicción de nuevos brotes epidémicos de la enfermedad, así como para definir el comportamiento de *V. cholerae* O1/O139 durante los brotes epidémicos del cólera.

**Tabla 2.** Características de los fagos utilizados en el esquema de tipificación de *Vibrio cholerae* O1 propuesto por Chattopadhyay en 1993.

Bacteriófago	Diámetro de la cápside (nm)	Tamaño de la cola (nm)	Familia
I	60.71 +/- 0.90	8.60 +/- 4.0	<i>Podoviridae</i>
II	59.70 +/- 1.60	9.00 +/- 4.3	<i>Myoviridae</i>
III	72.90 +/- 1.50	9.00 +/- 4.6	<i>Podoviridae</i>
IV	71.10 +/- 1.60	8.60 +/- 4.6	<i>Siphoviridae</i>
V	60.40 +/- 1.70	9.00 +/- 3.9	-----
N4	72.10 +/- 1.40	9.70 +/- 2.0	<i>Podoviridae</i>
S5	71.20 +/- 1.50	12.90 +/- 5.7	<i>Podoviridae</i>
S20	77.30 +/- 1.80	16.00 +/- 2.2	<i>Podoviridae</i>
M4	97.70 +/- 0.03	109.60 +/- 0.2	<i>Myoviridae</i>
D10	62.95 +/- 0.06	101.40 +/- 0.3	<i>Myoviridae</i>

## 2.5 Los fagos en la epidemiología del cólera.

En áreas endémicas del cólera, las epidemias se presentan de manera cíclica, usualmente dos veces por año, con relación a los países afectados por las pandemias, tanto la intensidad como la duración de los brotes epidémicos, así como los periodos interepidémicos pueden variar de una región a otra.

La aparición del cólera y su comportamiento epidemiológico antes del siglo XIX son temas de debate polémico entre los historiadores (Pollitzer 1954). Sin embargo, con el inicio de la primera pandemia de cólera oficialmente reconocida en 1817, se inicia una nueva era en la epidemiología de la enfermedad, dando como resultado el estudio formal del cólera.

Hasta el momento se han documentado siete pandemias de cólera desde 1817 (Tabla 3), de las cuales, las primeras 6 se originaron en Asia en el delta del río Ganges

en la frontera entre India y Bangladesh, mientras que la séptima se originó en la isla de Sulawesy en Indonesia (Kaper *et al.*, 1995). Esta última se caracterizó por ser la más extensa, ya que para 1970 se había reportado cólera en 36 países, en 28 de los cuales se presentaba por primera vez. Además, ha sido la de mayor duración ya que comenzó en 1961 y después de 30 años hizo nuevamente su aparición en el continente americano en donde persiste después de 23 años (Mukhopadhyay *et al.*, 2014: Siddique & Cash, 2014). Otro aspecto importante en la séptima pandemia fue el surgimiento de un nuevo biotipo, ya que mientras las primeras 6 pandemias fueron relacionadas con el biotipo Clásico, la séptima pandemia se asoció con el biotipo El Tor.

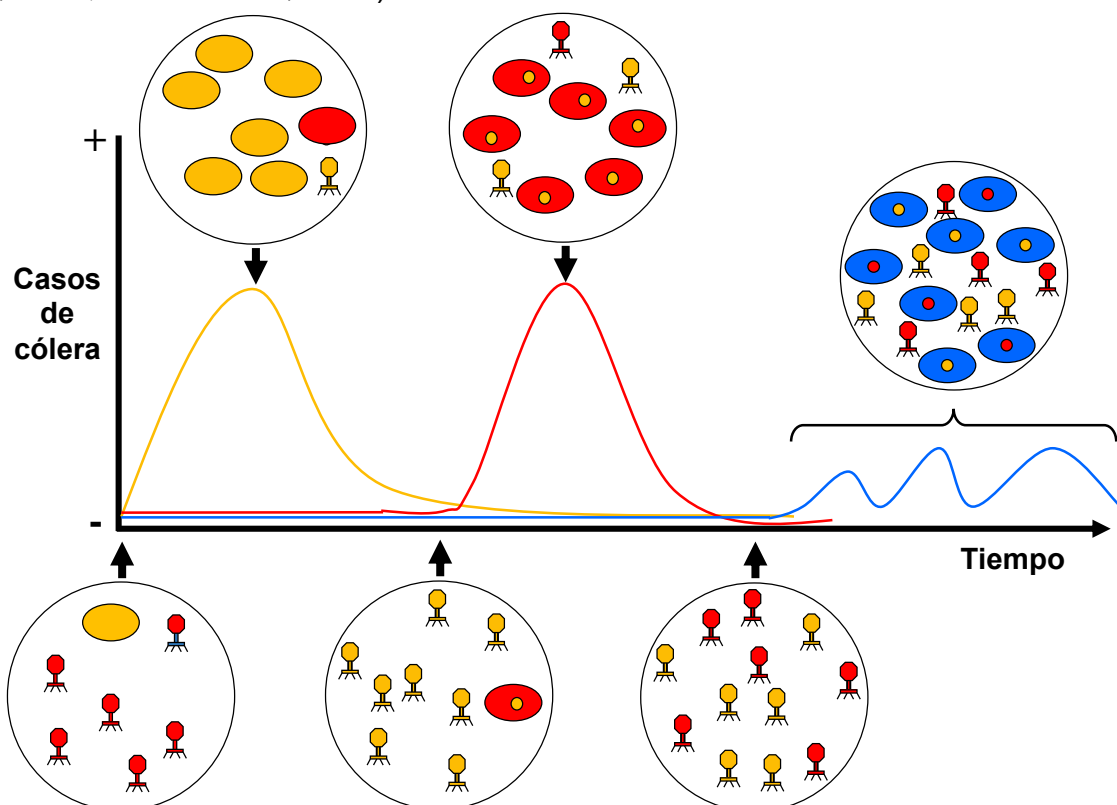
**Tabla 3.** Pandemias de cólera registradas desde 1817 a la fecha.

Pandemia	Publicado por:		Reciente
	Pollitzer, 1954	Wilson, 1984	
1 <sup>a</sup>	1817 – 1823	1817 – 1823	
2 <sup>a</sup>	1829 – 1851	1826 – 1837	
3 <sup>a</sup>	1852 – 1859	1846 – 1862	
4 <sup>a</sup>	1863 – 1879	1864 – 1875	
5 <sup>a</sup>	1881 – 1896	1883 – 1896	
6 <sup>a</sup>	1899 – 1923/5	1899 – 1923	
7 <sup>a</sup>			1961 - ...

Con el comienzo de la séptima pandemia surgió la necesidad de contar con herramientas que permitieran monitorear a la bacteria en el ambiente y poder implementar medidas para prevenir la enfermedad. Así surgió el interés por utilizar a los fagos para éste fin, los cuales han sido ampliamente utilizados como indicadores

para ubicar la presencia de bacterias patógenas en el ambiente (Dukta *et al.*, 1987; Gotuzzo *et al.*, 1994). Adicionalmente, en un estudio realizado por Madico (1996) en Perú, se demuestra la presencia en aguas residuales de fagos específicos de *V. cholerae* O1, señalando la existencia de una correlación entre el número de casos de cólera en las áreas de estudio, la existencia de *V. cholerae* y la presencia de éstos fagos, por lo que se propone el uso de estos fagos para monitorear la presencia de la bacteria.

En un estudio realizado en muestras de agua cercanas a la ciudad de Dhaka, capital de Bangladesh, se observó que existe una relación inversa entre la presencia de fagos específicos y la de *V. cholerae* O1 y O139, así como con la incidencia de casos de cólera en la zona, tal situación muestra que existe una dinámica en la temporalidad y la aparición cíclica de las epidemias de cólera [Figura 6] (Faruque *et al.*, 2005; Jensen *et al.*, 2006).



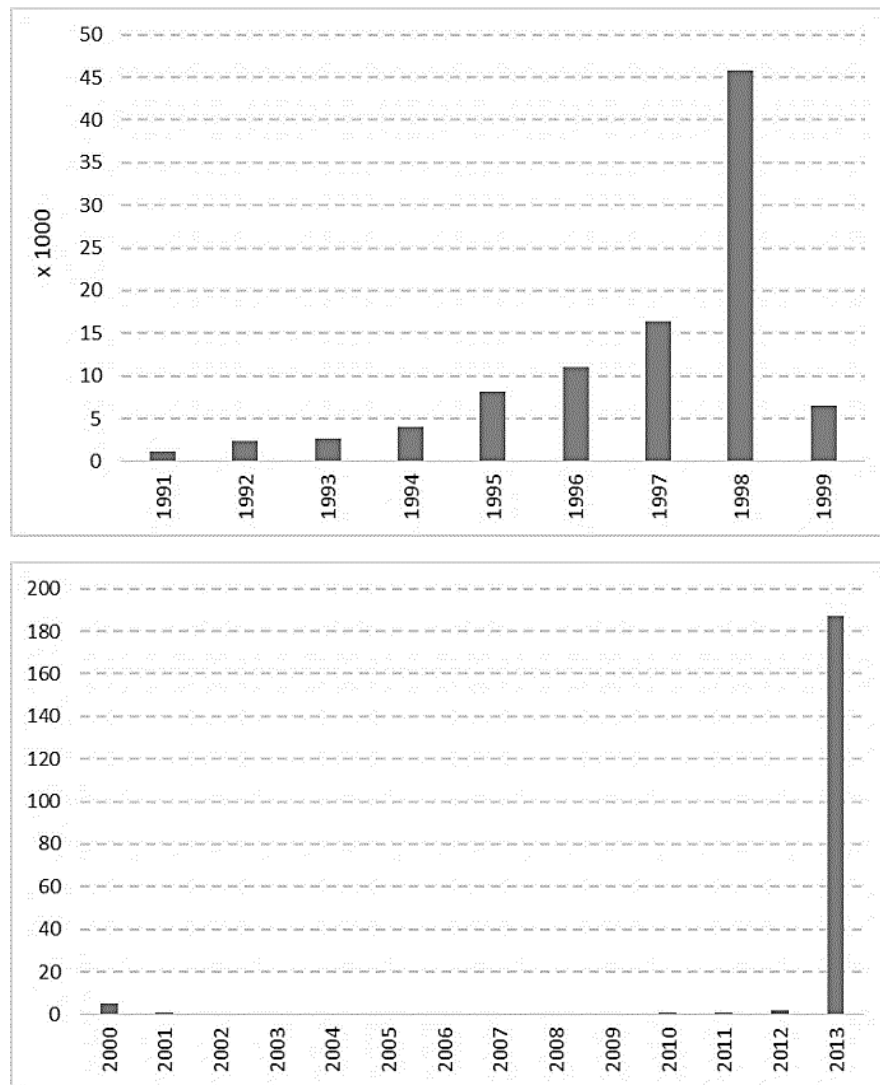
**Figura 6.** (Página anterior) Modelo para explicar la influencia que pueden tener los fagos en el comportamiento cíclico del cólera y la dominancia de serogrupos específicos. La densidad de poblaciones de bacterias y fagos, de tal forma que la ausencia de fagos específicos (amarillos) permite que bacterias de un serogrupo determinado (óvalos amarillos) se desarrollen y den lugar a una epidemia estacional. Sin embargo, eventualmente los fagos (amarillos) se amplifican y aumenta su número en el ambiente, las bacterias de éste serogrupo son lisadas lo que termina con la epidemia. Al mismo tiempo, cepas de la bacteria con un serogrupo diferente (óvalos rojos) resistentes al fago amarillo empiezan a proliferar, lo que origina un nuevo brote de cólera ahora relacionado con el serogrupo rojo, hasta que los fagos rojos se amplifican y lisan a la bacteria lo que permite que la nueva epidemia sea controlada. En la propuesta se sugiere también que en los períodos interepidémicos otros serogrupos (óvalos azules) no virulentos de la bacteria, pero resistentes a los fagos amarillos y rojos proliferan. Estas cepas se adaptan mejor a las condiciones del ambiente que las cepas virulentas, además pueden albergar profagos que atacan a los serogrupos virulentos (ovalos rojos y amarillos) y pueden adquirir factores de virulencia mediante transferencia horizontal. Tal situación da lugar a que las bacterias azules ahora con factores de virulencia proliferen y originen un nuevo serogrupo con potencial epidémico. (Modificado de Faruque *et al.*, 2005).

## **2.6 El cólera hoy en día, un problema no resuelto**

En octubre de 2010 después de casi 100 años, se reporta un brote epidémico de cólera en Haití. En pocas semanas, la enfermedad había sido identificada en todas las provincias del país y para el final del año se habían notificado más de 150 000 casos y 3 500 muertes (Andrews 2011). En Haití los problemas relacionados con la mala infraestructura de agua potable y residual, y el deterioro de éstos por efectos del terremoto que había sufrido la región 9 meses atrás, fueron factores condicionantes para la rápida propagación de la epidemia de cólera. De acuerdo a datos de la organización mundial de la salud, hasta marzo de 2014 se habían reportado más de 700,000 casos y 8 500 muertes (<http://www.who.int/cholera/statistics>).

En México en 2010 el Sistema Nacional de Vigilancia Epidemiológica (SINAVE) reportó un caso de cólera en el municipio de Navolato, Sinaloa y en 2011 otro más en

la localidad de Otatillos, municipio de Badiraguato, Sinaloa (<http://colera.sinave.gob.mx>). En septiembre de 2013 se presentó un brote de cólera asociado con *V. cholerae* O1 El Tor (WHO 2014), que afectó los estados de Hidalgo, Puebla y Veracruz con un total de 187 casos, uno de ellos mortal, el problema se mantuvo hasta el 15 de noviembre de 2013 (Figura 7).



**Figura 7.** Casos de cólera reportados en México desde 1991 hasta 2013.

# 3. Justificación



A pesar de los esfuerzos realizados por la organización mundial de la salud mediante programas de vacunación, prevención y control del cólera, la enfermedad continúa siendo un problema de salud pública sin resolver. Ante tal situación surge la necesidad de desarrollar nuevas estrategias que permitan realizar una prevención y control efectivos del cólera. Aunque, el saneamiento ambiental y la introducción de servicios de agua y drenaje a las comunidades que carecen de dichos servicios, pueden contribuir de manera efectiva al control del padecimiento, su implementación en muchas comunidades resulta complicado. Ante dicho panorama una posible alternativa es la implementación de procedimientos que sean de fácil aplicación, efectivos y sin posibles efectos secundarios. El control biológico de cepas de *V. cholerae* mediante el empleo de fagos específicos con actividad lítica puede ser un procedimiento útil que limite el crecimiento de la bacteria y de manera consecuente la enfermedad. En áreas en las que el cólera es endémico se han venido desarrollando algunos estudios en los que se ha buscado intencionalmente la presencia de fagos con actividad lítica. Sin embargo, en países en los que el cólera no es endémico, pero cuya aparición cíclica representa un problema de salud, no existen estudios respecto a la existencia de fagos líticos específicos de *V. cholerae*. En este trabajo se propone la búsqueda de fagos líticos con actividad específica contra *V. cholerae*, y la posterior caracterización de los mismos incluida la secuencia de su genoma. Los resultados de nuestro trabajo permitirán definir la existencia de este grupo de microorganismos en el ambiente y plantear su posible participación en la aparición cíclica del cólera en nuestro país.

# 4. Hipótesis

En México se ha comprobado la presencia de *Vibrio cholerae* en diversos hábitats acuáticos por lo que en estos lugares existen bacteriófagos específicos para la bacteria que participan regulando la aparición cíclica de la bacteria y de los brotes de cólera.

# 5. Objetivos

**General:**

Realizar el aislamiento y caracterización de fagos líticos de *Vibrio cholerae* O1 a partir de muestras de aguas residuales colectadas en la presa Endhó en Tula en el Estado de Hidalgo.

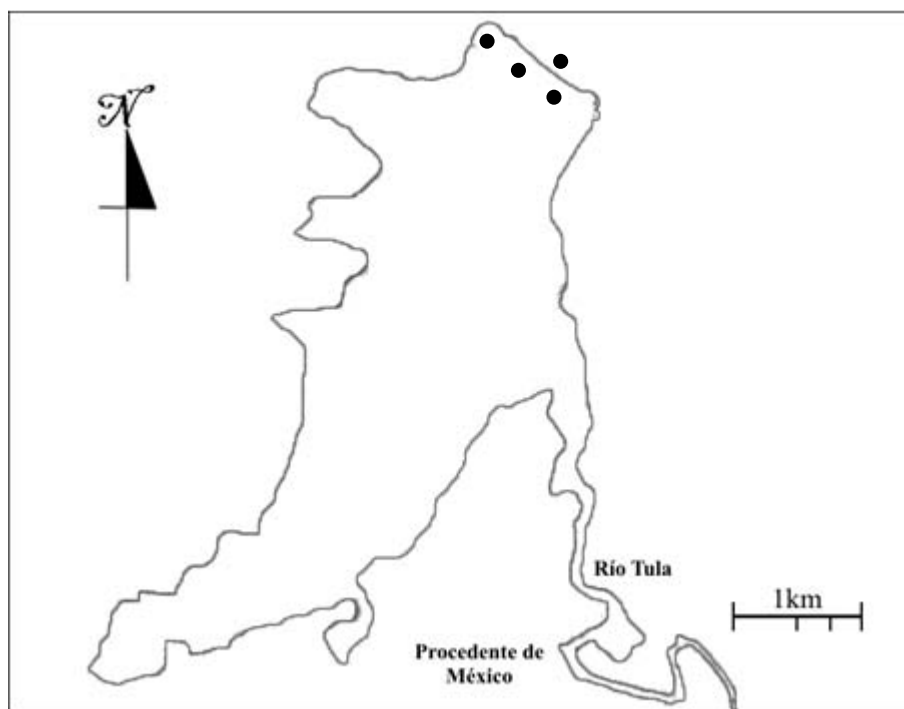
**Particulares:**

- I. Identificar en muestras de agua residual la presencia de bacterias del genero *Vibrio*.
- II. Serotipar los aislados de *V. cholerae*.
- III. Aislar fagos específicos para *Vibrio cholerae* a partir de las mismas muestras de agua.
- IV. Definir la especificidad de los fagos aislados contra cepas de *Vibrio cholerae* O1/O139
- V. Realizar la caracterización morfológica, fenotípica y de su genoma de uno de los fagos líticos específicos aislados.

# **6. Material y métodos**

## 6.1 Área de muestreo

Las muestras fueron colectadas en la presa Endhó en el Estado de Hidalgo. La presa Endhó es un embalse artificial de aguas residuales el cual tiene una extensión aproximada 1,260 hectáreas y una capacidad de más 198 millones de m<sup>3</sup>. Está ubicada en el municipio de Tula a 80 km al norte de la Ciudad de México (Figura 8), y recoge las aguas residuales y pluviales de la zona metropolitana de la Ciudad de México y parte del estado de Hidalgo (Siebe & Cifuentes 1995).



**Figura 8.** La presa se encuentra en la región central de México. Los puntos de muestreo están indicados con círculos negros (●).

## 6.2 Aislamiento de fagos

Cuatro muestras de agua de 200 ml c/u colectadas de la presa Endhó fueron transportadas a temperatura ambiente en botellas de vidrio y procesadas el mismo día.

De cada una de las muestras de agua, se tomaron 50 ml, se centrifugaron a  $16,000 \times g$  (Rotor RC5, Thermo Scientific, Pittsburg, PA, USA) el sobrenadante se filtró a través de una membrana de  $0,22 \mu\text{m}$  (Millex-PES, Millipore, Bedford, MA, USA). El aislamiento inicial de los fagos se realizó mediante un ensayo de doble placa con agar suave (Kropinski *et al.*, 2009). Se utilizaron como hospederos a seis cepas de *V. cholerae*, de las cuales cuatro pertenecen al serogrupo O1 y dos al serogrupo O139 (Tablas 4 y 5). El ensayo se realizó, mezclando 1 ml de cada muestra de agua filtrada y  $100 \mu\text{l}$  de cultivo bacteriano de cada una de las cepas receptoras en fase media logarítmica con 4 ml de agar TOP (triptona 10 g/l, NaCl 10 g/l, agar 7 g/l) a  $52^\circ\text{C}$ ; la mezcla fue posteriormente vertida sobre una placa con agar nutritivo (16 g/l de agar). Finalmente, las placas se incubaron a  $37^\circ\text{C}$  durante 18 h.

### **6.3 Purificación de fagos**

Una vez detectadas zonas de inhibición o placas de lisis en los cultivos bacterianos, se procedió a la propagación de cada uno de los fagos. De cada ensayo, se seleccionaron cinco placas individuales, tomando el bocado de agar donde se encuentra cada placa de lisis con la punta de una pipeta Pasteur estéril. Cada bocado de agar se colocó en 1 ml de medio Luria Bertani suplementado con 10 mM de  $\text{CaCl}_2$  inoculado con *V. cholerae* E8021 ( $\sim 4 \times 10^8$  unidades formadoras de colonia [UFC]). Los cultivos se incubaron durante 18 horas a  $37^\circ\text{C}$ . Los restos fueron eliminados por centrifugación a  $9,000 \times g/10 \text{ min}/4^\circ\text{C}$  (5415C, Eppendorf, Hamburgo, Alemania) y el sobrenadante filtrado por membrana de  $0,22 \mu\text{m}$  (Millipore, Bedford, MA, USA). Posteriormente, se realizó un segundo ciclo de amplificación en mayor volumen (50



ml). Los fagos fueron concentrados por precipitación con polietilenglicol 8,000 (5 %)/NaCl (0,25 M) (Yamamoto *et al.*, 1970), y centrifugando a 16,000 × *g* (RC5- rotor Thermo Scientific, Pittsburg, PA, USA) durante 20 min/4°C, el sedimento que contenía los fagos se resuspendió en 5 ml de buffer SM (Tris – HCl 50 mM, pH: 7; NaCl 100 mM y MgSO<sub>4</sub> 450 mM). Luego, los fagos fueron purificados por centrifugación en gradiente de CsCl<sub>2</sub> a 100,000 × *g*/18 h/4° C (rotor 41Ti, Beckman, Brea, CA, USA) y posteriormente dializados utilizando membrana de diálisis de 12-14, 000 MWCO (Spectra/Pore, Spectrum Laboratories, Rancho Dominguez, CA, USA). Una vez durante 30 min en 1000 volúmenes de buffer de diálisis A (NaCl 3 M; Tris -HCl 0.1 M, pH: 7.4); posteriormente 2 veces de 30 min c/u en 1,000 volúmenes de buffer de diálisis B (NaCl 0.3 M; Tris-HCl 0.1, pH: 7.4).

#### **6.4 Identificación bacteriana**

El aislamiento de *V. cholerae* se realizó mediante la técnica de enriquecimiento. De las muestras colectadas en la presa Endhó, utilizadas para aislar los fagos, se tomaron 10 ml de cada una para inocular 50 ml de agua de peptona alcalina (DeWitt *et al.*, 1971) y se incubaron a 37°C durante 6 h. Posteriormente, se tomaron alícuotas de la biopelícula superficial y se sembraron en agar Tiosulfato-Citrato-Bilis–Sacarosa [TCBS] (Becton-Dickinson, Sparks, MD, USA) y se incubaron a 37°C durante 18h. Se seleccionaron colonias amarillas y se identificaron mediante el sistema automatizado Vitek, utilizando la tarjeta GNI-Plus (bioMérieux, Marcy l' Etoile, Rhône, Francia). Las cepas identificadas como *V. cholerae* fueron tipificadas con sueros de conejo anti-O1

y anti-O139 en el laboratorio de patógenos entéricos de la Facultad de Medicina-UNAM.

### **6.5 Análisis de rango de hospederos**

La especificidad de ØVC8 fue probado contra 47 cepas bacterianas, de las cuales, 22 corresponden a *V. cholerae* O1, 3 a *V. cholerae* O139, 6 a *V. cholerae* No-O1/No-O139, 4 a *V. alginoliticus*, 7 a *V. parahaemolyticus*, 1 a *V. metschnikovii*, 3 a *Aeromonas veronii* y 1 *Escherichia coli* (Tablas 4 y 5). En análisis se realizó mediante un ensayo de gota en placa, agregando 10 µl de suspensión de fagos ( $10^8$  UFP/ml) sobre una placa de agar inoculada con 60 µl de cultivo bacteriano en fase media logarítmica, de cada una de las cepas receptoras. Las placas se incubaron durante 18 h/37° C.

### **6.6 Microscopía electrónica**

2 µl de una suspensión de ØVC8 purificada ( $\geq 1 \times 10^8$  UFP/ml) se colocaron en una rejilla de cobre recubierta con Formvar (EMS, Hatfield, PA, USA) y se dejó absorber durante 2 min, se retiró el exceso con papel filtro y se agregaron 6 µl de acetato de uranilo al 2% (EMS, Hatfield, PA, USA), se dejó adsorber por 2 min y el exceso se retiró con papel filtro, la rejilla se dejó secar al aire y se examinó con microscopio electrónico de transmisión JEM- 1200 EXII (JEOL, Tokio, Japón).

## **6.7 Análisis del Genoma**

### **6.7.1 Extracción de DNA**

El DNA total de ØVC8 fue extraído siguiendo el protocolo de extracción con fenol cloroformo (Sambrook & Russell, 2001). Una alícuota de 500µl de suspensión viral ( $1 \times 10^8$ /ml UFP) fue tratada con 100 U de DNasa I (Invitrogen, Carlsbad, CA, USA) y 50 µg de RNasa A (Sigma, St Louis, MO, USA) y se incubó 2 h/37°C. Posteriormente, se añadió 2 µl de proteinasa K (20mg/ml) (Vivantis, Oceanside, CA, USA) y 50 µl de SDS (10 %) y se incubó 1 h/56°C. Los fragmentos celulares fueron eliminados mediante extracción con fenol - cloroformo (1:1), el DNA fue precipitado con etanol frío y se suspendido en agua libre de DNAsas.

### **6.7.2 Análisis restricción de DNA**

Para el análisis de restricción, 500 µg de DNA purificado fue digerido con las endonucleasas de restricción *EcoRI*, *EcoRV*, *HincII*, *HindIII*, *HpaI*, *PstI*, *SspI* (New England BioLabs, Ipswich, MA, USA); *AluI*, *HinfI*, *DraI*, *HpaII*, *RsaI* (Invitrogen, Carlsbad, CA, USA); *HaeIII* y *XbaI* (Fermentas, Pittsburgh, PA, USA) de acuerdo con las recomendaciones del fabricante. Los fragmentos de DNA se resolvieron en geles de agarosa (1%), teñidos con bromuro de etidio y visualizados con luz UV. Para determinar el peso molecular de cada fragmento de restricción, se utilizó el programa Labworks versión 4.5 (UVP, Upland, CA, USA).

### 6.7.3 Clonación de fragmentos de restricción

Los fragmentos de DNA obtenidos con la enzima de restricción HpaII fueron clonados siguiendo el procedimiento descrito por Sambrook y Russell (2001), utilizando el vector pBluescript II (Fermentas, Pittsburgh, PA, USA) linearizado con la enzima ClaI (Invitrogen, Carlsbad, CA, USA). Posteriormente, células competentes TOP10 de *Escherichia coli* (Invitrogen, Carlsbad, CA, USA) fueron transformadas con la construcción pBluescript-vc8 y seleccionadas por color azul-blanco en placas de agar LB suplementado con ampicilina (100 µg/ml, concentración final) (IBI Scientific, Peosta, IA, USA), IPTG (2,5 mg/ml) (Invitrogen, Carlsbad, CA, USA) y Xgal (10 mM) (Invitrogen, Carlsbad, CA, USA). Los plásmidos recombinantes fueron secuenciados con el secuenciador ABI Prism 310 utilizando el kit ABI Prism Big Dye Terminator (Applied Biosystems, Grand Island, NY). Las secuencias obtenidas fueron analizadas utilizando el programa BLASTX (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

### 6.7.4 Secuenciación

El genoma de ØVC8 fue secuenciado utilizando el sistema Genome Analyzer Iix (Illumina, San Diego, California, USA) en la Unidad Universitaria de Secuenciación Masiva (UUSM) del Instituto de Biotecnología de la UNAM en Cuernavaca, Morelos (<http://uusmd.unam.mx>). El ensamble *de novo* y la anotación genética del genoma de ØVC8 fueron realizados por la empresa Winter Genomics (<http://www.wintergenomics.com>).

El genoma fue ensamblado *de novo* utilizando los programas Velvet (Zerbino & Birney, 2008), Abyss (Simpson *et al.*, 2009) y SOAP (Li *et al.*, 2010). Los resultados

fueron combinados utilizando el programa Minimus (Sommer *et al.*, 2007) en un solo “contig”. Para anotar el genoma de ØVC8, se utilizó una estrategia basada en un proyecto de secuenciación de alto rendimiento (Henn *et al.*, 2010). En primer lugar, se utilizó BLAST-X y PFAM/TIGR para identificar las posibles secuencias codificantes comparando el genoma de ØVC8 contra bases de datos de proteínas no redundantes y de familias de proteínas. Posteriormente se usaron los programas computacionales de predicción de genes, Glimmer3, MetaGeneAnotator, GeneMarkS, ZCURVE\_V y EasyGene, para identificar los ORFs tentativos. Por último, todas las predicciones fueron agrupadas en un único locus. Adicionalmente, se realizó la búsqueda de tRNAs y RNAs ribosomales usando tRNAscan-SE, Rfam y RNAmmer, mientras que la predicción de promotores se realizó utilizando el programa Promotor Hunter (Klucar *et al.*, 2010) y los posibles terminadores de la transcripción fueron predichos usando WebGeSTer (Mitra *et al.*, 2011).

