






## Characterization of CRISPR genetic sequences in microorganisms associated with infections in shrimp (*Litopenaeus vannamei*)

Caracterización de secuencias genéticas CRISPR en microorganismos asociados a infecciones en camarón (*Litopenaeus vannamei*)

Caracterização de sequências genéticas CRISPR em microrganismos associados a infecções em camarões (*Litopenaeus vannamei*)

Ángel Parra<sup>1</sup>, Carla Lossada<sup>4</sup>, Aleivi Pérez<sup>3</sup>, Johnny Navarrete<sup>5</sup> and Lenin González<sup>1,2\*</sup>

Departamento de Biología, Facultad Experimental de Ciencias. La Universidad del Zulia. <sup>1</sup>Laboratorio de Genética y Biología Molecular. Email: (AP) [angelparra115@gmail.com](mailto:angelparra115@gmail.com), ; Email: (LG) [lgonzalezpaz@gmail.com](mailto:lgonzalezpaz@gmail.com), . <sup>2</sup>Laboratorio de Citogenética. <sup>3</sup>Laboratorio de Microbiología General. Email: [aleiviciencias@gmail.com](mailto:aleiviciencias@gmail.com), . <sup>4</sup>Laboratorio de Caracterización Molecular y Biomolecular – Centro de Investigación y Tecnología de los Materiales – Instituto Venezolano de Investigación Científica - Zulia. Sección 526. Maracaibo, Venezuela. Email: [lossadacarla@gmail.com](mailto:lossadacarla@gmail.com), . <sup>5</sup>Escuela Superior Politécnica Agropecuaria de Manabí Manuel Félix López, Sitio El Limón, Campus Politécnico Calceta, Manabí, Ecuador. Email: [jnava\\_57@hotmail.com](mailto:jnava_57@hotmail.com), .

### Abstract

In shrimp farming, the family of proteobacteria Vibrionaceae, especially the species of the genus *Vibrio*, represent one of the main responsible for infections in shrimp production (*Litopenaeus vannamei*), generating great losses to this industry. Phagotherapy emerges as a novel alternative for the control of said infections in substitution to the use of antibiotics, thanks to the specific inhibitory activity of these viruses. However, it is necessary to take into account the presence in prokaryotes of genetic sequences called clusters of regularly interspaced short palindromic repeats (CRISPR) that act as an immune system against invasion of external mobile genetic elements such

Recibido el 23-08-2019 • Aceptado el 26-10-2020.

\*Autor de correspondencia. Correo electrónico: [lgonzalezpaz@gmail.com](mailto:lgonzalezpaz@gmail.com)

as phage or plasmids. Due to its characteristics, the CRISPR/Cas system is used as a tool for gene editing. This study presents the comparative analysis of 7 CRISPR loci found in 5 sequences of complete genomes, available in the database of NCBI/GenBank, to determine the potential use of the phage strategy in shrimp farming. The CRISPR systems corresponded to types I-E, I-F and III-D. 53 % of the spacers (75/142) presented homology with plasmids, while the remaining 47 % (67/142) showed homology with bacteriophages, mostly non-typical *Vibrio* infective viruses. The use of phage therapy is proposed as a treatment for infections caused by members of the family Vibrionaceae in shrimp cultures, due to the low occurrence of CRISPR systems in the species studied and the low immunity to their phages, thus ensuring greater sensitivity.