## **6.8 Análisis proteómico**

La separación de las proteínas estructurales de ØVC8 se realizó a partir de 500 µl de suspensión del fago ( $>1 \times 10^{10}$  UFP/ml). Los fagos fueron precipitados con PEG/NaCl como se mencionó anteriormente, y la pastilla resuspendida en solución de Laemmli (65.8 mM Tris-HCl, pH 6.8, 2.1% SDS, 26.3% (w/v) glicerol, 0.01% azul de bromofenol y 100 mM β-mercaptoetanol), se calentó a 100°C/5 min y posteriormente la muestra fue resuelta en gel de poliacrilamida-dodecil sulfato de sodio 10% (SDS-PAGE) teñido con azul de Coomassie (0.1% Coomassie Brilliant Blue R-250, 50% metanol y 10% ácido acético glacial) (Boulanger 2009).

Posteriormente se seleccionaron bandas aisladas de proteínas, las cuales fueron cortadas del gel, desteñidas, reducidas, carbaminometiladas y digeridas con tripsina porcina modificada (Promega, Madison, WI, USA). El volumen del extracto se redujo por evaporación al vacío a temperatura ambiente y después se ajustó a 20  $\mu$ l con ácido fórmico al 1%. El análisis de los péptidos se llevó a cabo en el Departamento de Bioquímica de la Facultad de Medicina-UNAM, utilizando un espectrómetro de masas QTRAP 3200 (Applied Biosystems/MDS Sciex, ON, Canadá), equipado con una fuente de nanoelectropulverización y un sistema LC de nanoflujo (Agilent 1100 Bomba Nano, Waldbronn, Alemania).

### **6.9 Análisis filogenético**

El análisis filogenético se realizó alineando la secuencia del ORF3 (subunidad mayor de la terminasa) con secuencias individuales relacionadas en fagos del orden *Caudovirales* obtenidas de GenBank. El alineamiento se hizo usando el programa ClustalW2 y los resultados fueron exportados al programa Mega 6.0, la construcción de los árboles se hizo por el método de neighbor-joining con una distribución gamma ( $\gamma=2$ ) y 1,000 réplicas de bootstrap, usando una corrección de distancias de Poisson.

# 7. Resultados

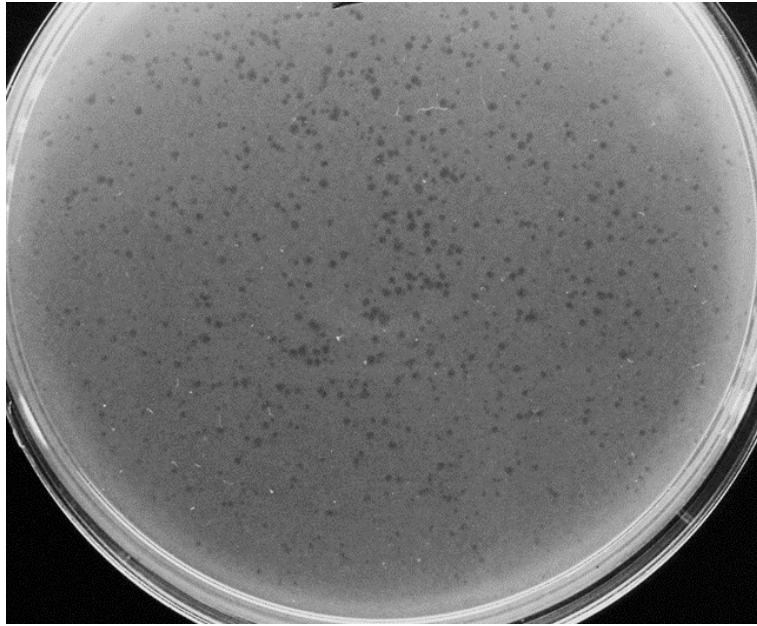
## 7.1 Aislamiento de fagos y bacterias

A partir de las muestras de agua obtenidas de la presa Endhó se aislaron 13 bacterias y 16 fagos. Las bacterias aisladas fueron identificadas como *Vibrio cholerae* No-O1/No-O139, cuatro *Vibrio alginoliticus* y tres *Aeromonas veronii* (Tabla 5). Las 13 bacterias se utilizaron para evaluar la especificidad de los fagos aislados. Los 16 fagos se obtuvieron con las cepas 96316 que presentó placas claras y opacas, CVD103 que presentó solo placas opacas y E8021 que presentó solo placas claras (Tabla 4). A cada uno de ellos se analizó el rango de hospedero sobre diferentes cepas de *Vibrio cholerae* O1/O139 y No-O1/No-O139. (Tablas 4 y 5).

## 7.2 Actividad lítica y especificidad de hospedero de los fagos aislados en la presa Endhó

En el ensayo para conocer la capacidad de infección de los fagos aislados se emplearon 53 cepas bacterianas (Tablas 4 y 5), en el 58.5% (31/53) de las bacterias analizadas no se observó la formación de algún tipo de placa (Tabla 5), mientras que en las 22 (41.5%) cepas restantes, se identificó la formación de placas claras y opacas. Al realizar la infección de los fagos en las cepas de *V. cholerae*, se observó que 12 de los 16 fagos generaban placas claras y opacas, esto solo en cepas de los serogrupos O1 y O139 (Tabla 4). Los fagos fueron nombrados con números consecutivos y solo el designado como ØVC8 mostró actividad en las 13 cepas de *V. cholerae* O1 (Clásicas y El Tor), produciendo placas claras de aproximadamente 1mm de diámetro (Figura 9).





**Figura 9.** Ensayo de doble placa con agar suave, donde se observa la morfología de las placas de lisis producidas por el fago ØVC8 sobre un cultivo bacteriano de la cepa E8021 de *V. cholerae* O1.

**Tabla 4.** Perfiles de infección de los fagos aislados de la presa Endhó en cepas de *V. cholerae*.

Cepa	Organismo	Fuente / Referencia	Fagos															
			ØVC1	ØVC2	ØVC3	ØVC6	ØVC7	ØVC8	ØVC9	ØVC11	ØVC18	ØVC19	ØVC20	ØVC28	ØVC30	ØVC31	ØVC33	ØVC34
87128	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87152	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87177	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87210	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87258	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87269	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87284	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87285	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87662	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88226	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88538	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88539	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88682	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88697	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
89207	<i>V. cholerae</i> O139	Tailandia <sup>1</sup>																
90500	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
<b>96316</b>	<i>V. cholerae</i> O139	México																
102223	<i>V. cholerae</i> O1	Republica Checa <sup>1</sup>																
97639/1	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
<b>CVD 103</b>	<i>V. cholerae</i> O1	1																
<b>E8021</b>	<i>V. cholerae</i> O1	HPA <sup>2</sup>																
N-16961	<i>V. cholerae</i> O1	3																

Placas opacas

Placas claras

Las cepas utilizadas en el aislamiento inicial de los fagos se muestran subrayadas y en negritas.

<sup>1</sup> Cepas aisladas de diferentes fuentes (y/o) proporcionadas por el Laboratorio de Bacteriología del Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional Autónoma de México.

<sup>2</sup> Health Protection Agency, Laboratory of Gastrointestinal Pathogens, London, England

<sup>3</sup> Provided by Dr. Shah M. Faruque of the Centre for Food and Water Borne Diseases at the International Centre for Diarrheal Disease Research, Bangladesh (ICDDR).

**Tabla 5.** Cepas que no fueron susceptibles a la infección por fagos.

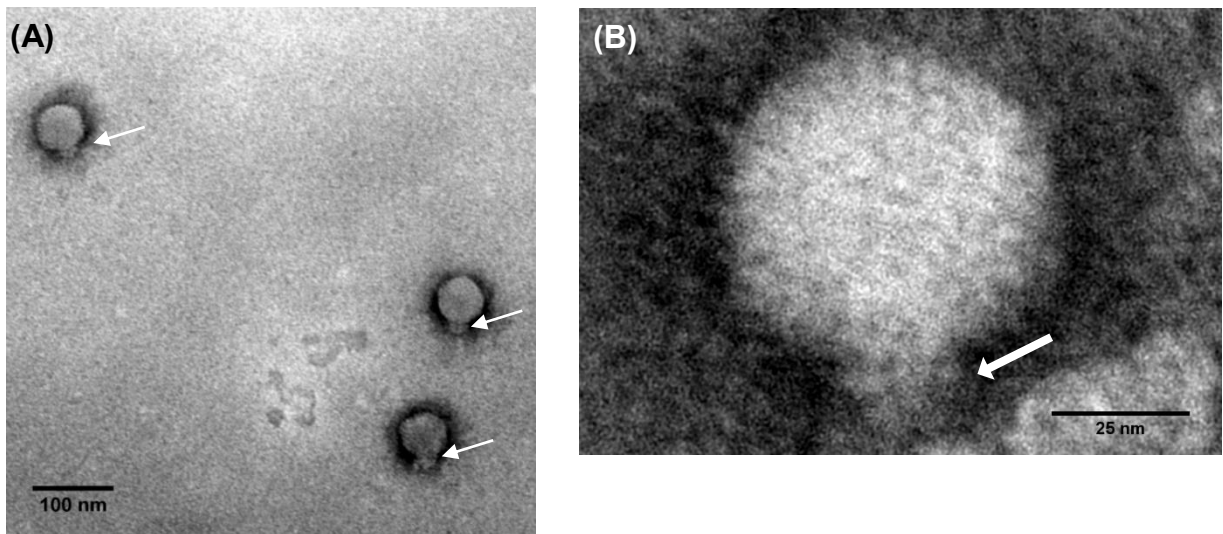
Cepa	Organismo	Fuente/Referencia
<b><u>2740</u></b>	<i>V. cholerae</i> O1	México <sup>1</sup>
<b><u>87145</u></b>	<i>V. cholerae</i> O1	<sup>1</sup>
87167	<i>V. cholerae</i> O1	México <sup>1</sup>
88229	<i>V. cholerae</i> O139	CDC Atlanta <sup>1</sup>
<b><u>88678</u></b>	<i>V. cholerae</i> O139	México <sup>1</sup>
88758	<i>V. metschnikovii</i>	México <sup>1</sup>
88845	<i>V. parahaemolyticus</i>	México <sup>1</sup>
88846	<i>V. parahaemolyticus</i>	México <sup>1</sup>
88859	<i>V. parahaemolyticus</i>	México <sup>1</sup>
89812	<i>V. parahaemolyticus</i>	México <sup>1</sup>
96558	<i>V. parahaemolyticus</i>	México <sup>1</sup>
97259	<i>V. parahaemolyticus</i>	México <sup>1</sup>
97276	<i>V. parahaemolyticus</i>	México <sup>1</sup>
AI-1852	<i>V. cholerae</i> O139	Faruque <i>et al.</i> , 2005
C-19385	<i>V. cholerae</i> O1	Faruque <i>et al.</i> , 2005
DK – 238	<i>V. cholerae</i> O1	<sup>1</sup>
DK – 60	<i>V. cholerae</i> O1	<sup>1</sup>
E1_1	<i>V. alginolyticus</i>	Este estudio
E1_5	<i>Aeromonas veronii</i>	Este estudio
E1_6	<i>V. alginolyticus</i>	Este estudio
E1_7	<i>Aeromonas veronii</i>	Este estudio
E1_8	<i>Aeromonas veronii</i>	Este estudio
E1_9	<i>V. alginolyticus</i>	Este estudio
E3_1	<i>V. cholerae</i> No-O1	Este estudio
E3_10	<i>V. cholerae</i> No-O1	Este estudio
E3_2	<i>V. cholerae</i> No-O1	Este estudio
E3_3	<i>V. cholerae</i> No-O1	Este estudio
E3_4	<i>V. alginolyticus</i>	Este estudio
E3_5	<i>V. cholerae</i> No-O1	Este estudio
E3_8	<i>V. cholerae</i> No-O1	Este estudio
HB101	<i>Escherichia coli</i>	México <sup>1</sup>

Las cepas utilizadas en el aislamiento inicial de los fagos se muestran subrayadas y en negritas.

<sup>1</sup>Cepas aisladas de diferentes fuentes (y/o) proporcionadas por el Laboratorio de Bacteriología del Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional Autónoma de México.

### 7.3 Análisis morfológico de ØVC8

El análisis por MET del fago ØVC8 mostró la presencia de una cápside isométrica de ~62nm y una cola corta de ~16 nm (Figura 10). Ambas características permiten incluir a este fago en el orden caudovirales (fagos con cola) dentro de la familia *Podoviridae* (fagos con cola corte).



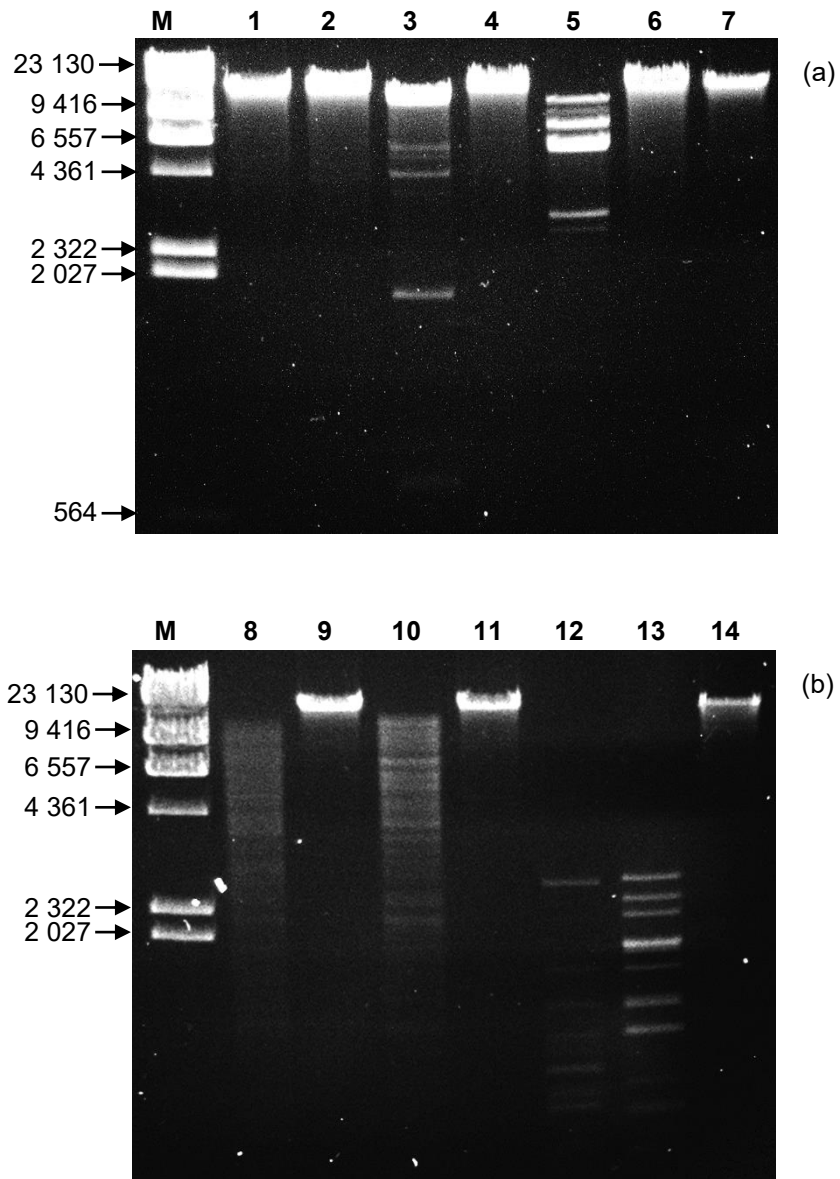
**Figura 10.** Micrografías electrónicas del fago ØVC8 aislado de muestras de agua de la presa Endhó. (A), Cápside isométrica y una cola corta marcadas por flechas blancas. (B) Amplificación de la imagen en la que se observa con mayor detalle la estructura de la cápside y la cola (flecha blanca). (A) 15,000 ×, (B) 50,000 ×. Tinción negativa con acetato de uranilo (2%).

### 7.4 Caracterización del genoma de ØVC8

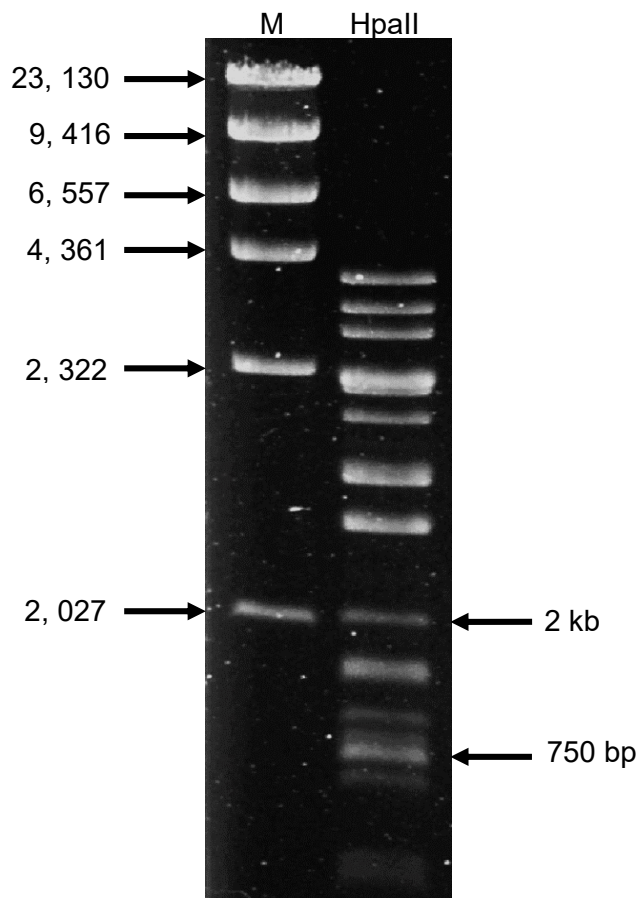
Como primer paso para conocer las características generales del genoma de ØVC8 su material genético fue tratado con DNasa I y Exonucleasa III, el resultado al respecto mostró que se trata de un genoma constituido por una cadena doble de DNA. Inicialmente para conocer el tamaño aproximado del genoma, el DNA purificado se trató con las enzimas de restricción PstI, EcoRV, EcoRI, HincII, HaeIII, RsaI, AluI, y HpaI. Los fragmentos de restricción obtenidos fueron separados por electroforesis

cuyos diferentes perfiles electroforéticos se muestran en la figura 11. El cálculo aproximado del tamaño del genoma del fago se hizo con la suma de los pesos de los diferentes fragmentos obtenidos en la restricción con las enzimas EcoRV y EcoRI. Así se definió que el tamaño del genoma del fago ØVC8 es de aproximadamente 39 kb.

En un primer intento por conocer la posible relación filogenética de ØVC8 con otros fagos secuenciados; dos fragmentos, uno de 750 pb y otro de 2 kb, obtenidos al digerir el DNA con la enzima HpaII fueron clonados y secuenciados (Figura 12). El análisis comparativo de las secuencias traducidas mediante BLAST-X, reveló una secuencia de 533 nt dentro del fragmento de 750 pb (GenBank, número de acceso: KM\_359784), cuya secuencia traducida presenta homología (93%) con la subunidad mayor de la proteína terminasa del fago VP2 de *V. cholerae*, y con menor homología con proteínas hipotéticas de diferentes organismos. Mientras que, dentro del fragmento de 2 kb, una secuencia de 626 nt pb (GenBank, número de acceso: KM\_359783), presentó homología (98%) con proteínas hipotéticas de los fagos VP2 y VP5 de *V. cholerae*, ambos pertenecientes a la familia *Podoviridae*.



**Figura 11.** Patrones de restricción de DNA de ØVC8 tratado con 14 enzimas de restricción. Carriles: 1, HpaI; 2, PstI; 3, EcoRV; 4, SspI; 5, EcoRI; 6, HincII; 7, HindII; 8, HaeIII; 9, XbaI; 10, RsaI; 11, DraI; 12, AluI; 13, HpaII; 14, Hinf I. M: marcador de peso molecular  $\lambda$  DNA/HindIII. Gel de agarosa 1% teñido con bromuro de etidio.



**Figura 12.** Patrón electroforético del DNA total de ØVC8 tratado con la enzima HpaI. Las flechas señalan los fragmentos que fueron secuenciados. El fragmento de 2 kb corresponde a una proteína hipotética (ORF17), mientras que el fragmento de 750 kb corresponde a la subunidad mayor de una terminasa (ORF 3). Gel de agarosa (1%), M: marcador de peso molecular  $\lambda$  DNA/HindIII.

### 7.5 Secuenciación del genoma de ØVC8

La secuencia completa del genoma de ØVC8, fue ensamblado en una sola secuencia a partir de un total de ~16 millones de secuencias de 36 pares de bases, con una cobertura de ~14,324  $\times$ . El análisis de la secuencia completa reveló que el genoma tiene 39, 422 pb de longitud con un contenido de G+C del 50,8%. Por otro lado, el análisis comparativo de secuencia de DNA, muestra que ØVC8 tiene ~85% de similitud con los fagos VP2 y VP5.

En el análisis bioinformático de la secuencia del genoma de ØVC8 se identificaron 48 posibles ORFs. De estos, el 30% (14/48) fueron anotados con una función conocida de acuerdo con la similitud que tuvieron con secuencias de otros fagos u organismos, el 70.83% (34/48) correspondieron a proteínas hipotéticas descritas en otros fagos y un solo ORF, el 2.08% (1/48) no tuvo similitud con ninguna secuencia reportada previamente (Tabla 6). Los 48 ORFs se encuentran distribuidos en ambas cadenas del DNA de ØVC8, una de ellas con 22 ORFs relacionados con funciones de empaquetamiento y proteínas estructurales, mientras que la cadena complementaria codifica para 26 ORF asociados con funciones de metabolismo, replicación y una región de función desconocida (Figura 13). En el mismo análisis no se identificó la presencia de tRNAs ni de rRNAs. La secuencia completa del genoma de ØVC8 fue depositada en la base de datos GenBank con el número de acceso: JF\_712866.

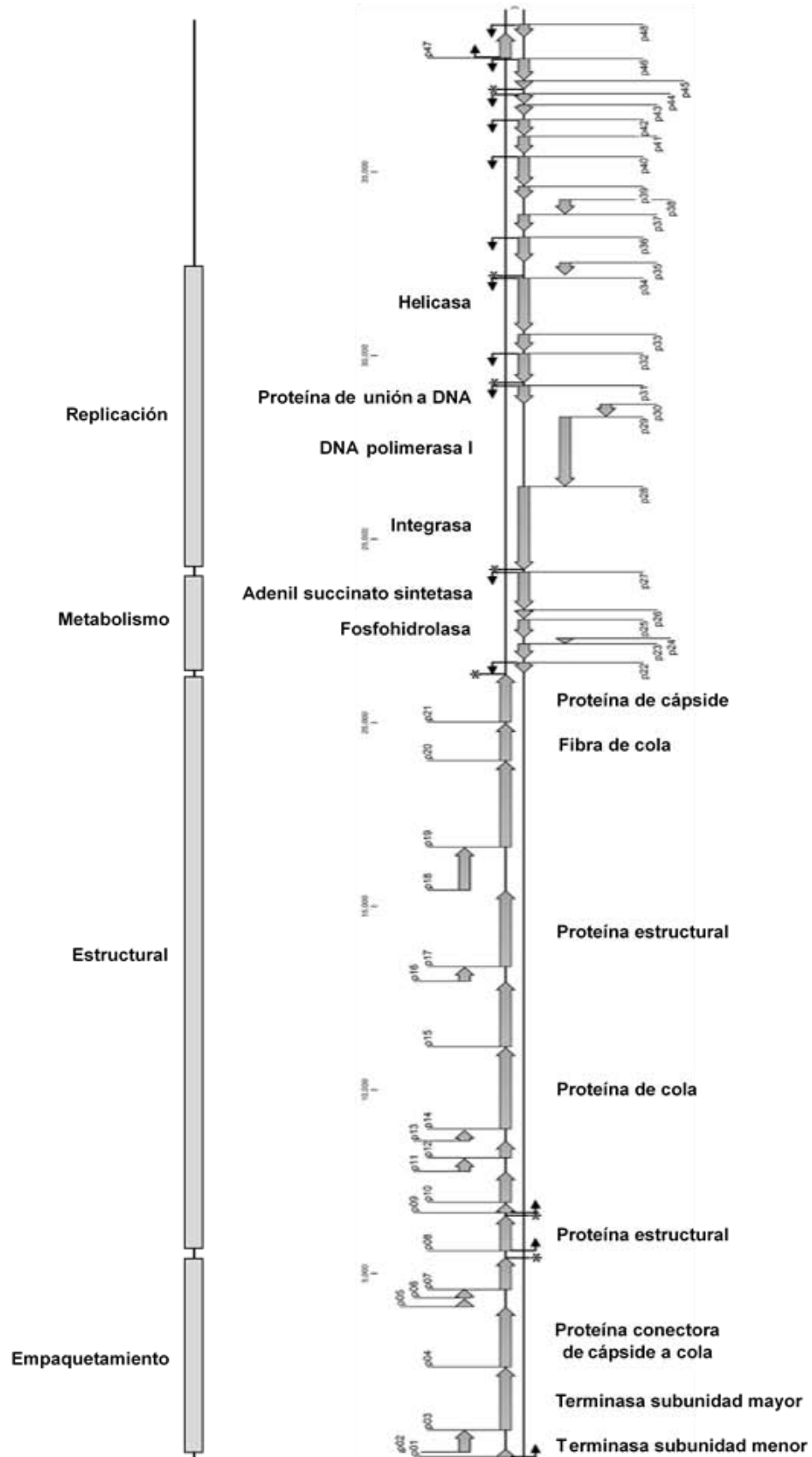
**Tabla 6.** Marcos de lectura abierta (ORFs) putativos en el genoma de ØVC8 y su función asignada.

ORF	Posición		Producto		Función	Fago/organismo relacionado	GenBank N° de ascensión	BLAST X	
	Inicio	Término	Tamaño (aa)	Peso molecular (kDa)				% Identidad	e-value
1	1	222	73	13.6	Proteína hipotética	Vibrio fago VP2	YP_052974	95	1 e-32
2	128	739	203	22.9	Terminasa subunidad menor	Vibrio fago VP2	YP_053007	94	4 e-131
3	726	2432	568	65.3	Terminasa subunidad mayor	Pelagibacter fago HTVC010P	YP_007517700	32	4 e-41
4	2443	4086	547	61.7	Proteína de conexión cápside-cola	Pelagibacter fago HTVC010P	YP_007517703	32	4 e-62
5	4086	4313	75	8	Proteína hipotética	Vibrio fago VP2	YP_052976	92	7 e-26
6	4326	4583	85	9.9	Proteína hipotética	Vibrio fago VP2	YP_052977	95	2 e-18
7	4552	5439	295	31.8	Proteína hipotética	Pelagibacter fago HTVC010P	YP_007517705	23	6 e-14
8	5607	6575	322	36.1	Proteína estructural	Pelagibacter fago HTVC010P	YP_007517707	24	1 e-21
9	6642	6911	89	9.2	Proteína hipotética	Vibrio fago VP2	YP_052980	100	7 e-35
10	6926	7774	282	31.1	Proteína hipotética	Vibrio fago VP5	YP_053013	91	1 e-46
11	7771	8136	121	13.4	Proteína hipotética	Vibrio fago VP2	YP_052982	96	1 e-59
12	8138	8611	157	18.2	Proteína hipotética	Vibrio fago KVP40	NP_899538	31	2 e-20
13	8604	8906	100	10.5	Proteína hipotética	Vibrio fago VP2	YP_052983	36	3 e-05



Continúa **Tabla 6.**

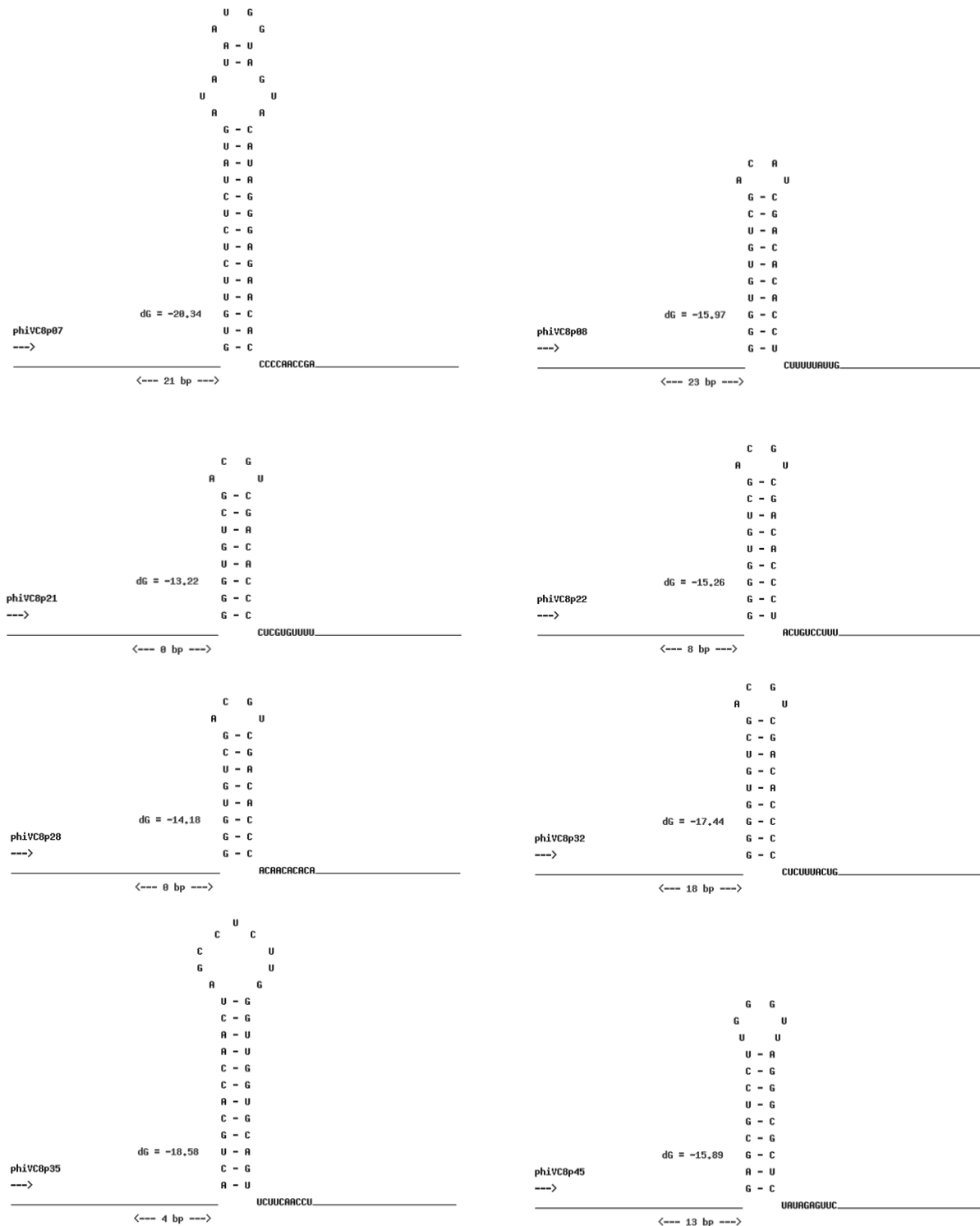
ORF	Posición		Producto		Función	Organismo relacionado	BLAST X		
	Inicio	Término	Tamaño (aa)	Peso molecular (kDa)			GenBank	BLAST X	
							N° de ascensión	% Identidad	e-value
14	8930	11155	741	74.1	Proteína de cola	Vibrio fago VP5	YP_024980	90	0
15	11164	12948	594	65.5	Proteína estructural	Vibrio fago VP5	YP_024422	98	0
16	12951	13349	132	13.9	Proteína hipotética	Vibrio fago VP2	YP_052984	97	1 e-41
17	13346	15430	694	76.8	Proteína estructural	Vibrio fago VP2	YP_052985	95	0
18	15430	16602	390	41.5	Proteína hipotética	Vibrio fago VP5	YP_053019	95	2 e-151
19	16604	18952	782	87.2	Proteína estructural	Escherichia fago phiV10	YP_512274	27	6 e-68
20	18956	19984	342	36.6	Fibra de cola	Vibrio fago VP5	YP_053020	96	2 e-162
21	20009	21322	437	44.9	Proteína de cápside	Vibrio fago VP2	YP_024425	31	1 e-35
22	21350	21646	98	11.2	Proteína hipotética	Pseudomonas fago PPpW-3	YP_008873205	45	9 e-16
23	21723	22160	145	11.2	Proteína hipotética	Vibrio fago VP2	YP_053022	96	2 e-79
24	22150	22308	52	11.2	Proteína hipotética	Vibrio fago VP2	YP_052988	81	1 e-19
25	22305	22817	170	19.2	Metal-dependiente fosfohidrolasa	Vibrio fago VP5	YP_024983	94	7 e-116
26	22819	23079	86	9.8	Proteína hipotética	Vibrio fago VP5	YP_053024	96	2 e-28
27	23081	24112	343	38.1	Adenil succinato sintetasa	<i>Solibacter usitatus</i>	YP_827486	33	1 e-31
28	24154	26442	762	87	Integrase	Vibrio fago VP2	YP_024428	99	0
29	26432	28333	633	71.2	DNA Polimerasa I	$\alpha$ -proteobacteria fago ØJL001	YP_223952	26	9 e-40
30	28330	28680	116	13.5	Proteína hipotética	Enterococcus fago EFDG1	AJP_61480	32	4 e-09
31	28677	29192	171	19.2	Proteína de unión a ssDNA	Vibrio fago VP2	YP_024430	93	8 e-95
32	29253	30068	271	30.3	Proteína hipotética	Pseudomonas fago F8	YP_001294468	32	1 e-23
33	30119	30586	155	17.8	Proteína hipotética	Vibrio fago VP2	YP_052991	98	5 e-108
34	30649	32118	489	55.4	DNA/RNA Helicasa de la superfamilia II	Thermoanaerobacterium fago THSA-485A	YP_006546319	36	2 e-77
35	32188	32538	116	12.6	Proteína hipotética	Vibrio fago VP5	YP_053027	84	1 e-26
36	32538	33230	230	24.5	Proteína hipotética	Vibrio fago CJY	AIZ01434	97	3 e-111
37	33381	33845	154	17.3	Proteína hipotética	Vibrio fago VP2	YP_052993	84	6 e-93
38	33842	34252	136	15.7	Proteína hipotética	Vibrio fago VP2	YP_052994	78	2 e-64
39	34254	34613	119	13.7	Proteína hipotética	Vibrio fago VP5	YP_053031	67	4 e-44
40	34621	35421	266	30.2	Proteína hipotética	Vibrio fago VP2	YP_052996	98	2 e-101
41	35490	35975	161	18.1	Proteína hipotética	Vibrio fago VP2	YP_052977	96	5 e-85
42	35985	36437	150	16	Proteína hipotética	Vibrio fago VP5	YP_053036	84	2 e-61
43	36543	36836	97	11	Proteína hipotética	Vibrio fago VP2	YP_053001	95	1 e-46
44	36848	37138	96	10.7	Proteína hipotética	Vibrio fago VP2	YP_053002	79	5 e-36
45	37245	37490	81	9.2	Proteína hipotética	Vibrio fago VP2	YP_053003	74	1 e-26
46	37495	38094	199	22.1	Proteína hipotética	Vibrio fago VP5	YP_053042	91	5 e-93
47	38093	38788	231	26.9	Proteína hipotética	Vibrio fago VP2	YP_053005	93	1 e-50
48	38669	39028	121	13.6	Proteína hipotética	Desconocido	---	---	---



**Figura 13.** (Página anterior) Organización del genoma de ØVC8. El fago ØVC8 posee un genoma linear el cual codifica para 48 posibles genes. El inicio fue asignado de manera arbitraria al ORF 1. Las flechas gruesas indican los genes predichos y su dirección de transcripción. La posición y dirección de los promotores está indicada por las flechas delgadas y los terminadores Rho-independientes con asteriscos. Los rectángulos señalan los módulos funcionales propuestos para el genoma de ØVC8.

### **7.5.1 Organización funcional del genoma de ØVC8**

El análisis del genoma de ØVC8 reveló la presencia de 15 secuencias promotoras hipotéticas presentes en regiones no codificantes (Tabla 7). Tres en la cadena positiva al inicio de los módulos funcionales de empaquetamiento y morfogénesis, mientras que los 12 restantes se localizan en la cadena complementaria. De estos, 3 en el módulo de replicación, 2 en el módulo de metabolismo y 7 en la región de función desconocida. Además, se identificaron 8 terminadores de tipo Rho-independiente, localizados uno en el módulo de empaquetamiento, 2 en el módulo de morfogénesis, 2 en el módulo de replicación, uno en el módulo de metabolismo y los tres restantes en la región de función desconocida (Figura 14).



**Figura 14.** Predicción de la estructura secundaria de los terminadores Rho-independientes en el genoma de ØVC8 usando el programa WebGeSTer. A la izquierda de cada una de las estructuras se muestran el ORF al cual está asociado el terminador, la distancia en pares de bases que separan el terminador del codón de término y la energía libre (dG).

**Tabla 7.** Regiones promotoras predichas en el genoma de ØVC8 usando el programa PromoterHunter. Las regiones marcadas en gris muestran las secuencias -10 y -35.

Promotor	ORF	Posición	5'	- 35	- 10	3'	
1	1	39373...39422	ATGTAACGGT	GGTTGACA	CAGAGCCAGAGGTGTG	CTATAAT	GGAGGTAAG
2	8	5551...5601	AGAGTCCGGA	AACGGGTA	GCTCAAATCGATATTA	ACAATCT	TTTAGGAAAA
3	9	6593...6641	GATTAGGG	TGTGTCGACA	TCGACACACCTCTTT	TTATTGG	AGAATACAT
4	22	21652...21701	AACCCTTGA	TTTTTTACC	GCACCGTGAAGGGTG	CAAAGAG	GCAGGAGAAA
5	27	24113...24153	AAGAATAGGG	TGTGACG	TCGACACCCACAAC	ACACAAGAG	GAAACACA
6	31	29193...29235	GGGGTGTGCA	CGTCGACA	CCCCCTCTTTACTGG	AGAGACT	TAA
7	32	30075...30118	C	AAGTAACG	AAGCAACTAGTAACGAAA	TAACTGT	CAATGGAGAC
8	34	32127...32178	CCAACTAGCC	TCTTGGGT	TGGTGCAGTTCTTCAAC	CTAACCC	AATGCAAGCA
9	36	33231...33280	TGTCGCCAC	CCTTTAGC	TATCCAACCAACCA	ATCGTAG	GAGCACGACG
10	40	35422...35472	CTGCAATAGA	GCTTGACG	CGCGCCCATATGTATGTT	ATAATAG	GTGTGTCG
11	42	36449...36500	TTGTCGCAAT	CTTTTATT	TGGGGGAGGGCGTGTCC	CTCCTGT	CTCAACTAGT
12	44	37139...37158	CAATGGTGTG	GACGGTGC	ACACACCCCGAACAC	GTAAGGA	GTCCGAA
13	46	38150...38200	TTGCAAGGGT	ACTTGACA	GGCATCCGAAAGTGTG	TTAAAAT	AAGAACACAA
14	47	38040...38092	CTTGCTTGCT	GACCACCA	CTCTGGTGTGTTAGTTCGT	AGCTGCT	TGTGCTGCTC
15	48	39031...38080	TCGGTCACAA	AATGTTTCG	AAATGTGACAAAAC	CACATCT	TTGACACACG

### 7.5.2 Módulo de empaquetamiento

En este módulo se identificaron siete ORFs, de éstos cuatro fueron anotados como proteínas hipotéticas. Los ORFs 2 y 3 fueron asignados como subunidad menor y subunidad mayor de la terminasa, respectivamente; mientras que al ORF 4 como proteína conectora de cápside a cola. El ORF2 mostró similitud con una proteína hipotética del fago VP5 (identidad= 94%) de *V. cholerae*. El ORF3, mostró similitud con una proteína terminasa del fago HTVC010P (identidad= 28%) de *Pelagibacter ubique*. El ORF4 mostró similitud con una proteína conectora de cápside a cola del fago

HTVC010P (identidad= 32%) de *Pelagibacter ubique*, por lo que se consideró que el ORF4 puede ser anotado con dicha función.

### **7.5.3 Modulo de morfogénesis de cápside y cola**

Corriente abajo del módulo de empaquetamiento se encuentra el módulo de morfogénesis de cápside y cola. Este módulo está constituido por catorce ORFs. Siete de ellos (9 – 13, 16 y 18) fueron anotados como proteínas hipotéticas. Cinco (14, 15, 19, 20 y 21) fueron identificados usando BLAST-X y anotados con una función definida (Proteína de cola, dos como proteínas estructurales, fibra de cola y proteína de cápside respectivamente) y dos (8, y 17) fueron identificados mediante espectrometría de masas como proteínas estructurales y proteína de cola. El ORF14 presentó similitud con la subunidad mayor de una proteína de cola del fago VP5 (identidad= 90%). En éste, además se identificó un dominio conservado de inmunoglobulina de la familia Ig\_2 (Pfam: 1 e-07). Las secuencias del ORF20 y 21 tuvieron similitud con fibras de cola del fago VP5 (identidad= 96%) y con una proteína de cápside del fago VP2 (identidad= 31%) respectivamente. En la secuencia del ORF21, se determinó la presencia del dominio conservado BNR/Asp-box con un motivo SxDxGxxW.

### **7.5.4 Módulo de metabolismo**

Adyacente al módulo de morfogénesis y en la cadena contraria se encuentra el módulo de metabolismo. Este módulo está formado por seis ORFs de los cuales cuatro (22, 23, 24 y 26) fueron anotados como proteínas hipotéticas. Mientras que el ORF25 presentó similitud con una hidrolasa del fago VP2 (identidad= 94 %) el cual, además, posee un dominio conservado HD-3 (Pfam: 9 e-11) con un motivo H-21x-HD, que

corresponde con una fosfohidrolasa metal-dependiente. El ORF27 posee un dominio conservado de adenocil succinato sintetasa (Pfam: 1.8 e-17), el cual tuvo similitud con una proteína de función similar del fago VP5 (identidad 99%).

### **7.5.5 Módulo de replicación**

Corriente arriba del módulo de metabolismo se encuentra el módulo de replicación, constituido por siete ORFs (28 – 34). Tres de ellos fueron anotados como proteínas hipotéticas (30, 32 y 33). Los ORFs 28, 29, 31 y 34 codifican para proteínas implicadas en funciones de replicación: una integrasa, una DNA polimerasa I, una proteína de unión a DNA de cadena sencilla y una helicasa de la superfamilia II, respectivamente. Los ORF28 y 31 mostraron similitud con una integrasa y con la proteína de unión a DNA de cadena sencilla (SSBP) del fago VP2 (identidad= 99% y 93%, respectivamente). En el análisis de la secuencia del ORF28, se identificó la presencia de un dominio de primasa/polimerasa bifuncional amino terminal (N-Ter prim/pol) además de un dominio de primasa carboxilo terminal-2 (PriCT-2). Con respecto al ORF29 éste presentó similitud con una DNA polimerasa I del fago ØJL001 de *Proteobacteria* (identidad= 26%) perteneciente a la familia *Siphoviridae*. El ORF34 tuvo similitud con una DNA/RNA helicasa de la superfamilia II del fago THSA-485A de *Thermoanaerobacterium saccharolyticum*, (identidad= 36%), perteneciente a la familia *Siphoviridae*. Adicionalmente, este ORF presenta una región conservada con el dominio SNF-2 (Pfam: 1 e-24) encontrado en diversas DNA/RNA helicasas ATP-dependientes.

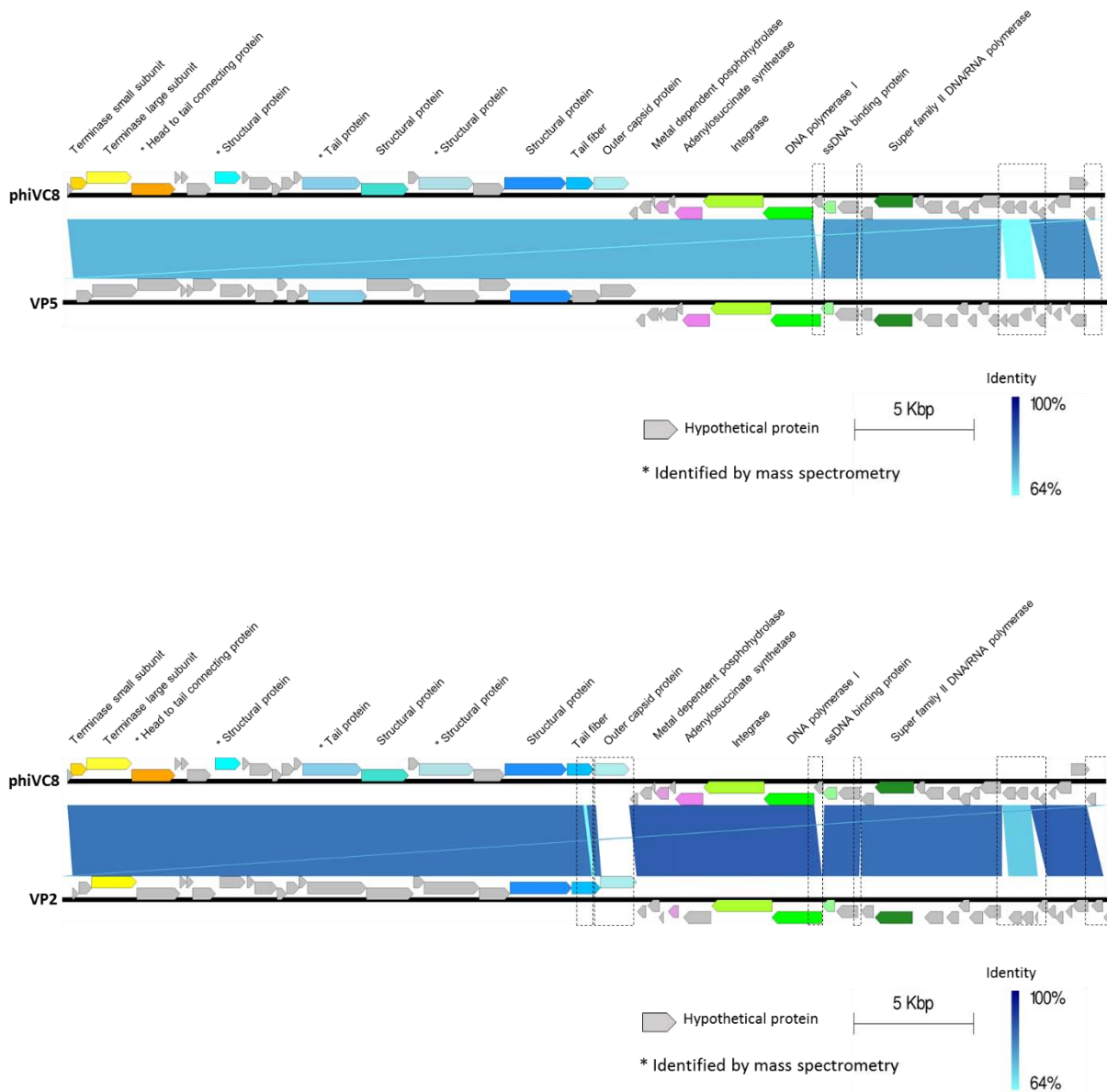
## **7.6 Análisis comparativo del genoma de ØVC8 con los de VP2 y VP5**

En los resultados de alineación con BLAST del genoma ØVC8 revelaron que posee un 86% y 85% de identidad con los genomas de los fagos VP2 y VP5, respectivamente. Las diferencias a nivel de secuencia de DNA se encuentran en seis regiones del genoma ØVC8 en comparación con el genoma VP2 y en cuatro regiones con el genoma VP5. Consistentemente, las proteínas putativas de seis ORFs localizados en estas regiones de la secuencia del genoma de ØVC8, muestran similitudes con diferentes porcentajes de identidad y posiciones a las proteínas de los fagos VP2 y VP5. Interesantemente, la probable proteína codificada por el ORF 30 no está presente en los genomas de VP2 y VP5 (Figura 15).

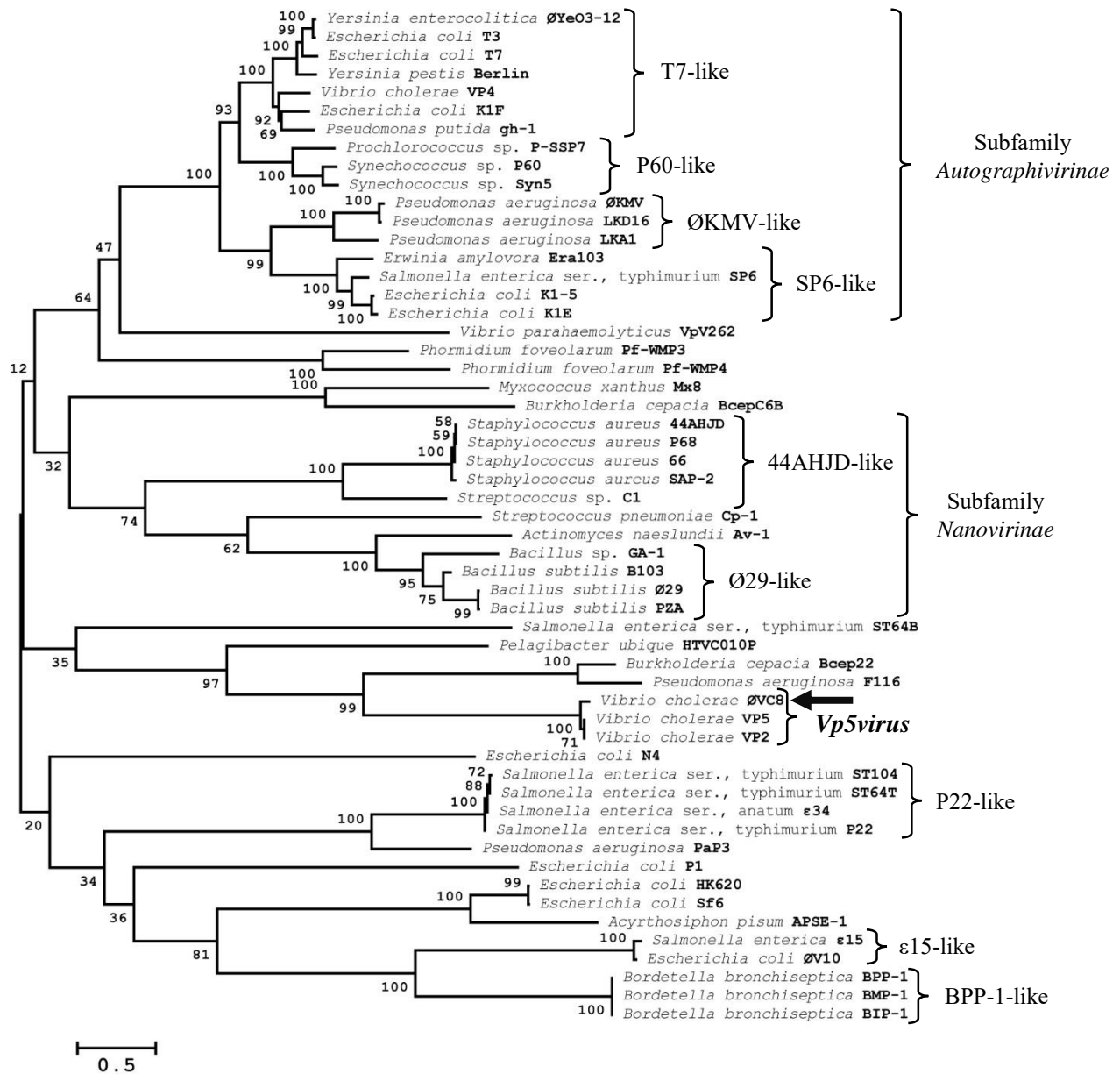
## **7.7 Relación filogenética de ØVC8 con fagos de la familia *Podoviridae***

El análisis morfológico y de secuencia de DNA, reveló que ØVC8 pertenece a la familia *Podoviridae*, por lo que, para definir su relación filogenética con otros miembros de ésta familia, se analizó la secuencia traducida del ORF3 que codifica para la subunidad mayor de la terminasa contra las secuencias de 55 fagos de la familia *Podoviridae*. El análisis mostró que ØVC8 se encuentra junto con los fagos VP2 y VP5 formando un grupo estrechamente relacionado (Figura 16).





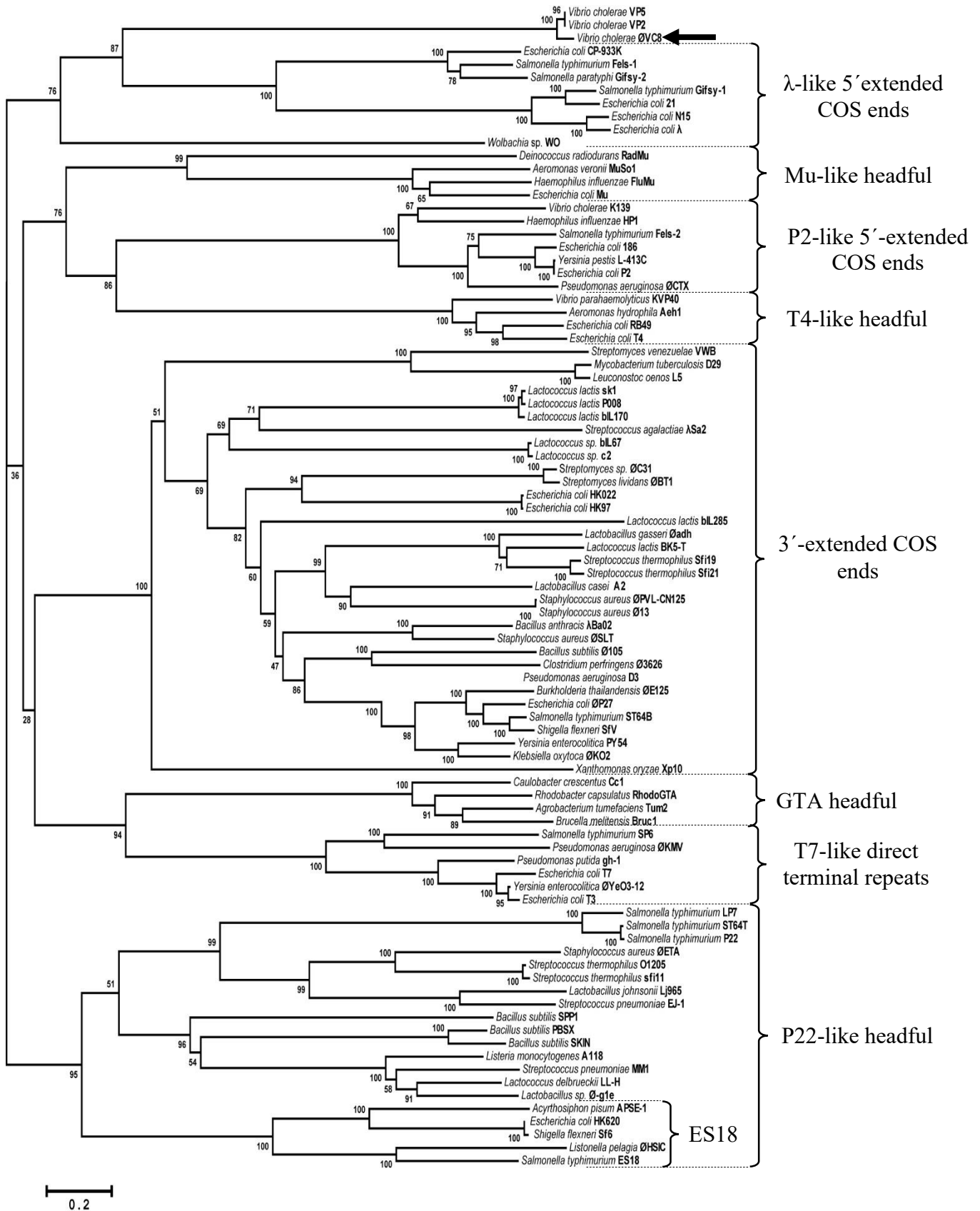
**Figura 15.** Comparación del genoma completo de ØVC8 contra los genomas de VP5 (arriba) y VP2 (abajo) utilizando el programa EasyFig ver. 2.1. La posición de cada ORF y la dirección de transcripción se representan como flechas de bloque. La función de cada ORF se muestra en color y en gris las proteínas hipótéticas. El grado de similitud de secuencia se indica mediante la intensidad del color, que indica los niveles de identidad de nucleótidos (de 64 a 100%). Las comparaciones se realizaron utilizando BLASTN. Las inserciones/delecciones así como los niveles de identidad menor que 64% se indican en recuadrados punteados.



**Figura 16.** Árbol Neighbor-Joining para comparar la secuencia de aminoácidos del ORF3 (subunidad mayor de la terminasa) de ØVC8 y de otros 54 fagos obtenidas de GenBank, que representan los principales grupos (marcados en llaves) de la familia *Podoviridae*. Al final de cada rama se encuentra el hospedero seguido del nombre del fago marcado en negritas. La flecha señala la posición de ØVC8. Los números en los nodos internos son los valores de bootstrap (0-100%) obtenidos de 1, 000 réplicas de bootstrap.