**Keywords:** CRISPR-Cas systems, *Vibrio*, phagotherapy, bioinformatic analysis.

### Resumen

En el cultivo del camarón, la familia de las proteobacterias Vibrionaceae, especialmente las especies del género *Vibrio*, representan uno de los principales responsables de las infecciones en la producción de camarones (*Litopenaeus vannamei*), generando grandes pérdidas a esta industria. La fagoterapia surge como una alternativa novedosa para el control de dichas infecciones en sustitución al uso de antibióticos, gracias a la actividad inhibitoria específica de estos virus. Sin embargo, es necesario tomar en cuenta la presencia en procariotas de secuencias genéticas denominadas repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR) que actúan como sistema inmune contra la invasión de elementos genéticos móviles externos como fagos o plásmidos. Dadas sus características, el sistema CRISPR/Cas se emplea como herramienta para la edición de genes. En este estudio se presenta el análisis comparativo de 7 loci CRISPR encontrados en 5 secuencias de genomas completos, disponibles en la base de datos del NCBI/GenBank, para determinar la potencialidad del empleo de la estrategia fágica en camaronicultura. Los sistemas CRISPR correspondieron a los tipos I-E, I-F y III-D. El 53 % de los espaciadores (75/142) presentó homología con plásmidos, mientras que el 47 % restante (67/142) mostró homología con bacteriófagos, en su mayoría virus no típicos infectivos de *Vibrio*. Se propone el uso de la fagoterapia como tratamiento de infecciones causadas por miembros de la familia Vibrionaceae en cultivos de camarón debido a la baja ocurrencia de sistemas CRISPR en las especies estudiadas y la baja inmunidad a sus fagos, asegurando así una mayor sensibilidad.

**Palabras clave:** sistemas CRISPR-Cas, *Vibrio*, fagoterapia, análisis bioinformático.

## Resumo

Na criação de camarões, a família das proteobacteria Vibrionaceae, especialmente as espécies agrupadas no género *Vibrio*, representam uma das principais responsáveis por infecções na produção de camarão (*Litopenaeus vannamei*), gerando grandes prejuízos a esta indústria. A terapia fágica surge como uma nova alternativa para o controle dessas infecções em substituição ao uso de antibióticos, graças à atividade inibitória específica desses vírus. No entanto, é necessário levar em consideração a presença em procariotos de sequências genéticas denominadas repetições palindrômicas curtas agrupadas e regularmente espaçadas (CRISPR) que atuam como um sistema imunológico contra a invasão de elementos genéticos móveis externos, como fagos ou plasmídeos. Pelas suas características, o sistema CRISPR / Cas é utilizado como ferramenta de edição de genes. Este estudo apresenta a análise comparativa de 7 locos CRISPR encontrados em 5 sequências de genomas completos, disponíveis no banco de dados do NCBI/GenBank, para determinar o potencial de utilização da estratégia de fagos na carcinicultura. Os sistemas CRISPR corresponderam aos tipos I-E, I-F e III-D. 53 % dos espaçadores (75/142) mostraram homologia com plasmídeos, enquanto os restantes 47 % (67/142) mostraram homologia com bacteriófagos, principalmente vírus infecciosos *Vibrio* não típicos. A utilização da terapia fágica é proposta como tratamento para infecções causadas por membros da família Vibrionaceae em cultivo de camarão devido à baixa ocorrência de sistemas CRISPR nas espécies estudadas e à baixa imunidade aos seus fagos, garantindo assim maior sensibilidade.

**Palavras chave:** sistemas CRISPR-Cas, *Vibrio*, fagoterapia, análise bioinformática.

## Introduction

In shrimp culture, the family of proteobacteria Vibrionaceae, those of the genus *Vibrio*, represent one of the main responsible for the infections with a high influence on larval production associated with shrimp culture (*Litopenaeus vannamei*). These Gram-negative bacteria are ubiquitous and widely distributed in aquatic environments, from brackish water to deep sea water, in association with marine animals, algae and detritus. *Vibrios* are also present in estuarine and marine

## Introducción

En el cultivo de camarón (*Litopenaeus vannamei*), la familia de proteobacterias Vibrionaceae, el género *Vibrio*, representa uno de los principales responsables de las infecciones con alta influencia en la producción larvaria asociadas al cultivo de camarón. Estas bacterias Gram-negativas son ubicuas y están ampliamente distribuidas en ambientes acuáticos, desde agua salobre hasta aguas profundas, en asociación con animales marinos, algas y detritos. Los *Vibrios* también

aquatic ecosystems where shrimp are naturally found and/or cultivated. Bacteria of this genus are considered part of the normal microbiota of shrimp, constituting the highest percentage of all bacteria isolated from the digestive tract, gills, cuticle and occasionally in the hemolymph. However, under conditions of stress or imbalance in the natural bacterial microbiota these bacteria can induce the development of infections in organisms such as vibriosis (or bacterial shrimp septicemia) (Cheng *et al.*, 2008; Aguirre *et al.*, 2010; Yooseph *et al.*, 2010).