## 7.8 Predicción de la estrategia de empaquetamiento del DNA de ØVC8

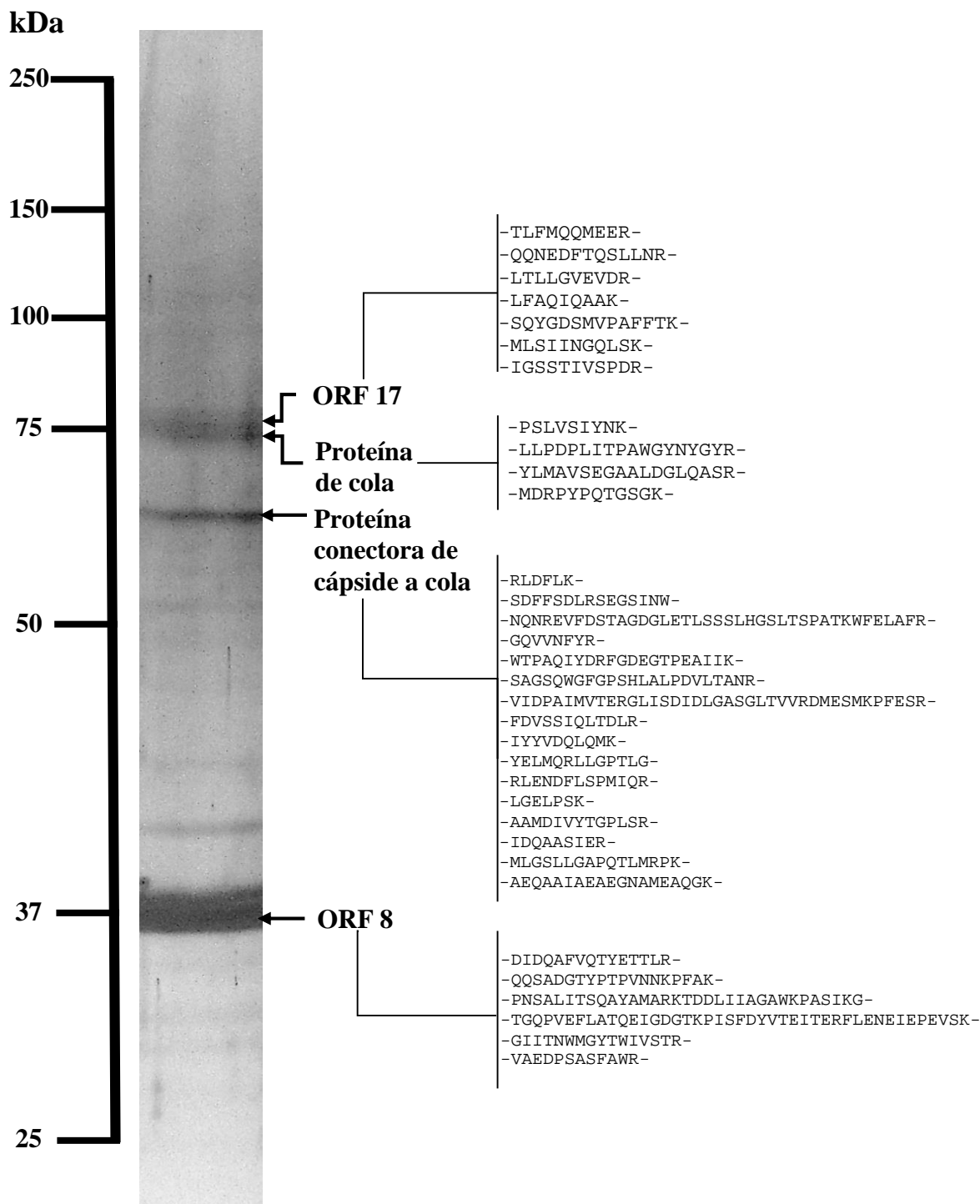
Una característica importante entre los miembros del orden *Caudovirales* (fagos con cola), es que todos poseen una copia de DNA lineal de doble cadena, la cual, durante el proceso de ensamblaje del fago es traslocado a la procápside viral. Este proceso compacta el DNA cargado negativamente hasta una densidad similar a la de un cristal de DNA, sin embargo, este proceso no es espontáneo y se lleva a cabo mediante un evento de translocación ATP dependiente, mediado por la proteína terminasa del fago. Por otra parte, la estrategia de replicación y empaquetamiento de DNA entre los miembros del orden *Caudovirales* no es la misma en todos los casos. Para determinar cuál es la posible estrategia de empaquetamiento del DNA que utiliza ØVC8, se comparó la secuencia de aminoácidos de la subunidad mayor de la terminasa contra la de otros 88 fagos que presentan las principales estrategias de empaquetamiento de DNA entre los *Caudovirales*. Los resultados mostraron que la terminasa de ØVC8 está relacionada con la de un grupo de fagos que utiliza una estrategia de empaquetamiento de tipo extremos cohesivos 5' descrito para fagos del grupo lambda-like, entre los que se encuentran fagos de *E. coli* y *Salmonella* spp., (Figura 17).



**Figura 17.** (Página anterior) Árbol Neighbor-Joining para comparar la secuencia de aminoácidos de la subunidad mayor de la Terminasa del ORF3 del ØVC8 y de otros 88 fagos pertenecientes al orden Caudovirales obtenidas de GenBank. La flecha señala la posición de ØVC8. Los números en los nodos internos son los valores de bootstrap (0-100%) obtenidos de 1, 000 réplicas. Los nombres de los fagos y profagos se muestran a la derecha de cada rama terminal (algunos nombres de profagos son los propuestos por Casjens (2011)). Los principales grupos de terminasas se encuentran marcados con llaves y líneas punteadas y la estrategia de empaquetamiento de DNA y el fago tipo para cada grupo se encuentra señalada a la derecha de cada llave.

## 7.9 Análisis proteómico

Proteínas estructurales de ØVC8 se obtuvieron y purificaron para finalmente hacer su análisis por espectrometría de masas. En la electroforesis en gel de poliacrilamida (Figura 18), se identificaron cuatro fragmentos con pesos moleculares de ~37, 60 y 75 kDa, en el gel se identificó que las fracciones de 37 y 75 kDa estaban conformadas por más de una banda con pesos moleculares muy cercanos. Al realizar el análisis por espectrometría de masas de estos tres fragmentos, se obtuvo la secuencia peptídica de cuatro proteínas. El análisis comparativo de las secuencias de éstas proteínas contra las secuencias traducidas de los ORFs identificados en el genoma de ØVC8, mostró que la de 61.7 kDa correspondía al ORF4 (proteína conectora de cápside a cola), la de 74.1 kDa al ORF14 (proteína de cola) y a los ORFs 8 y 17 las proteínas de 36.1 kDa y 76.8 kDa, respectivamente, ambas con identidad con proteínas hipotéticas. El análisis con BLAST-P reveló que las secuencias referidas se encuentran en proteínas de los fagos VP2 y VP5 con funciones similares.



**Figura 18.** Proteínas estructurales de ØVC8. A) Patrón electroforético de las proteínas estructurales de ØVC8 en gel de poliacrilamida-SDS al 12 %. B) Secuencia N-terminal de los péptidos obtenidos por espectrometría de masas y el ORF/función al cual está asociada.

# 8. Discusión

Los fagos son las entidades biológicas más abundantes y ampliamente distribuidas en el mundo. A la fecha se han descrito más de 6, 300 fagos y se ha obtenido la secuencia de aproximadamente 2, 000. Esto representa tan solo una pequeña fracción del total de fagos que se ha calculado que pueden estar presentes en la naturaleza  $\sim 1 \times 10^{30}$ , esto aunado al escaso conocimiento que se tiene de los fagos. En las últimas décadas se ha ido definiendo la importancia que los fagos tienen sobre el control y evolución de las comunidades bacterianas. Un claro ejemplo es el de *V. cholerae*, al respecto se ha reportado que la variación en las poblaciones de fagos influye en el comportamiento estacional de la bacteria y con el resurgimiento de nuevos brotes de cólera. En regiones en las que el cólera es endémico, cuando la población de fagos específicos para alguno de los serogrupos toxigénicos de *V. cholerae* es alta, la incidencia del padecimiento disminuye; por el contrario, cuando disminuyen dichas poblaciones, se incrementan los casos de cólera (Faruque *et al.*, 2005). Por otro lado, los fagos representan el mayor reservorio de secuencias genéticas nuevas, mismas que puede transferir a la bacteria favoreciendo cambios en su fenotipo.

El presente trabajo se realizó en una de las zonas que mayormente se vieron afectadas por el cólera, durante la séptima pandemia que afectó a nuestro país en 1991. Para la colecta de muestras de agua se seleccionó la Presa Endhó en el estado de Hidalgo, ya que dicho reservorio colecta las aguas residuales de la Ciudad de México, y de algunos municipios del Estado de México e Hidalgo. A partir de las muestras de agua se obtuvieron trece aislamientos de bacterias que fueron identificados como *Vibrio cholerae* No-O1/No-O139, *Vibrio alginolyticus* y *Aeromonas veronii*. Aunque, no se aisló ninguno de los serogrupos *V. cholerae* (O1/O139) implicados en la patogénesis del cólera, la presencia de cepas *V. cholerae* No-O1/No-



O139 da elementos para considerar que la bacteria es capaz de sobrevivir durante periodos prolongados en aguas residuales. En zonas endémicas de cólera, la presencia de cepas No-O1/No-O139 en el medio ambiente se ha relacionado con su posible participación como reservorios de fagos con actividad lítica contra *V. cholerae* O1/O139 (Faruque *et al.*, 2005). Estudios previos (Madico *et al.*, 1996), señalan la existencia de un número abundante de fagos en los sistemas de tratamiento de aguas residuales, aunque, al momento existe poca información sobre la dinámica poblacional y su interacción con la comunidad microbiana. En el trabajo desarrollado se aislaron 18 fagos que fueron específicos de *V. cholerae*, sin embargo, mostraron diversos patrones de infección definido esto por su capacidad para formar placas opacas o claras en las cepas utilizadas como receptoras. El 94% (17/18) de los fagos indujeron en diferentes cepas de *V. cholerae*, la formación de placas claras u opacas (Tabla 4), sin embargo, uno de los fagos que se designó como ØVC8 solo formó placas claras en cepas de *V. cholerae* O1. Actualmente, solo existen estudios sobre el impacto de los fagos sobre la aparición cíclica del cólera en áreas endémicas (Faruque *et al.*, 2005; Jensen *et al.*, 2006), por lo que el haber detectado la existencia de fagos líticos específicos para *V. cholerae* O1 en este estudio, resulta de gran importancia ya que éstos pudieran estar influyendo en el comportamiento epidemiológico del cólera en México.

El análisis por microscopía electrónica de ØVC8, mostró que presenta características morfológicas que corresponden al orden *Caudovirales*, mismo que representa el grupo más abundante de virus bacterianos (Ackermann 1998). De acuerdo con la clasificación de fagos de *Vibrio* propuesto por Ackermann y cols., (1984), ØVC8 corresponde al morfotipo C1 (fagos con cola corta o podovirus), en este

mismo grupo se incluyen los fagos I, III y 4996 de *Vibrio cholerae*, utilizados para la fagotipificación de la bacteria. Con relación a las dimensiones de la cápside y cola del fago (ØVC8), se identificó que son similares a las reportadas para los fagos I, III, 4996, S-2 y S20 que pertenecen a la familia *Podoviridae* (Mukerjee *et al.*, 1963; Chatterjee *et al.*, 1965; Ackermann *et al.*, 1983-1984; Chattopadhyay *et al.*, 1987; Chakrabarti *et al.*, 2000). Lo antes referido nos permite considerar que ØVC8 pertenece al orden *Caudovirales* en la familia *Podoviridae*. Interesantemente, algunos fagos de la familia *Podoviridae* como JSF3 y JSF6 de *V. cholerae* han sido aislados en áreas endémicas del cólera y están relacionados con la modulación de la enfermedad en la región del delta del río Ganges (Faruque *et al.*, 2005).

Con relación al análisis del genoma de ØVC8 se observó que al igual que los demás miembros del orden *Caudovirales*, este está compuesto por una cadena doble de DNA lineal con 39, 422 pb. En un ensayo previo a la secuencia completa del genoma, se realizó la secuencia y el análisis bioinformático de dos fragmentos de restricción (750 pb y 2 kb) del DNA del fago, obtenidos con la enzima *HpaII*. Los datos generados mostraron que ØVC8 tiene similitud con VP2 y VP5 fagos de *V. cholerae*, ambos pertenecientes al orden *Caudovirales* e incluidos originalmente en un subgrupo no clasificado de la familia *Podoviridae* denominada tentativamente como VP2-like por Lavigne y cols., (2008).

Estos tres fagos, ØVC8, VP2 y VP5 poseen diversas características en común como la ausencia de genes relacionados con la lisis bacteriana o genes putativos de lisis no reportados aun, la presencia de una adenilosuccinato sintetasa de la cual se han reportado homólogos en algunos fagos de *Mycobacterium* de la familia *Myoviridae* y en virus de Amoeba como megavirus de la familia *Mimiviridae* y de grupos no

clasificados pandoravirus, pithovirus y megavirus courdo7 (Xiao *et al.*, 2009; Philippe *et al.*, 2013; Legendre *et al.*, 2014). Estos resultados son la base para proponer un nuevo género de virus en la familia *Podoviridae* denominado *Vp5virus*.

La secuencia completa del genoma de ØVC8 mostró que está constituido por 48 marcos de lectura abierta (ORF) agrupados en 4 regiones con funciones relacionadas con el empaquetamiento, la morfogénesis de la cápside y la cola, el metabolismo y la replicación. Otro grupo de ORFs con función desconocida se localizan en una quinta región (río arriba del módulo de genes de replicación) (Figura 12 y tabla 6). La organización observada en el genoma de ØVC8 está de acuerdo con la teoría modular propuesta por Botstein en 1980, que refiere que las unidades de selección evolutiva en los fagos no son genes aislados sino módulos con una función específica que se intercambian por recombinación permitiendo la aparición de fagos con nuevas propiedades (Lima-Mendez *et al.*, 2011).

En el análisis bioinformático de los ORF de ØVC8 se identificó que algunos de éstos pueden poseer características adicionales descritas recientemente en otros fagos. En el módulo de empaquetamiento se identificaron siete ORFs, tres de ellos con funciones de la subunidad menor de la terminasa (ORF 2), de la subunidad mayor de la terminasa (ORF 3) y con la proteína conectora de cápside a cola (ORF4), todas funciones relacionadas con el empaquetamiento del DNA en la cápside del fago. Las terminasas son hetero-oligómeros compuestos por dos subunidades, menor y mayor, las cuales son responsables del empaquetamiento del DNA de doble cadena en la cápside del virión durante la etapa final de ensamblaje del fago (Jardine *et al.*, 2005; Rao & Feiss 2008). Estas enzimas han sido identificadas en todos los fagos con cola, y son de las proteínas mejor conservadas entre los miembros del orden *Caudovirales*

(Casjens 2003). La función de la subunidad menor es la de reconocer específicamente el sitio de inicio de empaquetamiento *cos* o *pac* (Roy *et al.*, 2012), y formar una estructura nucleoprotéica que ayuda a la subunidad mayor a cortar la cadena de DNA en un sitio específico (Gual *et al.*, 2000).

En un análisis filogenético de la subunidad mayor de la terminasa (ORF3) de ØVC8 contra 54 secuencias de fagos pertenecientes a la familia *Podoviridae*, se confirmaron los resultados obtenidos por análisis comparativo de las secuencias de restricción de los fragmentos de 750 pb y 2 kb obtenidos con la enzima *HpaI* del DNA del fago con respecto a la relación entre ØVC8 y los fagos VP2 y VP5 de *V. cholerae*. Con este resultado, además se puede predecir que ØVC8 posee una estrategia de empaquetamiento de tipo  $\lambda$ -like.

Durante la replicación de ØVC8, se requiere de la participación de una proteína terminasa que permita el empaquetamiento de su genoma y al mismo tiempo realice el corte de los concatémeros de DNA (Casjens *et al.*, 2011). En este sentido, la terminasa codificada por el ORF3 podría participar en este proceso. El análisis comparativo de dicha proteína, sugiere la participación de esta terminasa en el proceso de empaquetamiento a través de una estrategia de tipo "extremos cohesivos de cadena sencilla", descrito para fagos del tipo lambda. Los fagos con estas características tienen una secuencia que al ser cortada por la terminasa genera extremos de cadena sencilla complementarios llamados sitios cohesivos o COS; estos sitios son altamente conservados en el genoma y se localizan regularmente 1, 000 pb corriente arriba del gen que codifica para la subunidad menor de la terminasa (Casjens *et al.*, 2009). Los resultados anteriores nos llevaron a proponer que el sitio COS de

ØVC8 se encuentra en una secuencia repetida a 299 pb corriente arriba de la subunidad menor de la terminasa (Figura 13).

Por otro lado, corriente abajo se localiza el ORF4 que codifica para una proteína de unión de cápside a cola. La presencia de una conexión entre uno de los vértices de la procápside y la cola del fago maduro es fundamental para la translocación del genoma del fago a la cápside, ya que dicho conector forma un orificio axial a través del cual pasa la hebra de DNA del exterior hacia el interior de la procápside y viceversa (Carrascosa *et al.*, 1982). La presencia de la proteína de unión de cápside a cola (ORF4) en ØVC8, sugiere que ésta proteína participa en el proceso de empaquetamiento del cromosoma a la procápside durante la fase final de la replicación del fago y también de expulsarlo durante la fase infecciosa del mismo.

El módulo de morfogénesis de cápside y cola conformado por 14 ORFs tiene funciones relacionadas con la producción de componentes estructurales de la cápside y la cola de ØVC8. Por los resultados obtenidos en el análisis por BLAST-X se asignó una función a los ORFs 14, 20 y 21, mientras que por análisis proteómico se identificaron las proteínas estructurales correspondientes a los ORFs 8, 14 y 17. El ORF14 codifica para una proteína de cola de 74.1 kDa, la cual tiene un papel importante en todos los fagos con cola ya que es la responsable del reconocimiento de los receptores específicos en la célula hospedera (Auzat, *et al.*, 2008). En fagos como P22 se ha reportado que dicha proteína posee algunas funciones enzimáticas que contribuyen al corte específico de la ramnosil-1,3-galactosa presente en la pared celular de algunas bacterias (Sauer, *et al.*, 1982).

En el 2006, Fraser y cols., mediante análisis informático identificaron que, en las proteínas constituyentes de fagos de DNA de doble cadena con cola se encuentran

a menudo dominios Ig-like. Una de las primeras referencias que se tienen de dominios Ig-like corresponde a la proteína Hoc del fago T4 (Bateman 1997). Dichos dominios se dividen en tres familias, los que son similares a la inmunoglobulina de dominio clásico (I-Set), la fibronectina tipo 3 de repetición (FN3) y el dominio Ig-like bacteriano (Big2).

Resultados del análisis Pfam de la secuencia de aminoácidos del ORF14, muestran que tiene un dominio conservado de tipo I-set. A pesar de que este tipo de dominios han sido descritos en bacterias, se desconoce el origen de los mismos en secuencias de fagos. Al respecto, Fraser y cols., (2007) sugieren que los fagos pueden haber sido un vector importante en la propagación de dominios Ig-like a través de diversas especies de bacterias. Si bien se desconoce la función específica de los dominios Ig-like en los fagos, algunas hipótesis sugieren que pueden desempeñar un papel accesorio en la infección por fagos, por la interacción de dichos dominios con compuestos del tipo hidratos de carbono presentes en la pared celular de la bacteria. Por otra parte, en 2013, Barr y cols., demostraron que los dominios Ig-like son importantes para la unión de los fagos a la superficie de las mucosas en metazoos, tal situación plantea que la adherencia de los fagos a las mucosas pudiera participar en la captura de la bacteria y proporcionar una actividad inmune independiente de la propia del hospedero. La propuesta de la fagoterapia como un recurso para el tratamiento de las infecciones por bacterias, es un hecho que se ha planteado desde el descubrimiento de los fagos por Felix D'Herelle a principios del siglo XX. Sin embargo, en el caso de las enfermedades infecciosas del intestino uno de los principales inconvenientes hasta el momento, es la barrera del sistema inmune al interior del intestino. La presencia de dominios Ig-like en ØVC8, plantean que, si el fago es capaz de adherirse a la superficie de las mucosas en humanos este podría

considerarse como una alternativa en el control de *V. cholerae* en pacientes que presenten cólera.

El ORF20 localizado también en el módulo de morfogénesis, codifica para una proteína de fibra de cola. Sin embargo, en el análisis por microscopía electrónica de ØVC8 no mostró la presencia de este tipo de fibras, una explicación al respecto es que el fago presente una morfología de la cola similar a la de T7 que está constituida por homo-trímeros elongados compuestos de tres copias de la proteína gp17, las cuales están retraídas hacia la cápside y no son visibles hasta que sucede el reconocimiento irreversible de los receptores del hospedero (García-Doval *et al.*, 2012).

Junto a la fibra de la cola, corriente abajo se encuentra el ORF21, asociado a una proteína de cápside, de acuerdo con los resultados de BLAST-X, esta secuencia posee un dominio conservado BNR/Asp-box (neuranimidasa bacteriana o sialidasa). Las Asp-box se han reportado en más de nueve familias proteicas en las que se incluyen: ribonucleasas bacterianas, sulfito oxidasas, reelinas, netrinas, algunos receptores lipoprotéicos y una amplia variedad de glicosil hidrolasas (Copley *et al.*, 2001). Estas se caracterizan por poseer secuencias repetidas de un motivo de 8 a 12 residuos de los cuales 5 son altamente conservados (S-x-D-x-G-x-T-W) (Roggenting *et al.*, 1993), lo que sugiere que el producto del ORF21 de ØVC8 tiene actividad de sialidasa. La actividad de sialidasa en algunos fagos como 63D, se ha relacionado con su capacidad para degradar los exopolisacáridos de la cápsula en algunas bacterias (Machida *et al.*, 2000a-2000b).

El módulo de metabolismo está constituido por 6 ORFs (22 al 27), de éstos, dos ORF25 y ORF27, se relacionaron con funciones involucradas con el metabolismo de *V. cholerae*. El ORF25 codifica para una fosfohidrolasa metal-dependiente que

pertenece a la superfamilia HD cuyos miembros son metaloenzimas que poseen una combinación distintiva de residuos metal-quelantes principalmente de histidina y aspartato (HD) cuyos sustratos son moléculas que contienen enlaces fosfoéster (Aravind *et al.*, 1998). Los dominios HD poseen además actividad de nucleasa y fosfodiesterasa que participan de forma importante en el metabolismo de nucleótidos y en señalización (Zimmerman *et al.*, 2008). Por su parte el ORF27 codifica para una adenosil succinato sintetasa, enzima que participa en la biosíntesis de purinas mediante la catálisis GTP-dependiente de la inosina monofosfato (IMP) a adenosina monofosfato (AMP) (Wiesmüller *et al.*, 1991). Esta enzima ha sido descrita tanto en procariotas como en *Escherichia coli* (gen *purA*), como en tejidos de vertebrados (Stayton *et al.*, 1983). Típicamente, los fagos pueden contener múltiples genes relacionados con la biosíntesis de ácidos nucleicos y replicación, los cuales son procesos críticos para la replicación de su genoma (Miller *et al.*, 2003).

El último módulo funcional identificado fue el de replicación. Este está integrado por siete ORFs (28 al 34), de éstos el 28 (integrasa), 29 (DNA polimerasa I), 31 (SSB) y 34 (Helicasa), poseen funciones relacionadas con la replicación del DNA. El ORF28 tiene función de integrasa, este tipo de enzimas tienen función de recombinasas sitio-específicas (corte y empalme) entre sitios de reconocimiento de entre 30 a 40pb entre los genomas del fago y su hospedero. Las integrasas han sido identificadas tanto en procariotas como eucariotas y no poseen relaciones evolutivas con las integrasas retrovirales. Por la secuencia de aminoácidos, así como la presencia de residuos catalíticos, las integrasas de fagos se clasifican en dos familias: integrasas de tirosina e integrasas de serina (Groth *et al.*, 2004). Las de tirosina se caracterizan por poseer un residuo de tirosina (T342 en el fago  $\lambda$ ), además de otros tres residuos que



corresponden a  $\lambda$  R212, H308 y R311 que comprenden la triada clásica en las integrasas de tirosina, características que no fueron encontradas en el ORF28 de ØVC8 ni en ninguna de las secuencias más cercanas de acuerdo con los resultados de BLAST. En el ORF 28 se identificaron además dos dominios que corresponden a una primasa/polimerasa bifuncional amino terminal (N-Ter prim/pol) y a una primasa carboxilo terminal 2 (PriCT-2). En un análisis con Pfam, se identificó que tanto en ØVC8 como en VP2 y VP5, poseen los mismos dominios conservados de primasa/polimerasa bifuncional.

Las N-Ter prim/pol son un grupo de enzimas originalmente descritas para la bacteria termoacidófila *Sulfolobulus islandicus* y codificada por el plásmido pRN1 en el ORF904. La N-Ter prim/pol es un dominio enzimático multifuncional con actividad ATPasa, primasa, DNA polimerasa y helicasa en la misma región N- terminal de la proteína (Lipps *et al.*, 2004). A pesar de que los dominios N-terminal han sido identificados en diversos genomas de fagos (Weigel *et al.*, 2006), hasta el momento solo tres proteínas han sido estudiadas a detalle en los últimos años, el ORF904 de *S. islandicus*, RepA de *Sulfolobus solfataricus*, ambas codificadas por plásmidos de archeas y gp43 del fago BFK20 de *Brevibacterium flavum* (Halgasova *et al.*, 2012). Una de las principales características del dominio prim/pol es que en ambos casos utiliza preferentemente deoxiribonucleotidos como sustratos (Lipps *et al.*, 2004). Por su parte, PriCT-2 es un dominio que pertenece a la superfamilia de las primasas archaeo-eucariotica, al clado de las prim/pol-like y se encuentra ampliamente distribuido entre bacterias, fagos, plásmidos y algunas archeas (Iyer *et al.*, 2005).

El análisis bioinformático ORF29 que codifica para una DNA-Polimerasa, sugiere que la polimerasa de ØVC8 pertenece a la familia A, también conocida como

polimerasa I. La familia de la DNA polimerasa I, se encuentra principalmente en organismos relacionados con procariotas e incluye la DNA polimerasa I procariota, la polimerasa mitocondrial “g” y varias polimerasas de bacteriófagos, incluyendo las de los fagos T3 (*Mioviridae*), T5 (*Siphoviridae*) y T7 (*Podoviridae*) (Franklin *et al.*, 2001; Patel *et al.*, 2001).

El ORF31 codifica para una proteína de unión a cadena sencilla (SSBP). Las proteínas SSB se encuentran ampliamente distribuidas desde virus hasta eucariotas, no poseen actividad enzimática y su papel es la de cubrir y proteger los intermediarios de cadena sencilla de DNA (ssDNA) que se forman durante los eventos de replicación, reparación, transcripción y recombinación. Adicionalmente, las SSBs de organismos filogenéticamente alejados, poseen poca homología a nivel de secuencia de aminoácidos, sin embargo, estas proteínas han sido tipificadas a partir de una estructura conservada denominada plegamiento de unión oligonucleótido/oligosacárido (OB) la cual está involucrada en la unión a DNA (Bochkarev *et al.*, 2004). Por su parte, diversas proteínas de unión a DNA de cadena sencilla (SSBP) codificadas por fagos han sido identificadas basados en la similitud de secuencias de aminoácidos con las SSB de *E. coli*, mientras que otras han sido identificadas con base en su actividad (Szczepanska *et al.*, 2007). Con respecto a ØVC8 la función de la proteína se definió por la similitud que mostró en la base de datos con la reportada para el fago ØEcoM-GJ1 de *E. coli*.

Con relación al ORF34 éste codifica para una helicasa que presenta un dominio N-terminal de la familia SNF-2. Éste tipo de proteínas intervienen en procesos celulares como son: regulación transcripcional, estabilidad del cromosoma y en aspectos relacionados con daño celular, incluyendo reparación, recombinación y

reparación post-replicativa (Eisen *et al.*, 1995). Las proteínas con dominios SNF-2 son una familia helicasa-like ATP dependientes involucradas en la remodelación de la estructura de la cromatina (Hall *et al.*, 1999; Ryan *et al.*, 2011).

Para completar el ciclo de vida de cualquier fago, el siguiente paso importante es la liberación de la progenie viral. En este sentido, los fagos con cola han desarrollado diversas estrategias que permiten la lisis del hospedero. De manera interesante, el análisis bioinformático de ØVC8 no mostró la presencia de secuencias relacionadas con holinas ni lisinas, proteínas con actividad enzimática relacionadas con la liberación del fago durante la fase final del ciclo lítico. En fagos que infectan bacterias Gram-negativas, tanto las endolisinas como las holinas son proteínas relativamente pequeñas, en el caso de las endolisinas, éstas presentan un solo dominio responsable del corte en los enlaces del péptidoglicano de la pared celular bacteriana. Por su parte, las holinas son mucho más diversas y frecuentemente poseen una secuencia única de aminoácidos. Sin embargo, en estudios recientes se ha mostrado que los fagos también pueden explotar la maquinaria de secreción de la célula huésped para que libere las endolisinas a su objetivo además de activar los mecanismos de auto-lisis de la bacteria, de tal forma en que los fagos puedan ser liberados (Catalão *et al.*, 2013). Los datos obtenidos en el estudio, señalan que ØVC8 es un fago virulento con actividad lítica contra varias cepas de *V. cholerae* O1, sin embargo, ninguno de los genes identificados en el genoma de ØVC8, parecen estar implicados en la lisis bacteriana.

El análisis estructural de ØVC8 permitió identificar 4 proteínas cuyas secuencias corresponden a los ORFs 4, 8, 14 y 17. El análisis de dichas proteínas reveló que estas se distribuyen entre los módulos de morfogénesis y de estructura, de tal forma que las

proteínas codificadas por los ORFs 4, 8, y 17 participan en el ensamblaje del profago y la proteína codificada por el ORF 14 en el reconocimiento del hospedero. Adicionalmente, la proteína de 36,1 kDa cuya secuencia se ubica en el ORF8 es una de las proteínas estructurales más abundantes, lo que sugiere que se trata de una proteína con un alto número de copias, por lo que presumiblemente sea la principal proteína de la cápside.

No obstante que la selección del área de estudio para la colecta de las muestras de agua, se apoyó en que en dicha región (Hidalgo, Méx.) correspondió a una de las más afectadas por el cólera durante la emergencia de la séptima pandemia en nuestro país, no se logró el aislamiento de *V. cholerae* O1. Lo anterior plantea una serie de preguntas referentes a los factores físicos y/o biológicos que influyen en la supervivencia y abundancia de las cepas de *V. cholerae* responsables del cólera. Sin embargo, la presencia de *V. cholerae* No-O1/No-O139 tanto en estudios previos (Isaac-Márquez et al., 1998; Solís-Sánchez 2005) como en el presente, muestra que la bacteria se encuentra en lugares como la presa Endhó donde existen las condiciones adecuadas para su supervivencia. La pregunta de por qué el cólera no tiene un carácter endémico en nuestro país y la razón de su aparición tras largos periodos (hasta 100 años), aún no tiene una respuesta clara.

Sin embargo, la existencia de fagos específicos para *V. cholerae* O1 y O139 es un fuerte indicativo de la presencia de dichos serogrupos en el ambiente. El motivo de que no causen la enfermedad puede estar relacionado con la pérdida del profago CTXØ y de manera consecuente de la capacidad para producir la toxina responsable del cuadro clínico del cólera. En los últimos años, los fagos han contribuido en gran medida a la comprensión de la epidemiología del cólera (Brüssow *et al.*, 2004),

sirviendo como vectores del principal factor de patogenicidad de la bacteria y probablemente modulando las poblaciones de la bacteria en el hábitat acuático y por consecuencia la incidencia de casos de cólera (Faruque *et al.*, 2005; Zahid *et al.*, 2008; Seed *et al.*, 2011). En este contexto, fagos líticos específico de *V. cholerae* O1 como pudiera ser ØVC8, pueden ser un factor clave en la regulación de este serogrupo en México y por lo tanto en surgimiento de los brotes epidémicos.

# 9. Conclusiones

- Los resultados obtenidos en el presente estudio muestran la presencia de bacteriófagos específicos para *V. cholerae* de los serogrupos O1 y O139, los cuales son los primeros que se aíslan e identifican en una región no endémica del cólera.
- El aislamiento de cepas de *V. cholerae* No-O1/No-O139 demuestran que existen lugares donde hay condiciones adecuadas para la supervivencia de la bacteria en México.
- La presencia de fagos específicos para *V. cholerae* O1, son un indicativo de la presencia de ésta bacteria en México.
- ØVC8 es un fago lítico específico para *V. cholerae* O1, el primero que se ha caracterizado en México, pertenece a la familia *Podoviridae* y por sus características forma parte de un nuevo género propuesto denominado *Vp5-virus*.

# 10. Perspectivas



México es considerado como un área no endémica del cólera y desde el fin de la última pandemia en América Latina en 2001 y hasta el 2010, no se habían reportado nuevos casos de cólera. Durante éste periodo interepidémico se planteó la hipótesis de que la bacteria desaparece, sin embargo, diversos estudios incluyendo éste, muestran que en el país existen condiciones adecuadas para la persistencia de la bacteria en el ambiente. Ante tal hecho, se hace necesario mantener un monitoreo continuo de *V. cholerae* para establecer la presencia de la bacteria principalmente del serogrupo O1, empleando un marcador que facilite su identificación sin recurrir al aislamiento de la bacteria. En este sentido, los fagos aislados pueden ser utilizados para el monitoreo de *V. cholerae* O1 en el ambiente. Al mismo tiempo, es necesario realizar estudios que permitan conocer la relación ecológica entre estos fagos y *V. cholerae* y si participan como factor regulador de la bacteria como sucede en áreas endémicas del cólera.

Aunque en nuestro trabajo se utilizó un gran número de cepas para la caracterización de los fagos, consideramos que es necesario ampliar el número de bacterias a probar en el ensayo de perfil de infección, principalmente todas aquellas que se han aislado en nuestro país.

En el presente estudio se muestra la caracterización del genoma de ØVC8 el cual aporta nuevos datos en el estudio de los fagos de *V. cholerae*, sin embargo, resulta de interés comprobar experimentalmente la función de los ORFs encontrados en su genoma. Otro aspecto relevante de los fagos es su posible empleo como agentes terapéuticos, una vez que los fagos fueron identificados se consideró su uso en el control de las enfermedades, en los últimos años ha resurgido dicha propuesta. El cólera es una enfermedad que no ha podido ser controlada y a pesar de las medidas mundiales para controlar la enfermedad, el número de casos ha ido en aumento en los

últimos 10 años. Ante tal situación, el estudio de los fagos permitirá evaluar la factibilidad de su empleo terapéutico. En el caso específico de ØVC8 la presencia de un dominio de Ig que puede conferir la capacidad de adherirse a diferentes sustratos de mucosas (incluyendo el intestino humano), plantea la posibilidad de utilizar a ØVC8 para el tratamiento del cólera en fagoterapia, sin embargo, es necesario realizar los estudios correspondientes para determinar esta posibilidad. Por otro lado, es posible que fagos como ØVC8 estén participando como un control biológico natural de la bacteria, esta sería otra línea a investigar para el control del cólera.

## 11. Bibliografía

- Abedon ST, LeJeune JT.** 2007. Why bacteriophage encode exotoxins and other virulence factors. *Evol. Bioinform. Online.* 1:97–110.
- Ackermann HW, Furniss AL, Kasatiya SS, Lee JV, Mbiguino A, Newman FS, Takeya K, Vieu JF.** 1983. Morphology of *Vibrio cholerae* typing phages. *Ann. Inst. Pasteur Virol.* 134:387-397, 399-404.
- Ackermann HW, Kasatiya SS, Kawata T, Koga T, Lee JV, Mbiguino A, Newman FS, Vieu JF, Zachary A.** 1984. Classification of *Vibrio* bacteriophages. *Intervirology* 22:61-71.
- Ackermann HW.** 1998. Tailed bacteriophages: the order caudovirales. *Adv. Virus. Res.* 51:135-201.
- Ackermann HW.** 2007. 5500 Phages examined in the electron microscope. *Arch. Virol.* 152: 227–243.
- Ackermann HW.** 2011a. Bacteriophage taxonomy. *Microbiol. Australia.* 32: 90-94.
- Ackermann HW.** 2011b. The first phage electron micrographs. *Bacteriophage* 1:225-227.
- Ackermann HW.** 2012. Bacteriophage electron microscopy. *Adv. Virus Res.* 82: 1-32.
- Alam M, Sultana M, Nair GB, Siddique AK, Hasan NA, Sack RB, Sack DA, Ahmed KU, Siddique A, Watanabe H, Grim CJ, Huq A, Colwell RR.** 2007. Viable but nonculturable *Vibrio cholerae* O1 in biofilms in the aquatic environment and their role in cholera transmission. *Proc. Natl. Acad. Sci. USA.* 104:17801-17806.
- Anderson ES, Ward LR, Saxe MJ, de Sa JD.** 1977. Bacteriophage-typing designations of *Salmonella typhimurium*. *J. Hyg. (Lond).* 78:297-300.
- Andrews JR, Basu S.** 2011. Transmission dynamics and control of cholera in Haiti: an epidemic model. *Lancet.* 377:1248-1255.
- Aravind L, Koonin EV.**1998. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.* 23:469-72.
- Auzat I, Dröge A, Weise F, Lurz R, Tavares P.** 2008. Origin and function of the two major tail proteins of bacteriophage SPP1. *Mol. Microbiol.* 70:557-569.
- Barr JJ, Auro R, Furlan M, Whiteson KL, Erb ML, Pogliano J, Stotland A, Wolkowicz R, Cutting AS, Doran KS, Salamon P, Youle M, Rohwer F.** 2013. Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc. Natl. Acad. Sci. USA.* 110:10771-10776.
- Basu S, Mukerjee S.** 1968. Bacteriophage typing of *Vibrio* El Tor. *Experientia* 24:299-300.
- Bateman A, Eddy SR, Mesyanzhinov VV.** 1997. A member of the immunoglobulin superfamily in bacteriophage T4. *Virus Genes* 14:163-165.
- Bik EM, Bunschoten AE, Gouw RD, Mooi FR.** 1995. Genesis of the novel epidemic *Vibrio cholerae* O139 strain: evidence for horizontal transfer of genes involved in polysaccharide synthesis. *EMBO J.* 14:209-216.
- Bochkarev A, Bochkareva E.** 2004. From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr. Opin. Struct. Biol.* 14:36-42.

- Borroto RJ, Martinez-Piedra R.** 2000. Geographical patterns of cholera in México 1991-1996. *Int. J. Epidemiol.* 29:764-772.
- Borroto RJ.** 1997. Ecology of *Vibrio cholerae* serogroup O1 in aquatic environments. *Rev. Panam. Salud Publica* 1:3-8.
- Borroto RJ.** 1998. Supervivencia de *Vibrio cholerae* O1 en agua dulce superficial y cólera endémico: una hipótesis geoecológica. *Rev. Panam. Salud Pública*, 4:371-374.
- Botstein D.** 1980. A theory of modular evolution for bacteriophages. *Ann. NY Acad. Sci.* 354:484-490.
- Boulanger P.** 2009. Purification of bacteriophages and SDS-PAGE analysis of phage structural proteins from ghost particles. *Methods Mol. Biol.* 502:227-238.
- Boyd EF, Brüssow H.** 2002. Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. *Trends Microbiol.* 10:521-529.
- Brüssow H, Canchaya C, Hardt WD.** 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* 68:560–602.
- Brüssow H, Hendrix RW.** 2002. Phage genomics: small is beautiful. *Cell.* 108: 13–16.
- Burnet FM.** 1933. The classification of dysentery-coli bacteriophages. III. A correlation of the serological classification with certain biochemical tests. *J. Pathol. Bacteriol.* 37:179–184.
- Caprioli T, Zaccour F, Kasatiya SS.** 1975. Phage typing scheme for group D Streptococci isolated from human urogenital tract. *J. Clin. Microbiol.* 2:311-317.
- Carrascosa JL, Viñuela E, García N, Santisteban A.** 1982. Structure of the head-tail connector of bacteriophage phi 29. *J. Mol. Biol.* 154:311-24.
- Casjens SR, Gilcrease EB.** 2009. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. *Methods. Mol. Biol.* 502:91-111.
- Casjens SR.** 2003. Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.* 49:277-300.
- Casjens SR.** 2011 The DNA-packaging nanomotor of tailed bacteriophages. *Nat. Rev. Microbiol.* 9:647-657.
- Catalão MJ, Gil F, Moniz-Pereira J, São-José C, Pimentel M.** 2013 Diversity in bacterial lysis systems: bacteriophages show the way. *FEMS Microbiol. Rev.* 37:554-571.
- Cenens W, Makumi A, Mebrhatu MT, Lavigne R, Aertsen A.** 2013. Phage–host interactions during pseudolysogeny. *Bacteriophage.* 3:e25029.
- Chakrabarti AK, Ghosh AN, Nair GB, Niyogi SK, Bhattacharya SK, Sarkar B.** 2000. Development and evaluation of a phage typing scheme for *Vibrio cholerae* O139. *J. Clin. Microbiol.* 38:44-49.
- Chatterjee SN, Das J, Barua D.** 1965. Electron microscopy of cholera phages. *Ind. J. Med. Res.* 53:934-936.
- Chattopadhyay DJ, Sarkar BL, Ansari MQ, Chakrabarti BK, Roy MK, Ghosh AN, Pal SC.** 1993. New phage typing scheme for *Vibrio cholerae* O1 biotype El Tor strains. *J. Clin. Microbiol.* 31:1579-1585.

- Chattopadhyay S, Kinchington D, Ghosh RK.** 1987. Characterization of *Vibrio* El Tor typing phages: properties of the El tor Phage e4. *J. Gen. Virol.* 68:1411-1416.
- Chibani-Chennoufi S, Bruttin A, Dillmann ML, Brussow H.** 2004. Phage-host interaction: an ecological perspective. *J. Bacteriol.* 186:3677-3686.
- Chin CS, Sorenson J, Harris JB, Robins WP, Charles RC, Jean-Charles RR, Bullard J, Webster DR, Kasarskis A, Peluso P, Paxinos EE, Yamaichi Y, Calderwood SB, Mekalanos JJ, Schadt EE, Waldor MK.** 2011. The origin of the Haitian cholera outbreak strain. *N. Engl. J. Med.* 364:33-42.
- Clokier MRJ, Millard AD, Letarov AV, Heaphy S.** 2011. Phages in nature. *Bacteriophage* 1:31-45.
- Copley RR, Russell RB, Ponting CP.** 2001. Sialidase-like Asp-boxes: sequence-similar structures within different protein folds. *Protein Sci.* 10:285-292.
- DeWitt WE, Gangarosa EJ, Huq I, Zarifi A.** 1971. Holding media for the transport of *Vibrio cholerae* from field to laboratory. *Am. J. Trop. Med. Hyg.* 20:685-688.
- Dutka BJ, El Shaarawi A, Martins MT, Sanchez PS.** 1987. North and south american studies on the potential of coliphage as a water quality indicator. *Water Res.* 21:1127-1134.
- Eisen JA, Sweder KS, Hanawalt PC.** 1995. Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic. Acids. Res.* 23:2715-2723.
- Faruque SM, Islam MJ, Ahmad QS, Faruque ASG, Sack DA, Nair GB, Mekalanos JJ.** 2005. Self-limiting nature of seasonal cholera epidemics: Role of host-mediated amplification of phage. *Proc. Natl. Acad. Sci. USA.* 2005. 102:6119-6124.
- Franklin MC, Wang J, Steitz TA.** 2001. Structure of the replicating complex of a Pol alpha family DNA polymerase. *Cell.* 105:657-667.
- Fraser JS, Maxwell KL, Davidson AR.** 2007. Immunoglobulin-like domains on bacteriophage: weapons of modest damage?. *Curr. Opin. Microbiol.* 10:382-387.
- Fraser JS, Yu Z, Maxwell KL, Davidson AR.** 2006. Ig-like domains on bacteriophages: a tale of promiscuity and deceit. *J. Mol. Biol.* 359:496-507.
- García-Doval C, van Raaij MJ.** 2012. Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Proc. Natl. Acad. Sci. USA.* 109:9390-9395.
- Gotuzzo E, Cieza J, Estremadoyro L, Seas C.** 1994. Cholera, lessons from the epidemic in Peru. *Infect. Dis. Clin. N. Am.* 8:183-205.
- Groth AC, Calos MP.** 2004. Phage integrases: biology and applications. *J. Mol. Biol.* 335:667-678.
- Gual A, Camacho AG, Alonso JC.** 2000. Functional analysis of the terminase large subunit, G2P, of *Bacillus subtilis* bacteriophage SPP1. *J. Biol. Chem.* 275:35311-35319.
- Guttman B, Raya R, Kutter E.** 2005. Basic Phage Biology. In: Kutter, E., Sulakvelidze, A. (Eds) *Bacteriophages biology and applications*. Boca raton, CRC Press. Pp
- Hagens S, Loessner MJ.** 2014. Phages of *Listeria* offer novel tools for diagnostics and biocontrol. *Front. Microbiol.* 5:159

- Halgasova N, Mesarsova I, Bukovska G.** 2012. Identification of a bifunctional primase-polymerase domain of corynepophage BFK20 replication protein gp43. *Virus Res.* 163:454-460.
- Hall MC, Matson SW.** 1999. Helicase motifs: the engine that powers DNA unwinding. *Mol. Microbiol.* 34:867-877.
- Hendrix RW, Smith MC, Burns RN, Ford ME, Hatfull GF.** 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl. Acad. Sci. USA.* 96:2192-2197.
- Henn MR, Sullivan MS, Stange-Thomann N, Osburne MS, Berlin AM, Kelly L, Yandava C, Kodira C, Zeng Q, Weiland M, Sparrow T, Saif S, Giannoukos G, Young SK, Nusbaum C, Birren BW, Chisholm SW.** 2010. Analysis of high-throughput sequencing and annotation strategies for phage genomes. *PLoS ONE*, 5:1-12.
- Huq A, Sack RB, Nizam A, Longini IM, Nair GB, Ali A, Morris JG, Khan MNH, Siddique AK, Yunus M, Albert MJ, Sack DA, Colwell RR.** 2005. Critical factors influencing the occurrence of *Vibrio cholerae* in the environment of Bangladesh. *App. Environ. Microbiol.* 71:4645- 4654.
- Isaac-Márquez AP, Lezama-Dávila CM, Eslava-Campos C, Navarro-Ocaña A, Cravioto-Quintana A.** 1998. Serotypes of *Vibrio cholerae* non-O1 isolated from water supplies for human consumption in Campeche, Mexico and their antibiotic susceptibility pattern. *Mem. Inst. Oswaldo Cruz*, 93:17-22.
- Iyer LM, Koonin EV, Leipe DD, Aravind L.** 2005. Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. *Nucleic. Acids. Res.* 33:3875-3896.
- Jardine JJ, Anderson DW.** 2005. DNA packaging in dsDNA phages. In *The Bacteriophages*, 2nd edn. Calendar, R. (ed) New York: Oxford University Press pp 49-65.
- Jensen MA, Faruque SM, Mekalanos JJ, Levin BR.** 2006. Modeling the role of bacteriophage in the control of cholera outbreaks. *Proc. Natl. Acad. Sci. USA.* 103:4652-4657.
- Kaper JB, Morris JG, Levine MM.** 1995. Cholera. *Clin. Microbiol. Rev.* 8:48-86.
- Karaolis DK, Somara S, Maneval DRJR, Jonson JA, Karper JB.** 1999. A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* 399:375-379.
- Khakhria R, Duck D, Lior H.** 1990. Extended phage-typing scheme for *Escherichia coli* O157:H7. *Epidemiol. Infect.* 105:511–520.
- Khakhria R, Lior H.** 1992 Extended phage-typing scheme for *Campylobacter jejuni* and *Campylobacter coli*. *Epidemiol. Infect.* 108:403-414.
- Khemayan K, Pasharawipas T, Puiprom O, Sriurairatana S, Suthienkul O, Flegel TW.** 2006. Unstable lysogeny and pseudolysogeny in *Vibrio harveyi* siphovirus-like phage 1. *App. Env. Microbiol.* 71:1355–1363.
- Klucar L, Stano M, Hajduk M.** 2010. phiSITE: Database of gene regulation in bacteriophages. *Nucleic. Acids. Res.* 38:D366-D370.

- Kropinski AM, Mazzocco A, Waddell TE, Lingohr E, Johnson RP.** 2009. Enumeration of bacteriophages by double agar overlay plaque assay. In *Bacteriophages: Methods and Protocols*. Clokie, M. R. J. & Kropinski, A. M. (eds). New York: Humana Press, 501:69–76.
- Kruger DH, Schneck P, Gelderblom HR.** 2000. Helmut Ruska and the visualisation of viruses. *The Lancet*. 355: 1713-1717.
- Lavigne R, Seto D, Mahadevan P, Ackermann HW, Kropinski AM.** 2008. Unifying classical and molecular taxonomic classification: analysis of the *Podoviridae* using BLASTP-based tools. *Res. Microbiol.* 159:406-414.
- Legendre M, Bartoli J, Shmakova L, Jeudy S, Labadie K, Adrait, A, Lescot M, Poirot O, Bertaux L, Bruley C, Couté Y, Rivkina E, Abergel C, Claverie JM.** 2014. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proc. Natl. Acad. Sci. USA*. 111:4274-4279.
- Li M, Shimada T, Morris JG Jr, Sulakvelidze A, Sozhamannan S.** 2002. Evidence for the emergence of Non-O1 and Non-O139 *Vibrio cholerae* strains with pathogenic potential by exchange of O-antigen biosynthesis regions. *Infect. Immun.* 70:2441-2453.
- Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J.** 2010. *De novo* assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20:265-2672.
- Lima-Mendez G, Toussaint A, Leplae R.** 2011. A modular view of the bacteriophage genomic space: identification of host and lifestyle marker modules. *Res. Microbiol.* 162:737-746.
- Lipp EK, Huq A, Colwell RR.** 2002. Effects of global climate on infectious disease: the cholera model. *Clin. Microbiol. Rev.* 15:757-770.
- Lipps G, Weinzierl AO, von Scheven G, Buchen C, Cramer P.** 2004. Structure of a bifunctional DNA primase-polymerase. *Nat. Struct. Mol. Biol.* 11:157-162.
- Loessner MJ.** 2005. Bacteriophage endolysins – current state of research and applications. *Curr. Op. Microbiol.* 8:480-487.
- Lwoff A, Horner R, Tournier P.** 1962. A system of viruses. *Cold Spring Harb. Symp. Quant. Biol.* 27: 51-55.
- Machida Y, Hattori K, Miyake K, Kawase Y, Kawase M, Iijima S.** 2000a. Molecular cloning and characterization of a novel bacteriophage-associated sialidase. *J. Biosci. Bioeng.* 90:62-68.
- Machida Y, Miyake K, Hattori K, Yamamoto S, Kawase M, Iijima S.** 2000b. Structure and function of a novel coliphage-associated sialidase. *FEMS Microbiol. Lett.* 182:333-337.
- Madico G, Checkley W, Gilman RH, Bravo N, Cabrera L, Calderon M, Ceballos A.** 1996. Active surveillance for *Vibrio cholerae* O1 and vibriophages in sewage water as a potential tool to predict cholera outbreaks. *J. Clin. Microbiol.* 34:2968-2972.
- Mann NH.** 2005. The third age of phage. *PLoS Biol.* 3:e182.
- McLeod SM, Kimsey HH, Davis BM, Waldor MK.** 2005. CTXphi and *Vibrio cholerae*: exploring a newly recognized type of phage-host cell relationship. *Mol. Microbiol.* 57:347-356

- Miller ES, Kutter E, Mosig G, Arisaka F, Kunisawa T, Rüger W.** 2003. Bacteriophage T4 genome. *Microbiol. Mol. Biol. Rev.* 67:86-156.
- Miller JF.** 2003. Bacteriophage and the evolution of cholera. *Infect. Imm.* 71:2981-2982.
- Miller RV.** 2001. Environmental bacteriophage-host interactions: factors contribution to natural transduction. *Antonie Van Leeuwenhoek.* 79:141-147.
- Mitra A, Kesarwani AK, Pal D, Nagaraja V.** 2011. WebGeSTer DB-a transcription terminator database. *Nucleic. Acids. Res.* 19:1-7.
- Molineux IJ.** 2001. No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. *Molec. Microbiol.* 40:1-8.
- Mukerjee S, Phil D.** 1963. Bacteriophage typing of cholera. *Bull. World Health Organ.* 28:337-345.
- Mukhopadhyay AK, Takeda Y, Nair GB.** 2014. Cholera outbreaks in the El Tor biotype era and the impact of the new El Tor variants. *Curr. Top.Microbiol. Immunol.* 379:17-47
- Noguerola I, Blanch AR.** 2008. Identification of *Vibrio spp.* with a set of dichotomous keys. *J. Appl. Microbiol.* 105:175-185.
- Orata FD, Keim PS, Boucher Y.** 2014. The 2010 cholera outbreak in Haiti: how science solved a controversy. *PLoS Pathog.* 10:e1003967.
- Patel PH, Suzuki M, Adman E, Shinkai A, Loeb LA.** 2001. Prokaryotic DNA polymerase I: evolution, structure, and "base flipping" mechanism for nucleotide selection. *J. Mol. Biol.* 308:823-837.
- Paul JH.** 2002. Marine phage genomics. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* 133:463-476.
- Philippe N, Legendre M, Doutre G, Couté Y, Poirot O, Lescot M, Arslan D, Seltzer V, Bertaux L, Bruley C, Garin J. Claverie JM, Abergel C.** 2013. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science.* 341:281-286.
- Pollitzer, R.** 1954. Cholera studies 1. History of the disease. *Bull. World Health Organ.* 10:421-461.
- Rao VB, Feiss M.** 2008. The bacteriophage DNA packaging motor. *Annu. Rev. Genet.* 42:647-681.
- Rees CED, Loessner MJ.** 2005. Phage for the Detection of Pathogenic Bacteria. In: Kutter E and Sulakvelidze A (eds) *Bacteriophages: Biology and Applications.* CRC Press, Boca Raton, pp 267-284
- Roggentin P, Schauer R, Hoyer LL, Vimr ER.** 1993. The sialidase superfamily and its spread by horizontal gene transfer. *Mol. Microbiol.* 9:915-921.
- Rohwer F.** 2003. Global phage diversity. *Cell.* 113:141.
- Roy, A., Bhardwaj, A., Datta, P., Lander, G. C., Cingolani, G.** 2012. Small terminase couples viral DNA binding to genome-packaging ATPase activity. *Structure.* 20:1403-13.
- Ryan DP, Owen-Hughes T.** 2011. Snf2-family proteins: chromatin remodellers for any occasion. *Curr. Opin. Chem. Biol.* 15:649-656.
- Sack DA, Sack RB, Nair GB, Siddique AK.** 2004. Cholera. *Lancet.* 363:223-233.
- Sambrook J, Russell DW.** 2001. *Molecular cloning: a laboratory manual.* 3th edn. New York, USA, Cold Spring Harbor Laboratory Press. 1:1.116-1.118.



- Sauer RT, Krovatin W, Poteete AR, Berget PB.** 1982. Phage P22 tail protein: gene and amino acid sequence. *Biochemistry*. 21:5811-5815.
- Saunders JR, Allison H, James CE, McCarthy AJ, Sharp R.** 2001. Phage-mediated transfer of virulence genes. *J. Chem. Tech. Biotech.* 76:662-666.
- Schmelcher M., Loessner M. J.** 2008. Bacteriophage: Powerful Tools for the Detection of Bacterial Pathogens. In: Zourob, Mohammed, Elwary, Sauna, Turner, Anthony P. F. (Eds.) *Principles of bacterial detection: biosensors, recognition receptors and microsystems*. Springer, New York. pp. 731-754.
- Seed KD, Bodi KL, Kropinski AM, Ackermann HW, Calderwood SB, Qadri, F, Camilli A.** 2011. Evidence of a dominant lineage of *Vibrio cholerae*-specific lytic bacteriophages shed by cholera patients over a 10-year period in Dhaka, Bangladesh. *mBio*, 2:1-9.
- Sell TL, Schaberg DR, Fekety FR.** 1983. Bacteriophage and bacteriocin typing scheme for *Clostridium difficile*. *J. Clin. Microbiol.* 17:1148-1152.
- Sepulveda J. Gomez-Dantes H. Bronfman M.** 1992. Cholera in the Americas: an overview. *Infection* 20:243-248.
- Shattock, F.** 1955. The use of serology in the classification of micro-organisms. *J. Gen. Microbiol.* 12:367-374.
- Siddique AK, Cash R.** 2014. Cholera outbreaks in the classical biotype era. *Curr. Top. Microbiol. Immunol.* 379:1-16.
- Siebe C, Cifuentes E.** 1995. Environmental impact of wastewater irrigation in central Mexico: an overview. *Inter. J. Environ. Health. Res.* 5:161-173.
- Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Inan G.** 2009. Birol. ABySS: a parallel assembler for short read sequence data. *Genome Res.* 19:1117-1123.
- Smith GP.** 1988. Filamentous phage assembly: morphogenetically defective mutants that do not kill the host. *Virology.* 67:156-165.
- Solís-Sánchez GA.** 2005. Aislamiento y caracterización de *Vibrio cholerae* de los canales de Xochimilco. Tesis de Licenciatura. Facultad de Estudios Superiores Iztacala, UNAM. México.
- Sommer DD, Delcher AL, Salzberg SL, Pop M.** 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics*, 8:1-11.
- Stayton MM, Rudolph FB, Fromm HJ.** 1983. Regulation, genetics, and properties of adenylosuccinate synthetase: a review. *Curr. Top. Cell. Regul.* 22:103-141.
- Summers WC.** 2011. In the beginning. *Bacteriophage* 1:50-51.
- Suttle C. A.** 2005. Viruses in the sea. *Nature.* 437:356-361.
- Szczepanska AK, Bidnenko E, Płochocka D, McGovern S, Ehrlich SD, Bardowski J, Polard P, Chopin MC.** 2007. A distinct single-stranded DNA-binding protein encoded by the *Lactococcus lactis* bacteriophage bIL67. *Virology.* 363:104-112.
- Verma R.** 2012. Cholera vaccine: new preventive tool for endemic countries. *Hum. Vaccin. Immunother.* 8:682-684.

- Wagner PL, Waldor MK.** 2002. Bacteriophage control of bacterial virulence *Infect. Immun.* 70:3985-3993.
- Waldor MK, Mekalanos JJ.** 1994. Emergence of a new cholera pandemic: molecular analysis of virulence determinants in *Vibrio cholerae* O139 and development of a live vaccine prototype. *J. Infect. Dis.* 170:278-283.
- Waldor MK, Mekalanos JJ.** 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science.* 272:1910-1914.
- Wang Q, Kan B, Wang R.** 2013. Isolation and characterization of the new mosaic filamentous phage VFJΦ of *Vibrio cholerae*. *PLoS One.* 8:e70934.
- Ward LR, de Sa JD, Rowe B.** 1987. A phage-typing scheme for *Salmonella enteritidis*. *Epidemiol. Infect.* 99:291-294.
- Weigel C, Seitz H.** 2006. Bacteriophage replication modules. *FEMS Microbiol. Rev.* 30:321-381.
- Weinbauer, M.** 2004. Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* 28:127–181.
- Wiesmüller L, Wittbrodt J, Noegel AA, Schleicher M.** 1991. Purification and cDNA-derived sequence of adenylosuccinate synthetase from *Dictyostelium discoideum*. *J. Biol. Chem.* 266:2480-2485.
- Wilson G.** 1984. Cholera, p. 446–457. In G. Wilson, A. Miles, and M. T. Parker (ed.), *Topley and Wilson's principles of bacteriology, virology and immunity*, 7th ed., vol. 3. Williams and Wilkins, Baltimore.
- World Health Organization.** 2013. Cholera annual report. *Wkly. Epidemiol. Rec.* 88:321–336.
- World Health Organization.** 2014. Cholera annual report. *Wkly. Epidemiol. Rec.* 89:345–356.
- Xiao C, Kuznetsov YG, Sun S, Hafenstein SL, Kostyuchenko VA, Chipman PR, Suzan-Monti M, Raoult D, McPherson A, Rossmann MG.** 2009. Structural studies of the giant mimivirus. *PLoS. Biol.* 7:e92.
- Yamamoto KR, Alberts BM, Benzinger R, Lawhorne L, Treiber G.** 1970. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. *Virology* 40:734-744.
- Young R.** 2002. Bacteriophage holins: deadly diversity. *J. Mol. Microbiol. Biotechnol.* 4:21-36.
- Zahid MSH, Udden SMN, Faruque ASG, Calderwood SB, Mekalanos JJ, Faruque SM.** 2008. Effect of phage on the infectivity of *Vibrio cholerae* and emergence of genetic variants. *Infect. Immun.* 76:5266-5273.
- Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821-829.
- Zimmerman MD, Proudfoot M, Yakunin A, Minor W.** 2008. Structural insight into the mechanism of substrate specificity and catalytic activity of an HD-domain phosphohydrolase: the 5'-deoxyribonucleotidase YfbR from *Escherichia coli*. *J. Mol. Biol.* 378:215-226.

## 12. Artículo original

Solís-Sánchez et al. *Virology Journal* (2016) 13:47  
DOI 10.1186/s12985-016-0490-x

Virology Journal

RESEARCH

Open Access



# Genetic characterization of ØVC8 lytic phage for *Vibrio cholerae* O1

Alejandro Solís-Sánchez<sup>1,2</sup>, Ulises Hernández-Chiñas<sup>1,2</sup>, Armando Navarro-Ocaña<sup>1</sup>, Javier De la Mora<sup>3</sup>, Juan Xicohtencatl-Cortes<sup>4</sup> and Carlos Eslava-Campos<sup>1,2\*</sup>

### Abstract

**Background:** Epidemics and pandemics of cholera, a diarrheal disease, are attributed to *Vibrio cholerae* serogroups O1 and O139. In recent years, specific lytic phages of *V. cholerae* have been proposed to be important factors in the cyclic occurrence of cholera in endemic areas. However, the role and potential participation of lytic phages during long interepidemic periods of cholera in non-endemic regions have not yet been described. The purpose of this study was to isolate and characterize specific lytic phages of *V. cholerae* O1 strains.

**Methods:** Sixteen phages were isolated from wastewater samples collected at the Endhó Dam in Hidalgo State, Mexico, concentrated with PEG/NaCl, and purified by density gradient. The lytic activity of the purified phages was tested using different *V. cholerae* O1 and O139 strains. Phage morphology was visualized by transmission electron microscopy (TEM), and phage genome sequencing was performed using the Genome Analyzer IIx System. Genome assembly and bioinformatics analysis were performed using a set of high-throughput programs. Phage structural proteins were analyzed by mass spectrometry.

**Results:** Sixteen phages with lytic and lysogenic activity were isolated; only phage ØVC8 showed specific lytic activity against *V. cholerae* O1 strains. TEM images of ØVC8 revealed a phage with a short tail and an isometric head. The ØVC8 genome comprises linear double-stranded DNA of 39,422 bp with 50.8 % G + C. Of the 48 annotated ORFs, 16 exhibit homology with sequences of known function and several conserved domains. Bioinformatics analysis showed multiple conserved domains, including an Ig domain, suggesting that ØVC8 might adhere to different mucus substrates such as the human intestinal epithelium. The results suggest that ØVC8 genome utilize the “single-stranded cohesive ends” packaging strategy of the lambda-like group. The two structural proteins sequenced and analyzed are proteins of known function.

**Conclusions:** ØVC8 is a lytic phage with specific activity against *V. cholerae* O1 strains and is grouped as a member of the VP2-like phage subfamily. The encoding of an Ig domain by ØVC8 makes this phage a good candidate for use in phage therapy and an alternative tool for monitoring *V. cholerae* populations.

**Keywords:** *Vibrio cholerae*, Bacteriophage, *Caudovirales*, *Podoviridae*, ØVC8

\* Correspondence: carlos\_01eslava@yahoo.com.mx

<sup>1</sup>Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional Autónoma de México, Circuito Escolar S/N, Ciudad Universitaria, Delegación Coyoacán 04510, México, D.F., México

<sup>2</sup>Laboratorio de Patogenicidad Bacteriana, Unidad de Hemato-Oncología e Investigación, Hospital Infantil de México Federico Gómez/Facultad de Medicina, UNAM, Dr. Márquez No. 162, Col Doctores, Delegación Cuauhtémoc 06720, México, D.F., México

Full list of author information is available at the end of the article



© 2016 Solís-Sánchez et al. **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

## Background

Cholera is a clinical-epidemiologic syndrome caused by ingestion of water contaminated with *Vibrio cholerae* serogroups O1 and O139. This disease is considered an important public health problem worldwide, though it mainly affects developing countries and alters the economies of these regions [1]. From 1991 to 2001, the seventh pandemic of cholera affected Latin America, including Mexico. In October 2010, a cholera epidemic in Haiti resulted in over 180,000 cases in three months and spread rapidly to other countries, such as the Dominican Republic and Cuba [2]. *V. cholerae* has the ability to survive in aquatic habitats of different characteristics, including wastewater. During the process of adaptation to conditions of extreme pH, salinity, temperature, and nutrient insufficiency as well as predation by heterotrophic protists and bacteriophages, the expression of different genes is activated. A viable but non-culturable state or biofilm is then induced, which contributes to adaptation by the bacterium for survival in different environmental conditions [3]. Bacteriophages or phages (bacterial viruses) are mobile genetic elements that participate in horizontal gene transfer in bacteria, thereby contributing to their environmental adaptation and evolution. In addition, several bacterial virulence genes are present in phage genomes, and the mobile nature of phages can promote the emergence of new epidemic strains.

One of the main virulence factors of *V. cholerae* is cholera toxin (CT), which is encoded by CTX $\phi$ , a lysogenic filamentous phage that has contributed to bacterial evolution through lysogenic conversion and genomic rearrangement [4]. The *ctxAB* genes present in the CTX $\phi$  genome of toxigenic *V. cholerae* favor the conversion of nonpathogenic strains into toxigenic strains via CTX $\phi$  acquisition. The first vibrio phages were described in 1926 by d'Herelle, and in the 1950s, several distinct types of *V. cholerae* phages were described [5]. The use of bacteriophages as a tool for strain differentiation has contributed significantly to our understanding of cholera epidemiology [6]. In addition, the first phage-typing scheme for *V. cholerae* O1 was employed to study the spread of *V. cholerae* strains of the El Tor biotype [7]. Although, this phage-typing scheme has been used routinely for the classification of *V. cholerae* O1 strains due to its limitations, new phage-typing schemes for O139 strain classification have been developed [8]. Since 2007, more than 200 vibrio phages have been described; however, at present, only 17 genomes of *V. cholerae* phages have been sequenced and annotated in the GenBank database.

In recent years, lytic phages have been proposed as important factors modulating populations of *V. cholerae* serogroups O1 and O139 in the aquatic environment, thus affecting the seasonality and duration of cholera

epidemics in endemic areas [9]. In Bangladesh, which is considered an endemic cholera area, the prevalence of several predatory phages (JSF1 to JSF6) of *V. cholerae* has been partially characterized. Fluctuations in and the presence of the most prevalent phage types have been correlated with temporal changes in the cyclical appearance of cholera, acting as factors that modulate the epidemic cycle in the short period as well as outbreak severity [10]. In Mexico, conditions amenable to the survival of *V. cholerae* Non-O1/Non-O139 in aquatic reservoirs have been reported for several years [11]. However, the role of diverse phages in non-endemic cholera areas as elements that participate in the survival and occurrence of the bacterium during long interepidemic periods is not completely understood. In 2010, sporadic cholera cases were identified in Sinaloa State, México, and in 2013, an outbreak of 187 cases of cholera in Hidalgo State, México, was reported by the Secretaría de Salud de México ([www.epidemiologia.salud.gob.mx/dgae/boletin/intd\\_boletin.html](http://www.epidemiologia.salud.gob.mx/dgae/boletin/intd_boletin.html); [www.sinave.gob.mx/](http://www.sinave.gob.mx/)). Although the phages involved in the epidemiology of cholera in Mexico have not yet been characterized, predation of *V. cholerae* O1 by phages can be considered a key factor in understanding the long interepidemic periods of cholera in these regions. The main goal of this study was to isolate and characterize *V. cholerae* phages from wastewater of the Endo Dam in Hidalgo State, México, and to assess their lytic activity against *V. cholerae* O1 strains.

## Methods

### Sampling area

Samples were collected at the Endhó Dam in Hidalgo State, located 80 km north of Mexico City, Mexico. This ~1,260-hectare dam has a capacity of approximately 198 million m<sup>3</sup> and is the main reservoir of wastewater and rainwater from the metropolitan area of Mexico City and Hidalgo State [12].

### Isolation of phages

Four water samples of 200 ml each were collected from different points at the Endhó Dam, transported in glass bottles at room temperature, and processed on the same day. Briefly, 50 ml of each sample was centrifuged at 16,000  $\times$  g (RC5 rotor, Thermo Scientific, Pittsburgh, PA, USA), and the supernatant was filtered through a 0.22- $\mu$ m membrane (PVDF; Millipore, Bedford, MA, USA) to eliminate bacteria and/or diverse residues. The phages in the supernatants were isolated using *V. cholerae* O1 and O139 as receptor strains in double-layer plaque assays with soft agar (10 g/l tryptone, 10 g/l NaCl, and 7 g/l agar) as described by Kropinski et al. [13] (Table 1). In brief, 1 ml of each sample and 100  $\mu$ l of the receptor strain in exponential growth phase were mixed, and 4 ml of melted soft agar

**Table 1** Host range investigated by spot testing

Strain	ID	Source/Reference	ØVC1	ØVC2	ØVC3	ØVC6	ØVC7	ØVC8	ØVC9	ØVC11	ØVC18	ØVC19	ØVC20	ØVC28	ØVC30	ØVC31	ØVC33	ØVC34
87128	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87152	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87177	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87210	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87258	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87269	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87284	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87285	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
87662	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88226	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88538	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88539	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88682	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
88697	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
89207	<i>V. cholerae</i> O139	Thailand <sup>1</sup>																
90500	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
<b>96316</b>	<i>V. cholerae</i> O139	México																
102223	<i>V. cholerae</i> O1	Czech Republic <sup>1</sup>																
97639/1	<i>V. cholerae</i> O1	Mexico <sup>1</sup>																
<b>CVD 103</b>	<i>V. cholerae</i> O1	México <sup>1</sup>																
<b>E8021</b>	<i>V. cholerae</i> O1	HPA <sup>2</sup>																
N-16961	<i>V. cholerae</i> O1	México <sup>3</sup>																

Strains used for initial phage isolation are in bold

1. Strains isolated from different sources were provided by the Laboratory of Bacteriology, Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional Autónoma de México

2. Health Protection Agency, Laboratory of Gastrointestinal Pathogens, London, England

3. Provided by Dr. Shah M. Faruque of the Centre for Food and Water Borne Diseases at the International Centre for Diarrheal Disease Research, Bangladesh (ICDDR)

was added. The mixture was poured onto Petri dishes with nutrient agar (15 g/l agar) and incubated at 37 °C for 18 h.

#### Purification of phages

Lytic plaques were selected with a sterile Pasteur pipette and incubated in 50 ml of Luria-Bertani (LB) broth (LB was supplemented with 10 mM CaCl<sub>2</sub> and inoculated with ~4 × 10<sup>8</sup> colony-forming units (CFU/ml) of each receptor strain) at 37 °C for 18 h. The cellular debris was eliminated by centrifuging the samples twice at 9,000 × g/10 min/4 °C (5415C, Eppendorf, Hamburg, Germany). The supernatants containing the phages were filtered through a 0.22-µm membrane (Millipore PVDF, Bedford, MA, USA). The phages were concentrated adding 5 % polyethylene glycol 8000 (PEG) and 0.25 M NaCl at 4 °C overnight [14] and recovered by centrifugation at 16,000 × g (RC5 rotor, Thermo Scientific, Pittsburgh, PA, USA) for 20 min at 4 °C. The pellets were dissolved in 5 ml of SM buffer (50 mM Tris-HCl [pH 7], 100 mM NaCl and 50 mM MgSO<sub>4</sub>), loaded onto a cesium

chloride gradient (density, 1.3 to 1.6 g/ml), and centrifuged (41Ti rotor, Beckman, Brea, CA, USA) at 100,000 × g for 18 h at 4 °C to obtain relatively pure phages. The recovered phages were dialyzed three times against dialysis buffers "A" (3 M NaCl and 0.1 M Tris-HCl, pH 7.4) and "B" (0.3 M NaCl and 0.1 M Tris-HCl, pH 7.4) using a membrane with a molecular weight cut-off of 12,000 to 14,000 (Spectra/Pore, Spectrum Laboratories, Rancho Dominguez, CA, USA).

#### Isolation of *V. cholerae*

The four water samples previously collected from the Endhó Dam were also used for isolation of *V. cholerae* strains using an enrichment technique [15]. Briefly, 10 ml of each water sample was inoculated into 50 ml of alkaline peptone water and incubated at 37 °C for 6 h. The bacteria-enriched samples were cultured on thiosulfate-citrate-bile salts-sucrose agar (TCBS; Becton-Dickinson, Sparks, MD, USA) at 37 °C for 18 h. Yellow colonies produced by *V. cholerae* on TCBS agar were identified

using the GNI-Plus card (bioMérieux, Marcy l'Etoile, Rhône, France) and an automated Vitek system. Strains identified as *V. cholerae* were serotyped with anti-O1 and anti-O139 rabbit sera (Laboratorio de Patógenos Entéricos, UNAM, México).

#### Phage host range

The lytic activity of the purified phages was analyzed by a dot plaque assay employing the *V. cholerae* O1, *V. cholerae* O139, *V. cholerae* non-O1, *V. cholerae* non-O139, *V. alginolyticus*, *V. parahaemolyticus*, *V. metschnikovii*, *Aeromonas veronii*, *Escherichia coli* K-12 HB101, and *V. cholerae* isolates identified in this study (Table 1 and Additional file 1: Table S1). To assess phage activity, these strains were cultured on LB agar until log phase and incubated with 10  $\mu$ l of a phage suspension ( $10^8$  plaque-forming units [PFU]/ml) at 37 °C for 18 h [16]. The formation of turbid and/or clear plaques over the bacterial lawn was visualized with the naked eye and employed as the criterion for the selection of specific phages.

#### Visualization of the ØVC8 phage by transmission electron microscopy (TEM)

Two microliters of a purified phage ØVC8 suspension ( $1 \times 10^8$  PFU/ml) was placed on a Formvar-carbon-coated grid (EMS, Hatfield, PA, USA) for 2 min followed by negative staining with 6  $\mu$ l of 2 % uranyl acetate for 2 min. The morphology of the purified ØVC8 phage was examined under a JEM 1200 EXII transmission electron microscope (JEOL, Tokyo, Japan).

#### Isolation of phage DNA

Phage ØVC8 DNA was extracted as follows: a 500- $\mu$ l aliquot of phage ØVC8 suspension ( $1 \times 10^8$ /ml PFU) was treated with 100 U of *DNase* I (Invitrogen, Carlsbad, CA, USA), 2  $\mu$ l of proteinase K (20 mg/ml) (Vivantis, Oceanside, CA, USA), and 50  $\mu$ l of SDS (10 %) at 56 °C for 1 h. The purified DNA was treated with phenol-chloroform (1:1 ratio), precipitated with cold absolute ethanol, and resuspended in *DNase*-free water.

#### Genome sequencing and assembly of phage ØVC8

Total phage ØVC8 DNA was sequenced using Genome Analyzer Ix System (Illumina, San Diego, CA, USA) at the Massive Sequencing University Unit (UUSM, by its initials in Spanish) at the Institute of Biotechnology of UNAM (Cuernavaca, Morelos, Mexico). *De novo* assembly of ØVC8 was performed at Winter Genomics Company (www.wintergenomics.com) using the Velvet, Abyss, and SOAP programs. A consensus sequence from the results obtained with three programs was generated using the Minimus program [17].

#### Bioinformatics analysis of the ØVC8 genome

The consensus sequence of the ØVC8 genome was analyzed as described by Henn et al. [18]. To identify potential coding sequences, the programs BLAST-X and PFAM/TIGR were used to compare the ØVC8 genome sequence against the sequences of proteins reported in the databases. Open reading frames (ORFs) identified using the Glimmer3, Meta GeneAnnotator, GeneMarkS, ZCURVE\_V, and EasyGene gene prediction programs were grouped as a single locus. tRNAs and rRNAs were predicted using the tRNAscan-SE, Rfam, and ARNmmr programs. The pI/MW program was used to calculate the molecular weights of the identified proteins [19]. The transcriptional promoters and terminators were predicted using the PromoterHunter and WebGeSTer programs, respectively [20, 21].

#### Phylogenetic analysis of ØVC8

A phylogenetic analysis of ØVC8 was performed using the translated DNA sequence of ORF3 (terminase large subunit), which was aligned with 53 homologous sequences from phages of the *Podovirus* family using the ClustalW2 program [22]. Additionally, the phylogenetic tree was constructed with the Mega ver. 6.0 program using the neighbor-joining method, which employs a gamma distribution ( $\gamma = 2$ ) and 1,000 bootstrap replicates with Poisson distance correction [23].

#### Accession number of the nucleotide sequence

The genome sequence and the genetic annotation of genome ØVC8 were deposited in the GenBank database under accession number JF712866.

#### Structural proteins of phage ØVC8

Potential proteins in the phage ØVC8 capsid were identified according to the procedure described by Boulanger et al. [24]. The phage was precipitated with PEG/NaCl (as described above), mixed with Laemmli solution (65.8 mM Tris-HCl, [pH 6.8], 2.1 % SDS, 26.3 % [w/v] glycerol, 0.01 % bromophenol blue and 100 mM  $\beta$ -mercaptoethanol) and heated to 100 °C for 5 min. Proteins were separated by 10 % SDS-PAGE (polyacrylamide-sodium dodecyl sulfate gel electrophoresis) and visualized using Coomassie blue. The identified proteins were processed using a QTRAP 3200 mass spectrometer (Applied Biosystems/MDS Sciex, ON, Canada) at the Biochemistry Department of the Faculty of Medicine-UNAM.

## Results

#### Isolation of bacteria and phages

Thirteen isolates identified in wastewater samples from the Endhó Dam were characterized as *V. cholerae* non-O1/O139 (6 isolates), *V. alginolyticus* (4 isolates), and *A. veronii*

(3 isolates) (Additional file 1: Table S1). In addition, 16 phages were isolated from these wastewater samples.

#### Lytic activity and host specificity of the identified phages

The sixteen phages identified were tested using 53 gram-negative bacteria; although 58.49 % (31/53) of these bacteria were not infected by any phage (Table 1 and Additional file 1: Table S1). On the other hand, all the phages showed lytic and lysogenic activity when tested against 22 *V. cholerae* O1/O139 strains. In this assay, 75 % (12/16) of the phages produced clear and/or opaque plaques in both serotypes (*V. cholerae* O1 and O139), though four of the phages did not infect *V. cholerae* O139 (Table 1). Interestingly, only the phage designated as ØVC8 showed lytic activity against thirteen strains of *V. cholerae* O1, producing clear plaques ~1 mm in diameter without halos.

#### Morphology of the ØVC8 phage

The ØVC8 phage was stained with 2 % uranyl acetate and analyzed by TEM. Morphological analysis showed an isometric icosahedral capsid approximately 62 nm in diameter and a tail 16 nm in length (Fig. 1a and b). These characteristics are similar to the phages described in the *Podoviridae* family of the order *Caudovirales*.

#### Genome sequencing of ØVC8

The genomic sequence of ØVC8 was assembled into a single contig with a median coverage of 14,324x, composed of a double-stranded DNA molecule 39,422 bp in length with 50.8 % G + C content. Bioinformatics analysis of the ØVC8 genome sequence revealed 48 putative ORFs; of these, 30 % (14/48) can be assigned functions according to their homology to known sequences of other phages, 2.08 % (1/48) do not show similarity with any previously reported sequences, and 70.83 % (34/48) correspond to hypothetical proteins described in other phages (Table 2). The 48 ORFs are distributed on both DNA strands: one strand contains 22 ORFs related to

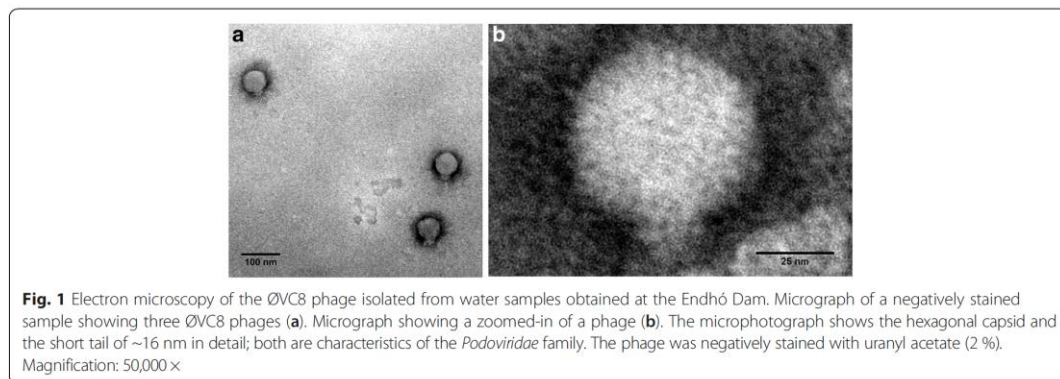
phage packaging functions and to mainly structural proteins, and the other 26 ORFs on the complementary strand are associated with metabolic, replication and unknown functions. tRNAs and rRNAs were not identified.

#### Functional organization of the ØVC8 genome

Hypothetical regulatory sequences were found in the intergenic regions of the ØVC8 genome: 15 correspond to promoter sequences (Table 3), and eight are associated with Rho-independent terminators (Fig. 2). Considering the specific regulatory sequence positions and ORF functions, the ØVC8 genome is organized into functional modules of packaging, head-tail morphogenesis, metabolism, and replication (Fig. 3). Additionally, two possible modules (ORFs 48 to 45 and 44 to 35) described as hypothetical proteins without assigned functions are located upstream of the replication module. Furthermore, five repeat sequences are present in the genome: four of these are inverted sequences of 20 bp, three are located in intergenic regions of both replication and metabolism modules, and the last is located in ORF39 (hypothetical protein). Additionally, a tandem direct repeats sequence of 54 bp was also found in the non-coding region 222 bp upstream of ORF48.

#### Packaging module

Seven ORFs were identified in this module, four of which are annotated as hypothetical proteins with unknown functions. Functions for ORFs 2, 3, and 4 can be ascribed to the small terminase subunit, large terminase subunit, and head-to-tail connecting protein, respectively (Fig. 3). The functional characteristics of the three ORFs and closely related homologs are described in Table 2. Characterization of the large terminase subunit as a protein widely conserved in the *Podoviridae* family is important for determining the close phylogenetic relationship of ØVC8 with *V. cholerae* phages VP2 and VP5 (Fig. 4).



**Table 2** Putative open reading frames in the ØVC8 genome and their assigned functions

ORF	Nucleotide position		Gene product		Presumed function	Related phage or organism	GenBank accession no	BLAST X	
	Start	End	Size (aa)	Molecular mass (kDa)				% Identity	e-value
1	1	222	73	13.6	Hypothetical protein	Vibrio phage VP2	YP_052974	95	1 e-32
2	128	739	203	22.9	Terminase small subunit	Vibrio phage VP5	YP_053007	94	4 e-131
3	726	2432	568	65.3	Terminase large subunit	Pelagibacter phage HTVC010P	YP_007517700	32	4 e-41
4	2443	4086	547	61.7	Head-to-tail connecting protein	Pelagibacter phage HTVC010P	YP_007517703	32	4 e-62
5	4086	4313	75	8	Hypothetical protein	Vibrio phage VP2	YP_052976	92	7 e-26
6	4326	4583	85	9.9	Hypothetical protein	Vibrio phage VP2	YP_052977	95	2 e-18
7	4552	5439	295	31.8	Hypothetical protein	Pelagibacter phage HTVC010P	YP_007517705	23	6 e-14
8	5607	6575	322	36.1	Structural protein	Pelagibacter phage HTVC010P	YP_007517707	24	1 e-21
9	6642	6911	89	9.2	Hypothetical protein	Vibrio phage VP2	YP_052980	100	7 e-35
10	6926	7774	282	31.1	Hypothetical protein	Vibrio phage VP5	YP_053013	91	1 e-46
11	7771	8136	121	13.4	Hypothetical protein	Vibrio phage VP2	YP_052982	96	1 e-59
12	8138	8611	157	18.2	Hypothetical protein	Vibrio phage KVP40	NP_899538	31	2 e-20
13	8604	8906	100	10.5	Hypothetical protein	Vibrio phage VP2	YP_052983	97	1 e-51
14	8930	11155	741	74.1	Tail protein	Vibrio phage VP5	YP_024980	90	0
15	11164	12948	594	65.5	Structural protein	Vibrio phage VP5	YP_024422	98	0
16	12951	13349	132	13.9	Hypothetical protein	Vibrio phage VP2	YP_052984	97	1 e-41
17	13346	15430	694	76.8	Structural protein	Vibrio phage VP2	YP_052985	95	0
18	15430	16602	390	41.5	Hypothetical protein	Vibrio phage VP5	YP_053019	95	2 e-151
19	16604	18952	782	87.2	Structural protein	Escherichia phage phiV10	YP_512274	27	6 e-68
20	18956	19984	342	36.6	Tail fiber	Vibrio phage VP5	YP_053020	96	2 e-162
21	20009	21322	437	44.9	Outer capsid protein	Vibrio phage VP2	YP_024425	31	1 e-35
22	21350	21646	98	11.2	Hypothetical protein	Pseudomonas phage PpW-3	YP_008873205	45	9 e-16
23	21723	22160	145	11.2	Hypothetical protein	Vibrio phage VP2	YP_053022	96	2 e-79
24	22150	22308	52	11.2	Hypothetical protein	Vibrio phage VP2	YP_052988	81	1 e-19
25	22305	22817	170	19.2	Metal dependent phosphohydrolase	Vibrio phage VP5	YP_024983	94	7 e-116
26	22819	23079	86	9.8	Hypothetical protein	Vibrio phage VP5	YP_053024	96	2 e-28
27	23081	24112	343	38.1	Adenylosuccinate synthetase	<i>Solibacter usitatus</i>	YP_827486	33	1 e-31
28	24154	26442	762	87	Integrase	Vibrio phage VP2	YP_024428	99	0
29	26432	28333	633	71.2	DNA Polymerase I	$\alpha$ -proteobacteria phage ØJL001	YP_223952	26	9 e-40
30	28330	28680	116	13.5	Hypothetical protein	Enterococcus phage EFDG1	AJP61480	53	8 e-14
31	28677	29192	171	19.2	ssDNA binding protein	Vibrio phage VP2	YP_024430	93	8 e-95
32	29253	30068	271	30.3	Hypothetical protein	Pseudomonas phage F8	YP_001294468	32	1 e-23
33	30119	30586	155	17.8	Hypothetical protein	Vibrio phage VP2	YP_052991	98	5 e-108
34	30649	32118	489	55.4	Superfamily II DNA/RNA helicases	Thermoanaerobacterium phage THSA-485A	YP_006546319	36	2 e-77



**Table 2** Putative open reading frames in the ØVC8 genome and their assigned functions (Continued)

35	32188	32538	116	12.6	Hypothetical protein	Vibrio phage VP5	YP_053027	84	1 e-26
36	32538	33230	230	24.5	Hypothetical protein	Vibrio phage CJY	AIZ01434	97	3 e-111
37	33381	33845	154	17.3	Hypothetical protein	Vibrio phage VP2	YP_052993	84	6 e-93
38	33842	34252	136	15.7	Hypothetical protein	Vibrio phage VP2	YP_052994	78	2 e-64
39	34254	34613	119	13.7	Hypothetical protein	Vibrio phage VP5	YP_053031	67	4 e-44
40	34621	35421	266	30.2	Hypothetical protein	Vibrio phage VP2	YP_052996	98	2 e-101
41	35490	35975	161	18.1	Hypothetical protein	Vibrio phage VP2	YP_052977	96	5 e-85
42	35985	36437	150	16	Hypothetical protein	Vibrio phage VP5	YP_053036	84	2 e-61
43	36543	36836	97	11	Hypothetical protein	Vibrio phage VP2	YP_053001	95	1 e-46
44	36848	37138	96	10.7	Hypothetical protein	Vibrio phage VP2	YP_053002	79	5 e-36
45	37245	37490	81	9.2	Hypothetical protein	Vibrio phage VP2	YP_053003	74	1 e-26
46	37495	38094	199	22.1	Hypothetical protein	Vibrio phage VP5	YP_053042	91	5 e-93
47	38093	38788	231	26.9	Hypothetical protein	Vibrio phage VP2	YP_053005	93	1 e-50
48	38669	39028	121	13.6	Hypothetical protein	Unknown	—	—	—

\*Predicted using the Compute pI/Mw tool, Swiss Institute of Bioinformatics

#### Head-tail morphogenesis module

Fourteen ORFs are located downstream of the packaging module, seven (ORFs 9 to 13, 16, and 18) of which are described as hypothetical proteins. In addition, five ORFs (ORFs 14, 15, 19, 20, and 21) were identified using BLAST-X, and two (ORFs 8 and 17) were identified by mass spectrometry. The hypothetical proteins of these ORFs show homology to structural proteins of the VP2 phage (Table 2). ORF14 codes for a 74.1-kDa protein similar to the tail protein of VP5 (identity = 90 %) and contains conserved traits of specific domains of the

immunoglobulin (Ig) superfamily. ORF21 shows 96 % identity with the capsid protein of VP2. Pfam analysis of the ORF21 sequence reveals a BNR/Asp-box repeat conserved domain of the bacterial neuraminidase or sialidase family.

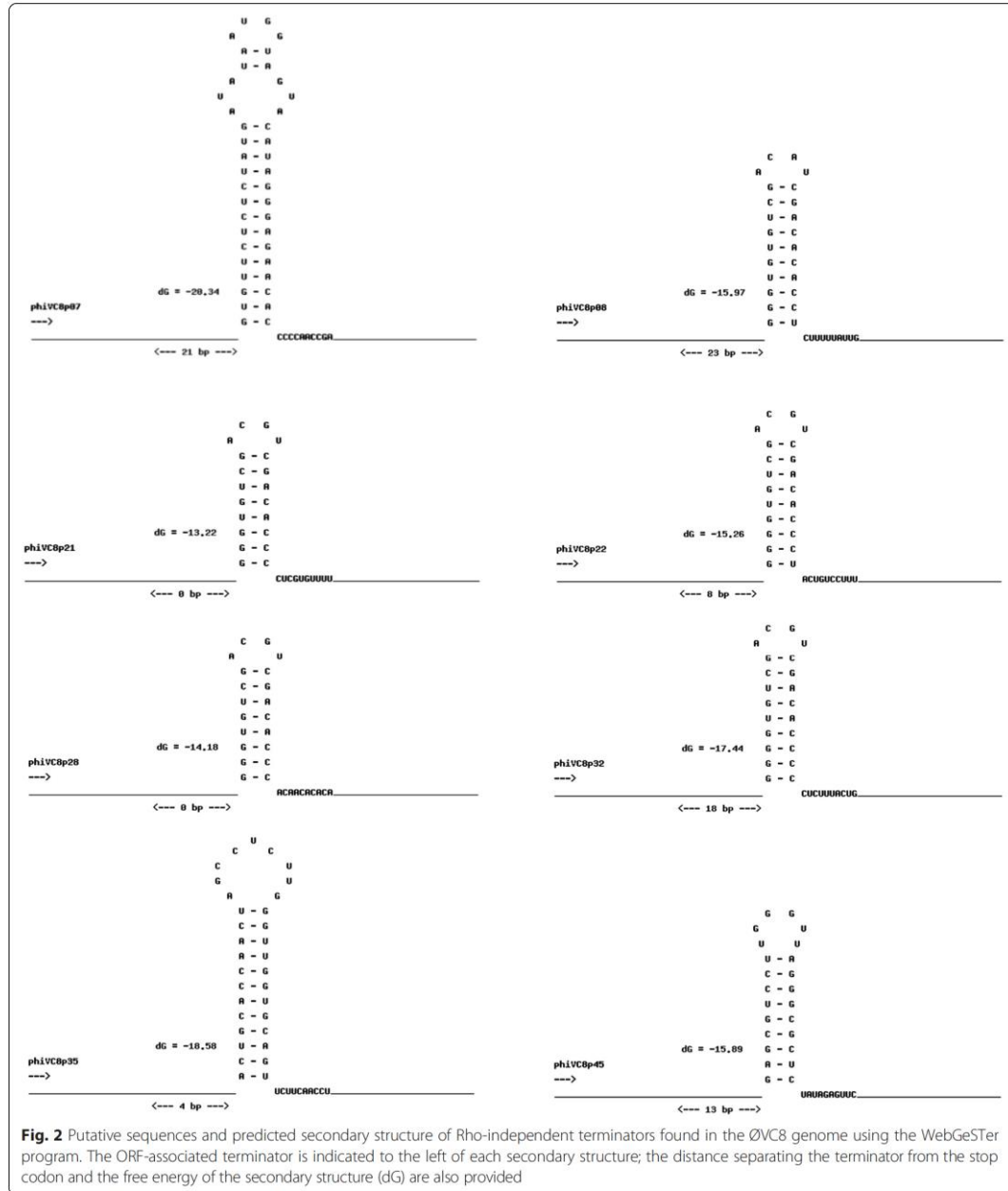
#### Replication module

Seven ORFs (34 to 28) located at the 3' end of the complementary strand encode proteins involved in ØVC8 phage replication. ORFs 34, 31, 29 and 28 are annotated as a helicase, single-strand DNA-binding protein (SSBP),

**Table 3** Predicted promoter sequences found in intergenic regions in the ØVC8 genome using the PromoterHunter program

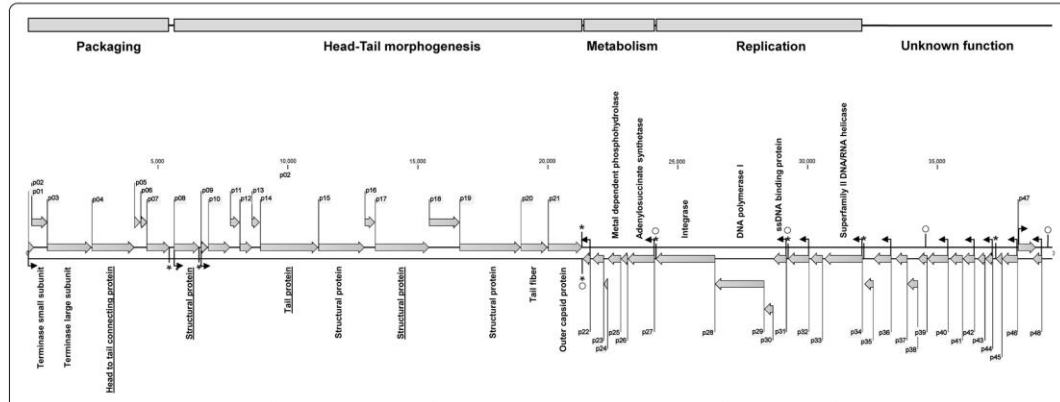
Promoter	ORF	Position	5'	- 35	- 10	3'
1	1	39373...39422	ATGTAACGGT- <b>GGTTGACA</b> -CAGAGCCAGAGGTGTG- <b>CTATAAT</b> -GGAGGTAAG			
2	8	5551...5601	AGAGTCCGGA- <b>AACGGGTA</b> -GCTCAAATCGATATTA- <b>ACAATCT</b> -TTTAGGAAAA			
3	9	6593...6641	GATTAGGG- <b>TGTGTGACA</b> -TCGACACACCTCTT- <b>TTAITGG</b> -AGAATACAT			
4	22	21652...21701	AACCCITGA- <b>TTTTTACC</b> -GCACCGTGAAGGGTG- <b>CAAAGAG</b> -GCAGGAGAAA			
5	27	24113...24153	AAGAATAGGG- <b>TGTCGACG</b> -TCGACACCCACAA- <b>ACACAAGAG</b> -GAAACACA			
6	31	29193...29235	GGGGTGTGCA- <b>CGTCGACA</b> -CCCCCTTTACTGG- <b>AGAGACT</b> -TAA			
7	32	30075...30118	C- <b>AAGTAACG</b> -AAGCAACTAGTAACGAAA- <b>TAACTGT</b> -CAATGGAGAC			
8	34	32127...32178	CCAAGTAGCC- <b>TCTGGGT</b> -TGTTGCAAGTTCTTCAAC- <b>CTAACCC</b> -AATGCAAGCA			
9	36	33231...33280	TGTCGCCAC- <b>CCTTTAGC</b> -TATCCAACCAACCA-ATCGTAG-GAGCAGCAGC			
10	40	35422...35472	CTGCAATAGA- <b>GCTTGACG</b> -CGCGCCCATATGTATGTT- <b>ATAATAG</b> -GTGTGTGCG			
11	42	36449...36500	TTGTGCAAT- <b>CTTTTATT</b> -TGGGGGAGGGCGTGTCC- <b>CTCCTGT</b> -CTCAACTAGT			
12	44	37139...37158	CAATGGTGTG- <b>GACGGTGC</b> -ACACACCCGAAACAC- <b>GTAAGGA</b> -GTCCGAA			
13	46	38150...38200	TTGCAAGGGT- <b>ACTTGACA</b> -GGCATCCGAAAGTGTG- <b>TTAAAT</b> -AAGAACACAA			
14	47	38040...38092	CTTGCTTGT- <b>GACCACCA</b> -CTCTGGTGTAGTTCGT- <b>AGCTGCT</b> -TGTGCTGCTC			
15	48	39031...38080	TCGGTCACAA- <b>AATGTTCCG</b> -AAATGTGACAAAAC- <b>CACATCT</b> -TTGACACACG			

Sequences in bold correspond to -10 and -35 regions

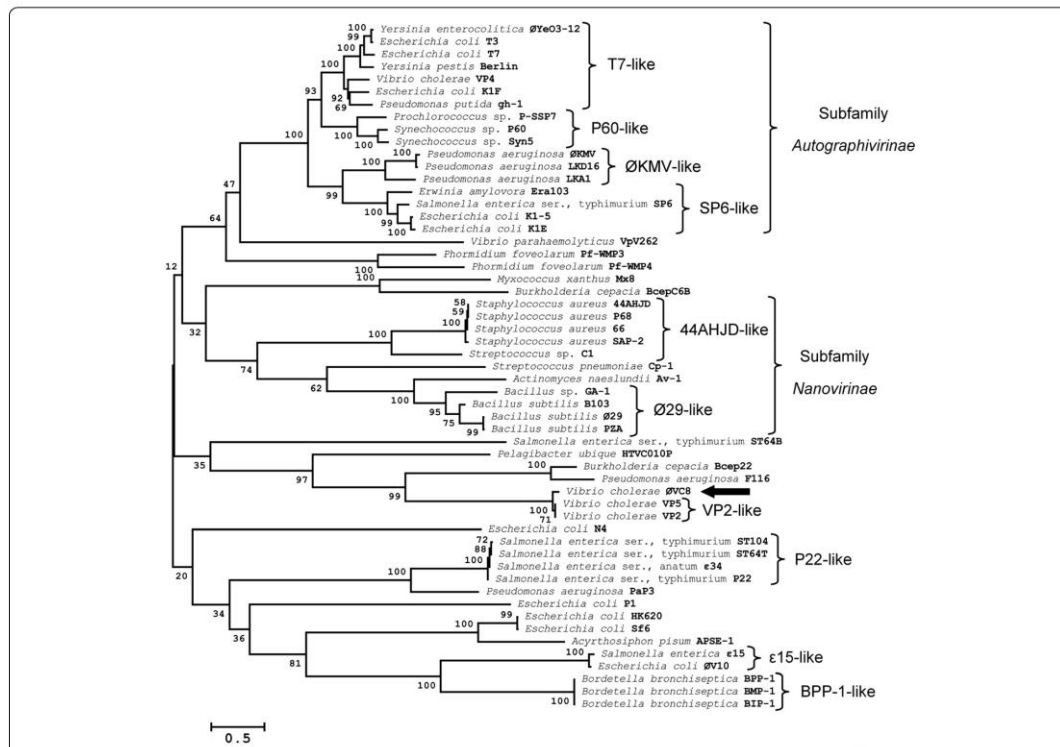


DNA polymerase I, and integrase functions, respectively (Fig. 3 and Table 2). ORF34 exhibits an amino-terminal conserved region with an SNF-2 domain that corresponded to the helicase-like ATP-dependent family. ORF31, which is located downstream, shows 93 %

identity with SSBPs of phage VP2. ORF29 codes for DNA polymerase I, which has also been described in T3 of *Myoviridae*, T5 of *Siphoviridae*, and T7 of *Podoviridae*. ORF28 encodes an integrase of VP2 and VP5 phages, with 99 and 56 % identity, respectively. A



**Fig. 3** ØVC8 genome organization. The start site was arbitrarily assigned to ORF 1. Arrows represent predicted genes and transcription direction. Promoter positions and directions are indicated by thin arrows; Rho-independent terminators, by asterisks; repeat sequences, by open circles; and underlined ORFs were also determined by mass spectrometry. Rectangles represent the proposed functional modules for the ØVC8 genome



**Fig. 4** Neighbor-joining tree for comparing the amino acid analysis of the terminase large subunit (ORF3) of ØVC8 and derived sequences in GenBank from 54 other phages of the Podoviridae family. The keys represent the major groups, and the respective hosts followed by the phage name are included at the end of each branch. Black arrow indicates the group of ØVC8 phage. The numbers in the internal nodes are bootstrap values (0-100 %) obtained from 1,000 bootstrap replicates

bifunctional-N-terminal primase/polymerase (N-Ter prim/pol) and a primase C-terminal-2 domain (PriCT-2) in ORF28 were also identified. ORFs 33, 32, and 30 are hypothetical proteins.

#### Metabolism module

The ØVC8 metabolism module is composed of six ORFs (22 to 27 ORFs). The sequences of ORFs 25 and 27 are enzymes involved in the metabolic pathways of amino acid synthesis. ORF25 encodes an HD-3 conserved domain with an H-21x-HD motif that corresponds to a metal-dependent phosphohydrolase from the HD superfamily, and ORF27 encodes an adenosyl succinate synthetase. ORFs 22, 23, 24, and 26 are hypothetical proteins (Fig. 3).

#### Comparative genomics of ØVC8 with VP2 and VP5

The BLASTX alignment results for the ØVC8 genome revealed 86 and 85 % identity compared with the VP2 and VP5 genomes, respectively [25]. Differences at the DNA sequence level include six regions of the ØVC8 genome compared with the VP2 genome and four regions compared with the VP5 genome. Consistently, the putative proteins of six ORFs localized in these regions of the ØVC8 genome display sequence similarities with different percentages of identity and positions compared to VP2 and VP5 phage proteins (Additional file 2: Figure S1). Interestingly, the putative protein of ORF30 is not present in the VP2 and VP5 genomes.

#### Phylogenetic analysis of the terminase large subunit

To investigate the possible role of the terminase subunit in the ØVC8 genome packaging process, bioinformatics analysis comparing the amino acid sequence of ORF3 with sequences of 88 phages of the order *Caudovirales* was performed. The phylogenetic tree grouped ØVC8 ORF3 into the same cluster with the terminases of VP2, VP5, CP-933 K, Fels-1, Gifsy-1, Gifsy-2, and Wo phages from the lambda-like group (Fig. 5).

#### Structural proteins

Structural proteins of ØVC8 phage were purified, analyzed, and compared with 48 ORFs of the ØVC8 phage genome (Table 2). Four main proteins (36.1, 61.7, 74.1, and 76.8 kDa) were identified by SDS-PAGE analysis and sequenced by mass spectrometry (Fig. 6). BLAST-P analysis of the amino-terminal sequences of the 61.7 and 74.1 kDa proteins showed matches to structural proteins of phages VP2 and VP5. In addition, 61.7 kDa protein peptides are homologous to specific peptides of head-to-tail (ORF4), and 74.1 kDa protein peptides are homologous to specific peptides of the tail protein (ORF14). In contrast, matches with hypothetical proteins of unknown functions were found for the 36.1- and 76.8-kDa

proteins, corresponding to ORFs 8 and 17; these proteins have been described in the VP2 and VP5 phages.

#### Discussion

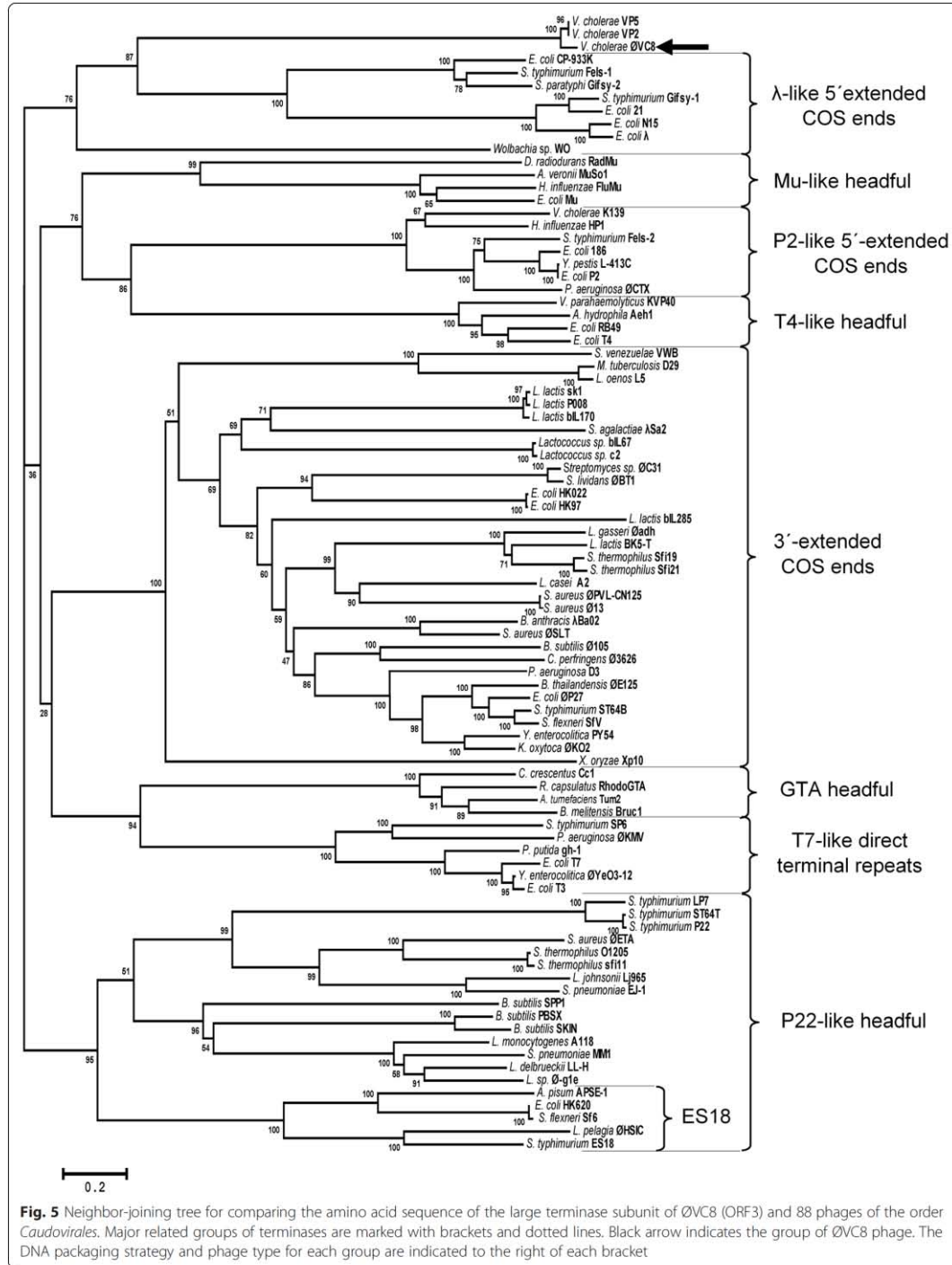
Phages as biological entities are abundant and widely distributed in the world and have great relevance in the control of bacterial communities. Fluctuations in phage populations during the seasonal behavior of cholera and the surveillance of *V. cholerae* in the aquatic environment are important factors that have been associated with cholera outbreaks [9]. In endemic cholera regions, *V. cholerae* phages have been detected in high frequency in different aquatic habitats, and these phages have been employed as strain markers for phage typing of *V. cholerae* O1 and O139 [8].

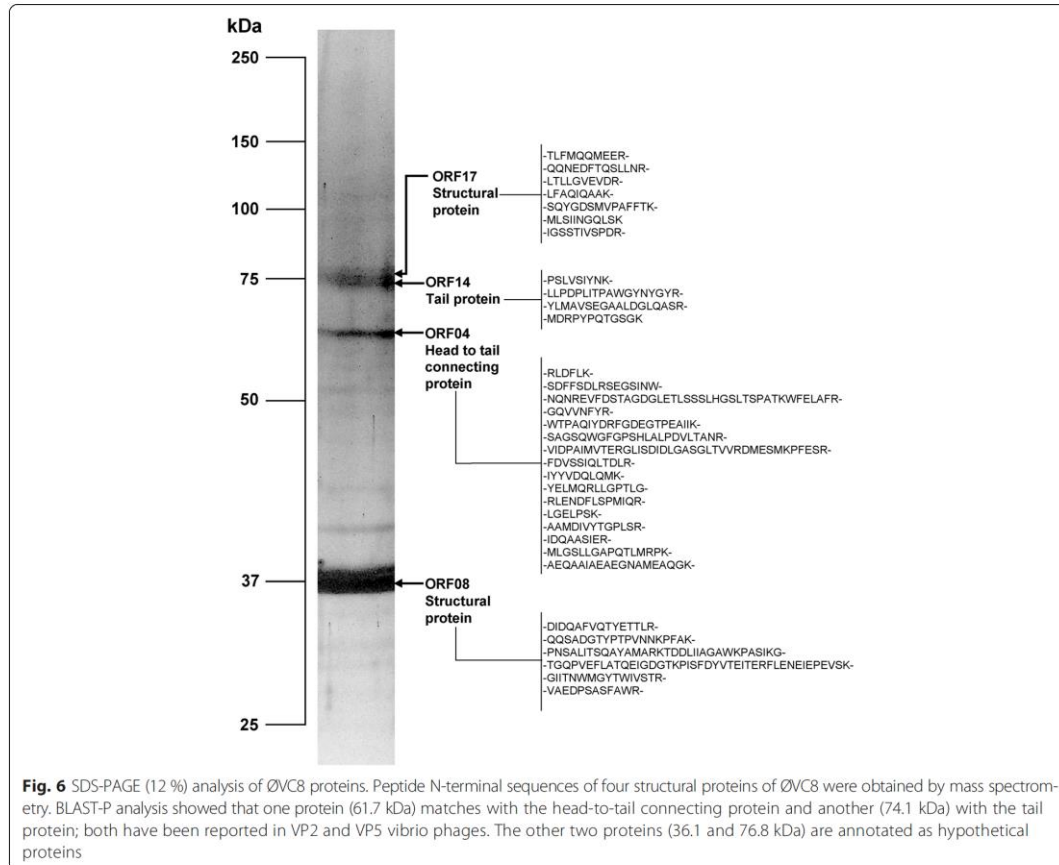
Thirteen bacterial isolates in wastewater samples collected from the Endhó Dam in Hidalgo State were identified as *V. cholerae* non-O1/O139, *V. alginolyticus* and *A. veronii*. Although toxigenic isolates of *V. cholerae* O1/O139 were not identified, the presence of non-O1/O139 *V. cholerae* strains is suggestive of the ability of these bacteria to survive for prolonged periods in sewage-polluted waters.

In endemic cholera areas, the presence of non-O1/O139 strains in the environment has been related to these bacteria serving as possible phage reservoirs with lytic activity against *V. cholerae* O1/O139 [9]. An abundant number of phages in wastewater treatment systems have been described, though little information regarding their population dynamics and their interaction with the microbial community has been published [26]. In the present work, only one phage, named ØVC8, showed lytic activity against *V. cholerae* O1 strains. *V. cholerae* predation by lytic phages has been proposed to be an important factor involved in the cyclical occurrence and severity of cholera outbreaks in endemic areas [10]. Thus, the presence of ØVC8, a lytic phage of *V. cholerae* O1 strains, could be involved in the epidemiology of cholera in Mexico, possibly regulating the presence of *V. cholerae* O1 strains for long periods; however, further studies are required to confirm this possibility.

The morphological characteristics of ØVC8 phage visualized by TEM showed a structure similar to *V. cholerae* phages JSF3 and JSF6 of the *Podoviridae* family, two phages that have been associated with the cyclic appearance of cholera in Bangladesh (Fig. 1). Considering the morphological classification of JSF3 and JSF6 phages, ØVC8 could be included in group III of the *V. cholerae* phage C1 morphotype, which includes OXN-100P, 4996, I, and III [5–27].

Sequencing of the ØVC8 genome revealed 48 putative ORFs distributed on both DNA strands and organized into packing, head-tail morphogenesis, metabolism,





replication, and unknown function modules. ORFs 2 and 3 encode for proteins of the terminase family, which are implicated in translocation of the viral capsid DNA during the final stage of phage assembly. Terminases are the most conserved proteins among caudoviruses, and they have been identified in all podoviruses [28]. Considering the presence of terminases as a potential marker of podoviruses, ORF3 of phage ØVC8 was analyzed by bioinformatics procedures. The results obtained allowed the identified terminases to be grouped in the same cluster with the terminases of phages VP2 and VP5 (Fig. 4). These data support that ØVC8 has a genomic organization that is similar to that of VP2 and VP5; therefore, ØVC8 could be included in the VP2-like subfamily proposed by Lavigne et al. [29].

During ØVC8 phage replication, a terminase protein is required for DNA packaging and for chromosomal end formation [30]. ORF3 encodes a large terminase subunit that could participate in this process. Indeed,

comparative analysis of the terminase protein coded by ORF3 suggests participation in the packing process via the “single-stranded cohesive ends” strategy, as has been described for lambda-like phages. Phages with these characteristics have a complementary sequence and generate protruding single strands called COS sites; these sites are highly conserved in the genome and are present in a region 1,000 bp upstream of the gene that encodes the small terminase subunit [28]. The above findings led us to propose that the COS site of ØVC8 is located in a tandem sequence 299 bp upstream of the terminase small subunit (Fig. 3). The interaction of head-to-tail connecting proteins with one of the procapsid vertices of the mature phage promotes formation of an axial pore for DNA translocation in both directions [31]. The presence of the head-to-tail connecting protein (ORF4) of ØVC8 suggests that this protein participates in this process during DNA packaging by translocating the chromosome into the procapsid and ejecting it during the infective stage

of the phage. These results indicate that ORF4 is the portal protein of phage ØVC8.

In some podovirus phages such as P22 from *Salmonella* sp., tail proteins have been described as molecules with the ability to recognize specific receptors during the initial stages of host infection [32]. In the head-tail morphogenesis module of ØVC8, ORF14 encodes a tail protein that could be involved in the recognition and infection of *V. cholerae* O1 strains. However, an Ig domain similar to that described for phage T4, which has been associated with functions of immune response and adhesion to eukaryotic cells, was also identified in ORF14 [33]. This domain has been grouped into the classic Ig domain (I-Set), which is widely distributed among bacteria, as well as the fibronectin type 3 (FN3) and bacterial Ig-like domain (Big2) families. Bioinformatics analysis of ORF14 revealed that its Ig domain corresponds to the I-Set family [34]. Recent studies have demonstrated that Ig-like domains are important in phage interaction with metazoan mucosal surfaces via specific adherence that might provide immunity independent of the host immune response [35]. Thus, the Ig domain of ORF14 may be an important element in ØVC8 phage interaction with the human intestinal mucosa, which is associated with the lytic activity of the phage in preventing *V. cholerae* O1 colonization.

Although ORF20 codes for a tail fiber protein, the presence of fibers in the ØVC8 phage was not observed by TEM. The tail fiber protein described for phage T7 consists of elongated homo-trimers that are responsible for the reversible initial recognition of a cell [36]. These structures are commonly composed of six fibers that are attached to the phage capsid, which hinders TEM analysis, and these fibers can be only visualized when host interaction occurs [37]. In our study, ØVC8 phage fibers were not observed, suggesting the possibility of a situation similar to that described for T7. In contrast, ORF21 of ØVC8 was annotated as a capsid protein, which encodes a BNR/Asp-box domain that has been described in the neuraminidase or sialidase family from bacteria and phages [38]. Proteins with sialidase activity are important for the degradation of bacterial polysaccharides; expression of these enzymes is an attractive feature for phage therapy [39]. Indeed, the presence of the BNR/Asp-box domain in ORF21 enables ØVC8 to be considered as a possible strategy for the treatment of cholera.

DNA/RNA helicases are widely distributed proteins that are required for the ATP-dependent unwinding of double-stranded DNA, an essential step in replication, expression, recombination, and DNA repair. In the replication module of the ØVC8 genome, ORF34 encodes a DNA/RNA helicase of the SNF-2 family with a conserved domain in its amino-terminal region that is involved in chromatin structure remodeling [40]. ORF31 codes for SSB, a protein that participates in replication,

recombination and DNA repair processes [41]. ORF29 encodes for a DNA polymerase I described in mitochondrial polymerase-g, prokaryotic DNA polymerase I, and diverse polymerases (T3, T5, and T7 phages) of the Pol A family [42]. Therefore, the helicase of ORF34 identified in this study may participate in transcription and replication processes of the ØVC8 phage genome.

ORF28 exhibits sequence homology with integrases of *V. cholerae* phages VP2 and VP5. Integrases achieve viral genome integration into the host genome via site-specific recombination of DNA sequences of 30 to 40 bp, with the first located on the phage chromosome (attP) and the second on the bacterial chromosome (attB). These enzymes are classified into two major families based on their amino acid sequence homology and catalytic residues, either tyrosine or serine. However, bioinformatics analysis shows no tyrosine or serine residues in the sequence of ØVC8 integrase or the corresponding sequences of VP2 and VP5 phages. Suggesting that the sequence does not correspond to an integrase or is a non-functional protein. Conserved bifunctional-N-terminal primase/polymerase domains (N-Ter prim/pol) and other primase C-terminal-2 domains (PriCT-2) were identified in the ORF28 sequence of the ØVC8 genome. N-Ter prim/pol is a multifunctional enzymatic domain with ATPase, primase, DNA polymerase, and helicase activity [43]. In contrast, the PriCT-2 domain belongs to the archaea-eukaryotic primase superfamily from the primase-polymerase clade (prim/pol-like) [44]. N-Ter prim/pol and PriCT-2 are essential domains of multifunctional replication proteins of the phage replication machinery. Accordingly, we speculate that ORF28 has bifunctional DNA primase/polymerase activity involved in ØVC8 phage replication.

In the metabolism module of the ØVC8 genome, ORFs 25 and 27 encode for enzymes involved in metabolic pathways of amino acid synthesis; these proteins are auxiliary metabolic molecules that may provide additional support in host metabolism steps, allowing successful phage infection [45]. HD-3, a conserved domain of ORF25, corresponds to proteins with a distinct combination of metal-chelating residues, nucleases and phosphodiesterase activities [46]. These data suggest that ORF25 could participate in ØVC8 phage signaling and nucleotide metabolism. ORF27 encodes an adenylosuccinate synthetase that participates in purine biosynthesis by catalyzing the GTP-dependent conversion of inosine monophosphate to adenosine monophosphate [47]. Notably, this enzyme is located at the same loci of VP2 (VP2p26) and VP5 (VP5\_gp26) chromosomes, and the presence of this enzyme constitutes one of the main distinguishing characteristics of the proposed VP2-like subfamily [29].

Recent studies have shown that some phages can employ alternative pathways of the classical holin-endolysin

lysis system by employing the host cell secretion machinery to deliver their endolysins [48]. Our observations showed that ØVC8 is a virulent phage with lytic activity against several *V. cholerae* O1 strains; however, none of the identified genes of the ØVC8 genome appear to be involved in bacterial lysis. One possible explanation for the lytic activity of ØVC8 is that this phage uses a lysis pathway that differs from the classic system.

Comparative genome analysis of ØVC8, VP2, and VP5 showed similar genome sequences and genetic organization. The presence of an adenylosuccinate synthetase and the lack of a lysis cassette are unique traits of these three phages. However, the genome of ØVC8 shows five insertion/deletions that have not been identified in the VP2 and VP5 genomes; these insertion/deletions are located mainly in the unknown function region and in the replication module (Additional file 2: Figure S1). The effect of these insertion/deletions on the phenotypes of VP2 and VP5 remain unknown, largely because of a lack of data regarding the characteristics of these phages.

Mass spectrometric analysis of the structural proteins of phage ØVC8 showed that these proteins are distributed among the packing and structural modules (ORFs 17, 14, 8, and 4), indicating that ØVC8 requires these four structural proteins for prophage assembly and potentially for initial host recognition. Additionally, ORF8, which encodes a protein of 36.1 kDa, was identified as one of the most abundant structural proteins, suggesting that this is a protein with a high copy number that is presumably the major capsid protein.

## Conclusions

The obtained results allow us to propose that ØVC8, which was identified in a non-endemic cholera area, is a new specific lytic phage for toxigenic *V. cholerae* O1 strains. Some of its features suggest that this phage could be considered a member of the VP2-like phage subfamily. Additionally, the presence of an Ig domain could confer to this phage the ability to adhere to different mucus substrates (including the human intestine), a situation that may influence the epidemiology of cholera. Certain features of phage ØVC8 may be employed as alternative tools for monitoring environmental populations of *V. cholerae* strains and suggest it as a potential candidate for phage therapy.

## Additional files

**Additional file 1: Table S1.** Strains used for the phage host range test. These strains were not susceptible to infection by any of the phages used in the assay. (DOC 50 kb)

**Additional file 2: Figure S1.** Whole-genome comparisons of ØVC8 against VP2 (up) and VP5 (down) using the Easyfig 2.1 program [49].

Predicted genes and transcription direction are represented as block arrows. ORFs are colored according to gene function, as indicated by the legend at the bottom. The degree of sequence similarity is indicated by color intensity, indicating the nucleotide identity levels (from 64 to 100 %). The comparisons were performed using BLASTn. Insertion/deletions and identity levels lower than 64 % are indicated by dotted squares. (DOC 22 kb)

## Competing interests

The authors declare that they have no competing interests.

## Authors' contribution

ASS and CEC conceived the study and designed the experiments. ASS performed the experiments and analyzed the data. UHU performed the data analysis. ANO performed the bacterial identification by biochemical and serological test. FMJ performed transmission electron microscopy experiments. ASS, CEC, and JXC, wrote the manuscript. CEC and JXC revised the manuscript. All authors read and approved the final manuscript.

## Acknowledgments

Guillermo Alejandro Solís Sánchez appreciates Doctorado en Ciencias Biomédicas-Universidad Nacional Autónoma de México, Facultad de Medicina, and Consejo Nacional de Ciencia y Tecnología (CONACyT) for Fellowship Program support (204332). We thank Dr. Edgar Oliver and M. en C. Esther Sánchez Espindola from Escuela Nacional de Ciencias Biológicas-IPN, and Dr. Fernando García Hernández from Unidad de Imagenología Instituto de Fisiología Celular, UNAM, for their technical assistance for this work.

## Author details

<sup>1</sup>Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional Autónoma de México, Circuito Escolar S/N, Ciudad Universitaria, Delegación Coyoacán 04510, México, D.F., México. <sup>2</sup>Laboratorio de Patogenicidad Bacteriana, Unidad de Hemato-Oncología e Investigación, Hospital Infantil de México Federico Gómez/Facultad de Medicina, UNAM. Dr. Márquez No. 162, Col Doctores, Delegación Cuauhtémoc 06720, México, D.F., México. <sup>3</sup>Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Ciudad Universitaria, Delegación Coyoacán 04510, México, D.F., México. <sup>4</sup>Laboratorio de Investigación en Bacteriología Intestinal, Unidad de Hemato-Oncología e Investigación Hospital Infantil de México Federico Gómez, Dr. Márquez No. 162, Col. Doctores, Delegación Cuauhtémoc 06720, México, D.F., México.

Received: 24 November 2015 Accepted: 18 February 2016

Published online: 22 March 2016

## References

- Glass RI, Black R. The epidemiology of cholera. In: Barua D, Greenough WB, editors. Cholera. New York: Plenum and Publishing House; 1992. p. 129–54.
- Andrews JR, Basu S. Transmission dynamics and control of cholera in Haiti: an epidemic model. *Lancet*. 2011;377:1248–55. doi:10.1016/S0140-6736(11)60273-0.
- Colwell RR. Viable but nonculturable bacteria: a survival strategy. *J Infect Chemother*. 2000;6:121–5. doi:10.1007/s101560000026.
- Waldor MK, Mekalanos JJ. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science*. 1996;272:1910–4. doi:10.1126/science.272.5270.1910.
- Ackermann HW, Kasatiya SS, Kawata T, Koga T, Lee JV, Mbiguino A, et al. Classification of *Vibrio* bacteriophages. *Intervirology*. 1984;22:61–71. doi:10.1159/000149535.
- Faruque SM. Role of phages in the epidemiology of cholera. *Curr Top Microbiol Immunol*. 2014;379:165–80. doi:10.1007/82\_2013\_358.
- Mukerjee S, Phil D. Bacteriophage typing of cholera. *Bull World Health Organ*. 1963;28:337–45.
- Chakrabarti AK, Ghosh AN, Nair GB, Niyogi SK, Bhattacharya SK, Sarkar BL. Development and evaluation of a phage typing scheme for *Vibrio cholerae* O139. *J Clin Microbiol*. 2000;38:44–9.
- Faruque SM, Naser IB, Islam MJ, Faruque AS, Ghosh AN, Nair GB, et al. Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages. *Proc Natl Acad Sci U S A*. 2005;102:1702–7. doi:10.1073/pnas.0408992102.



10. Jensen MA, Faruque SM, Mekalanos JJ, Levin BR. Modeling the role of bacteriophage in the control of cholera outbreaks. *Proc Natl Acad Sci U S A*. 2006;103:4652–7. doi:10.1073/pnas.0600166103.
11. Isaac-Márquez AP, Lezama-Dávila CM, Eslava-Campos C, Navarro-Ocaña A, Cravioto-Quintana A. Serotypes of *Vibrio cholerae* non-O1 isolated from water supplies for human consumption in Campeche, México and their antibiotic susceptibility pattern. *Mem Inst Oswaldo Cruz*. 1998;93:17–22. doi:10.1590/S0074-02761998000100004.
12. Siebe C, Cifuentes E. Environmental impact of wastewater irrigation in central Mexico: an overview. *Int J Environ Health Res*. 1995;5:161–73. doi:10.1080/09603129509356845.
13. Kropinski AM, Mazzocco A, Waddell TE, Lingohr E, Johnson RP. Enumeration of bacteriophages by double agar overlay plaque assay. In: Clokie MRJ, Kropinski AM, editors. *Bacteriophages, methods and protocols*. New York: Humana Press; 2009. p. 69–76.
14. Yamamoto KR, Alberts BM, Benzinger R, Lawhorne L, Treiber G. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. *Virology*. 1970;40:734–44. doi:10.1016/0042-6822(70)90218-7.
15. DeWitt WE, Gangarosa EJ, Huq I, Zarifi A. Holding media for the transport of *Vibrio cholerae* from field to laboratory. *Am J Trop Med Hyg*. 1971;20:685–8.
16. Powell R, Neilan J, Gannon F. Plaque dot assay. *Nucleic Acids Res*. 1986;14:1541. doi:10.1093/nar/14.3.1541.
17. Sommer DD, Delcher AL, Salzberg SL, Pop M. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics*. 2007;8:64. doi:10.1186/1471-2105-8-64.
18. Henn MR, Sullivan MB, Stange-Thomann N, Osburne MS, Berlin AM, Kelly L, et al. Analysis of high-throughput sequencing and annotation strategies for phage genomes. *PLoS One*. 2010;5:e9083. doi:10.1371/journal.pone.0009083.
19. Bjellqvist B, Hughes GJ, Pasquali C, Paquet N, Ravier F, Sanchez JC, et al. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis*. 1993;14:1023–31. doi:10.1002/elps.11501401163.
20. Klucar L, Stano M, Hajduk M. phiSITE database of gene regulation in bacteriophages. *Nucleic Acids Res*. 2010;38:D366–70. doi:10.1093/nar/gkp911.
21. Mitra A, Kesarwani AK, Pal D, Nagaraja V, WebGe S. Ter DB-a transcription terminator database. *Nucleic Acids Res*. 2011;19:1–7.
22. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23:2947–8. doi:10.1093/bioinformatics/btm404. PubMed: 17846036.
23. Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30:2725–9. doi:10.1093/molbev/mst197.
24. Boulanger P. Purification of bacteriophages and SDS-PAGE analysis of phage structural proteins from ghost particles. *Methods Mol Biol*. 2009;502:227–38. doi:10.1007/978-1-60327-565-1\_13.
25. Wang D, Wang M, Li Y, Dong H, Liu Z, Liu Y, et al. Complete genome sequence and analysis of *Vibrio cholerae* phage VP2. *Bing Xue Bao [Chinese J Virol]*. 2005;21:60–4.
26. Hantula J, Kurki A, Vuoriranta P, Bamford DH. Ecology of bacteriophages infecting activated sludge bacteria. *Appl Environ Microbiol*. 1991;57:2147–51.
27. Ackermann HW. Tailed bacteriophages: the order caudovirales. *Adv Virus Res*. 1998;51:135–201. doi:10.1016/S0065-3527(08)60785-X.
28. Casjens SR, Gilcrease EB. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. *Methods Mol Biol*. 2009;502:91–111. doi:10.1007/978-1-60327-565-1\_7.
29. Lavigne R, Seto D, Mahadevan P, Ackermann HW, Kropinski AM. Unifying classical and molecular taxonomic classification: analysis of the *Podoviridae* using BLASTp-based tools. *Res Microbiol*. 2008;159:406–14. doi:10.1016/j.resmic.2008.03.005.
30. Casjens SR, Gilcrease EB, Winn-Stapley DA, Schickmaier P, Schmieger H, Pedulla ML, et al. The generalized transducing *Salmonella* bacteriophage ES18: complete genome sequence and DNA packaging strategy. *J Bacteriol*. 2005;187:1091–104. doi:10.1128/JB.187.3.1091-1104.2005.
31. Carrascosa JL, Viñuela E, García N, Santisteban A. Structure of the head-tail connector of bacteriophage phi. *J Mol Biol*. 1982;29:311–24.
32. Berget PB, Potete AR. Structure and functions of the bacteriophage P22 tail protein. *J Virol*. 1980;34:234–43.
33. Fraser JS, Yu Z, Maxwell KL, Davidson AR. Ig-like domains on bacteriophages: A tale of promiscuity and deceit. *J Mol Biol*. 2006;359:496–507. doi:10.1016/j.jmb.2006.03.043.
34. Fraser JS, Maxwell KL, Davidson AR. Immunoglobulin-like domains on bacteriophage: weapons of modest damage? *Curr Opin Microbiol*. 2007;10:382–7. doi:10.1016/j.mib.2007.05.018.
35. Barr JJ, Auro R, Furlan M, Whiteson KL, Erb ML, Pogliano J, et al. Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc Natl Acad Sci U S A*. 2013;110:10771–6. doi:10.1073/pnas.1305923110.
36. García-Doval C, van Raaij MJ. Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Proc Natl Acad Sci U S A*. 2012;109:9390–5. doi:10.1073/pnas.1119719109.
37. Hu B, Margolin W, Molineux IJ, Liu J. The bacteriophage T7 virion undergoes extensive structural remodeling during infection. *Science*. 2013;339:576–9. doi:10.1126/science.1231887.
38. Copley RR, Russell RB, Ponting CP. Sialidase-like asp-boxes: sequence-similar structures within different protein folds. *Protein Sci*. 2001;10:285–92. doi:10.1110/ps.31901.
39. Yan J, Mao J, Xie J. Bacteriophage polysaccharide depolymerases and biomedical applications. *BioDrugs*. 2014;28:265–74. doi:10.1007/s40259-013-0081-y.
40. Ryan DP, Owen-Hughes T. Snf2-family proteins: chromatin remodellers for any occasion. *Curr Opin Chem Biol*. 2011;15:649–56. doi:10.1016/j.cbpa.2011.07.022.
41. Szczepanska AK, Bidnenko E, Plochocka D, McGovern S, Ehrlich SD, Bardowski J, et al. A distinct single-stranded DNA-binding protein encoded by the *Lactococcus lactis* bacteriophage bIL67. *Virology*. 2007;363:104–12. doi:10.1016/j.virol.2007.01.023.
42. Franklin MC, Wang J, Steitz TA. Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell*. 2001;105:657–67. doi:10.1016/S0092-8674(01)00367-1.
43. Halgasova N, Mesarosova I, Bukovska G. Identification of a bifunctional primase-polymerase domain of corynephage BFK20 replication protein gp43. *Virus Res*. 2012;163:454–60. doi:10.1016/j.virusres.2011.11.005.
44. Iyer LM, Koonin EV, Leippe DD, Aravind L. Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. *Nucleic Acids Res*. 2005;33:3875–96. doi:10.1093/nar/gki702.
45. Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe J, et al. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc Natl Acad Sci U S A*. 2011;108:E757–64. doi:10.1073/pnas.1102164108.
46. Aravind L, Koonin EV. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem Sci*. 1998;23:469–72. doi:10.1016/S0968-0004(98)01293-6.
47. Wiesmüller L, Wittbrodt J, Noegel AA, Schleicher M. Purification and cDNA-derived sequence of adenylosuccinate synthetase from *Dictyostelium discoideum*. *J Biol Chem*. 1991;266:2480–5.
48. Catalão MJ, Gil F, Moniz-Pereira J, São-José C, Pimentel M. Diversity in bacterial lysis systems: bacteriophages show the way. *FEMS Microbiol Rev*. 2013;37:554–71. doi:10.1111/1574-6976.12006.
49. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics*. 2011;27:1009–10. doi:10.1093/bioinformatics/btr039.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
www.biomedcentral.com/submit



### 13. Propuesta de creación de un nuevo género de virus



This form should be used for all taxonomic proposals. Please complete all those modules that are applicable (and then delete the unwanted sections). For guidance, see the notes written in blue and the separate document "Help with completing a taxonomic proposal"

Please try to keep related proposals within a single document; you can copy the modules to create more than one genus within a new family, for example.

#### MODULE 1: TITLE, AUTHORS, etc

<b>Code assigned:</b>	<b>2015.031a-dB</b>	(to be completed by ICTV officers)			
<b>Short title:</b> Create one (1) new genus, <i>Vp5virus</i> , including three (3) new species within the family <i>Podoviridae</i> (e.g. 6 new species in the genus <i>Zetavirus</i> )					
<b>Modules attached</b> (modules 1 and 10 are required)	1 <input checked="" type="checkbox"/>	2 <input checked="" type="checkbox"/>	3 <input checked="" type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>
	6 <input type="checkbox"/>	7 <input type="checkbox"/>	8 <input type="checkbox"/>	9 <input type="checkbox"/>	10 <input checked="" type="checkbox"/>

#### Author(s):

Andrew M. Kropinski – University of Guelph (Canada)  
Evelien M. Adriaenssens – University of Pretoria (South Africa)  
Guillermo A. Solís-Sánchez – Hospital Infantil de México “Federico Gómez”/ Facultad de Medicina UNAM (Mexico)  
Carlos A. Eslava Campos – Hospital Infantil de México “Federico Gómez”/ Facultad de Medicina UNAM (Mexico)

#### Corresponding author with e-mail address:

Andrew Kropinski [Phage.Canada@gmail.com](mailto:Phage.Canada@gmail.com)

#### List the ICTV study group(s) that have seen this proposal:

A list of study groups and contacts is provided at <http://www.ictvonline.org/subcommittees.asp> . If in doubt, contact the appropriate subcommittee chair (fungal, invertebrate, plant, prokaryote or vertebrate viruses)

Bacterial and Archaeal Virus Subcommittee

#### ICTV Study Group comments (if any) and response of the proposer:

Please note that we have chosen to refer to this new genus as *B4virus* rather than *B4likevirus* since the Bacterial and Archaeal Virus Subcommittee of ICTV has voted overwhelmingly in favour of eliminating “like” from phage genus names.

Date first submitted to ICTV:

June 2015

Date of this revision (if different to above):

#### ICTV-EC comments and response of the proposer:

--

MODULE 2: **NEW SPECIES**

creating and naming one or more new species.

If more than one, they should be a group of related species belonging to the same genus. All new species must be placed in a higher taxon. This is usually a genus although it is also permissible for species to be “unassigned” within a subfamily or family. Wherever possible, provide sequence accession number(s) for **one** isolate of each new species proposed.

Code	<b>2015.031aB</b>	(assigned by ICTV officers)
<b>To create 3 new species within:</b>		
Genus:	<i>Vp5virus (new)</i>	Fill in all that apply. <ul style="list-style-type: none"> <li>• If the higher taxon has yet to be created (in a later module, below) write “<b>(new)</b>” after its proposed name.</li> <li>• If no genus is specified, enter “<b>unassigned</b>” in the genus box.</li> </ul>
Subfamily:		
Family:	<i>Podoviridae</i>	
Order:	<i>Caudovirales</i>	
<b>Name of new species:</b>	<b>Representative isolate: (only 1 per species please)</b>	<b>GenBank sequence accession number(s)</b>
<i>Vibrio virus VP5</i>	Vibriophage VP5	AY510084
<i>Vibrio virus VC8</i>	Vibrio phage phiVC8	JF712866
<i>Vibrio virus VP2</i>	Vibriophage VP2	AY505112

<p><b>Reasons to justify the creation and assignment of the new species:</b></p> <ul style="list-style-type: none"> <li>• Explain how the proposed species differ(s) from all existing species.           <ul style="list-style-type: none"> <li>○ If species demarcation criteria (see module 3) have previously been defined for the genus, <b>explain how the new species meet these criteria.</b></li> <li>○ If criteria for demarcating species need to be defined (because there will now be more than one species in the genus), please state the proposed criteria.</li> </ul> </li> <li>• Further material in support of this proposal may be presented in the Appendix, Module 9</li> </ul> <p>We have chosen 95% DNA sequence identity as the criterion for demarcation of species in this new genus. Each of the proposed species differs from the others with more than 5% at the DNA level as confirmed with the BLASTN algorithm.</p>
--

MODULE 3: **NEW GENUS**

creating a new genus

Ideally, a genus should be placed within a higher taxon.

Code	<b>2015.031bB</b>	(assigned by ICTV officers)
<b>To create a new genus within:</b>		
Subfamily:		Fill in all that apply. <ul style="list-style-type: none"> <li>• If the higher taxon has yet to be created (in a later module, below) write “<b>(new)</b>” after its proposed name.</li> <li>• If no family is specified, enter “<b>unassigned</b>” in the family box</li> </ul>
Family:	<i>Myoviridae</i>	
Order:	<i>Caudovirales</i>	

naming a new genus

Code	<b>2015.031cB</b>	(assigned by ICTV officers)
<b>To name the new genus: <i>Vp5virus</i></b>		

Assigning the type species and other species to a new genus

Code	<b>2015.031dB</b>	(assigned by ICTV officers)
<b>To designate the following as the type species of the new genus</b>		
<i>Vibrio virus VP5</i>	Every genus must have a type species. This should be a well characterized species although not necessarily the first to be discovered	
The new genus will also contain any other new species created and assigned to it (Module 2) and any that are being moved from elsewhere (Module 7b). <b>Please enter here the TOTAL number of species (including the type species) that the genus will contain:</b>		
3		

**Reasons to justify the creation of a new genus:**

Additional material in support of this proposal may be presented in the Appendix, Module 9

These three phages are a phylogenetically unique group of *Vibrio cholerae*-specific viruses. Since descriptions of only one of these phages has been published (3), the following analysis is derived from the sequences in GenBank. While the presence of an integrase might lead one to suspect that these phages are temperate, a new analysis of these proteins reveals that they contain primase-polymerase domains (Prim\_Pol, cd04859; Prim-Pol, pfam09250; PriCT\_2, pfam08707; Prim-Pol, smart00943). These phages also contain an “adenylosuccinate synthetase” which is supported by motif analysis (Adenylsucc\_synt, smart00788; PRK04293; PurA, COG0104; purA, TIGR00184; Adenylsucc\_synt, pfam00709; AdSS, cd03108; and PLN02513, LN02513). Its closest homologs are to similarly named proteins from *Pithovirus sibericum* (YP\_009000992) and *Megavirus courdo7* (AEX62163).

“Adenylosuccinate synthetase (AdSS) catalyzes the first step in the de novo biosynthesis of AMP. IMP and L-aspartate are conjugated in a two-step reaction accompanied by the hydrolysis of GTP to GDP in the presence of Mg<sup>2+</sup>” (NCBI).

It also encodes a DNA polymerase I homolog with 3'-5' exonuclease and polymerase domains whose primary homologs are to marine viroplankton proteins.

It would appear that these phages have global distribution having been isolated in China and Mexico.

A phylogenetic analysis is not called for since most of the proteins are poorly related to any other phage protein.

**Origin of the new genus name:**

*Vibrio* phage VP5

**Reasons to justify the choice of type species:**

The first fully sequenced member of this genus

**Species demarcation criteria in the new genus:**

If there will be more than one species in the new genus, list the criteria being used for species demarcation and explain how the proposed members meet these criteria.

We have chosen 95% DNA sequence identity as the criterion for demarcation of species in this new genus. Each of the proposed species differs from the others with more than 5% at the DNA level as confirmed with the BLASTN algorithm.

MODULE 10: **APPENDIX**: supporting material

additional material in support of this proposal

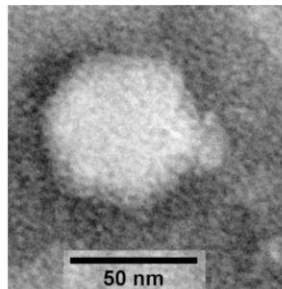
**References:**

1. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One. 2010; 5(6):e11147.
2. Turner D, Reynolds D, Seto D, Mahadevan P. CoreGenes3.5: a webserver for the determination of core genes from sets of viral and small bacterial genomes. BMC Res Notes. 2013; 6:140. doi: 10.1186/1756-0500-6-140.
3. Wang, D.; Wang, M.; Li, Y.; Dong, H.; Liu, Z.; Liu, Y.; Qi, G.; Gao, S.; Jin, W.; Kan, B. Complete genome sequence and analysis of *Vibrio cholerae* phage VP2. Bing Du Xue Bao. 2005, 21, 60-64.

**Annex:**

Include as much information as necessary to support the proposal, including diagrams comparing the old and new taxonomic orders. The use of Figures and Tables is strongly recommended but direct pasting of content from publications will require permission from the copyright holder together with appropriate acknowledgement as this proposal will be placed on a public web site. For phylogenetic analysis, try to provide a tree where branch length is related to genetic distance.

**Fig. 1** Electron micrograph of uranyl acetate negatively stained phage phiVC8 (provided by Dr. Eslava Campos)



**Table 1.** Properties of the three phages belonging to the genus *Vp5virus*

Phage	GenBank Accession No.	Genome size (bp)	Genome (mol%G+C)	No. CDS	No. tRNAs	DNA (% sequence identity)*	% Homologous proteins **
VP5	AY510084	39,786	50.5	48	0	100	100
phiVC8	JF712866	39,422	50.8	48	0	82	100
VP2	AY505112	39,853	50.6	47	0	92	91.7

\* Determined using BLASTN; \*\* Determined using CoreGenes

**Table 2.** Phages closely related to VP2

Vibrio phage QH	KM612259
Vibrio phage J2	KM612264
Vibrio phage J3	KM612265
Vibrio phage H3	KM612263
Vibrio phage H2	KM612262
Vibrio phage H1	KM612261
Vibrio phage CJY	KM612260

**Fig. 2.** progressiveMauve alignment (1) of the annotated genomes of members of the *Vp5virus* genus (from top to bottom: VP5, VP2, and phiVC8). Colored blocks indicate the regions of 1 to 1 best alignment with rearrangement breakpoints in a different random color. The degree of sequence similarity between regions is given by a similarity plot within the colored blocks with the height of the plot proportional to the average nucleotide identity (Aaron Darling, personal communication).