Of the genus *Vibrio*, the main virulent strains of shrimp identified in various larval stages are *V. alginolyticus*, *V. anguillarum*, *V. parahaemolyticus*, *V. harveyi* and *V. vulnificus* (Kalatzis *et al.*, 2018), generating high mortalities, economic losses in the production; they also cause public health problems in humans due to the consumption of derived marine products contaminated with these opportunistic pathogens (Cheng *et al.*, 2008; Won and Park, 2008; Aguirre *et al.*, 2010; Alagappan *et al.*, 2010). The prevention and control of diseases in shrimp farming is based primarily on the use of antibiotics (Lomelí and Martínez, 2014), however, the growing emergence of bacterial strains resistant to antibiotics, added to the concern induced by the prospect of not having effective antibiotics in the near future, has prompted the search for new antagonistic alternatives toward those that have not yet developed mechanisms of resistance, with fewer side effects, and with high

están presentes en ecosistemas acuáticos estuarinos y marinos donde los camarones se encuentran y/o cultivan naturalmente. Las bacterias de este género se consideran parte de la microbiota normal del camarón, constituyendo el porcentaje más alto de todas las bacterias aisladas del tracto digestivo, branquias, cutículas y ocasionalmente en la hemolinfa. Sin embargo, en condiciones de estrés o desequilibrio en la microbiota bacteriana natural, estas bacterias pueden inducir el desarrollo de infecciones en organismos como la vibriosis (o septicemia bacteriana del camarón) (Cheng *et al.*, 2008; Aguirre *et al.*, 2010; Yooseph *et al.*, 2010).

Del género *Vibrio*, las principales cepas virulentas de camarón identificadas en varios estadios larvales son *V. alginolyticus*, *V. anguillarum*, *V. parahaemolyticus*, *V. harveyi* y *V. vulnificus* (Kalatzis *et al.*, 2018), generando altas mortalidades, pérdidas económicas en la producción; también causan problemas de salud pública en humanos debido al consumo de productos marinos derivados contaminados con estos patógenos oportunistas (Cheng *et al.*, 2008; Won y Park, 2008; Aguirre *et al.*, 2010; Alagappan *et al.*, 2010). La prevención y control de enfermedades en la camaronicultura se basa principalmente en el uso de antibióticos (Lomelí y Martínez, 2014), sin embargo, la creciente aparición de cepas bacterianas resistentes a los antibióticos, sumado a la preocupación inducida por la perspectiva de no contar con antibióticos efectivos. en un futuro próximo, ha impulsado la

specificity, to avoid the disturbance of the intestinal microbiota of the human being and guarantee a greater effectiveness of the treatments (Rice and Stokes, 2008).

One of the promising strategies on the way to the objective of finding more effective, less toxic and specific antagonists to control microorganisms that cause infections, is represented by the viruses that exclusively infect the bacteria, called bacteriophages, and their enzymatic derivatives (enzymiotics). These are inhibitory agents that have been described as potential antimicrobial agents with therapeutic interest (São, 2018), including their application in shrimp farming (Oliveira *et al.*, 2012; Kalatzis *et al.*, 2018).

Although it is not clear if there is any mechanism of bacterial resistance against enzymiotics (phage lysins), studies on the use of phages as an alternative treatment or biological control show the opposite, and currently a large number of investigations are carried out in group of genomic sequences present in several prokaryotic microorganisms known as “Clustered Regularly Interspaced Short Palindromic Repeats” (CRISPR), which have shown a distinctive characteristic of most genomes of bacteria and archaea, that endows resistance against bacteriophages (Horvath and Barrangou, 2010; Kalatzis *et al.*, 2018).

In addition to the fact that research on the presence of CRISPRs has focused mainly on

búsqueda de nuevas alternativas antagónicas hacia aquellas que aún no han desarrollado mecanismos de resistencia, con menores efectos secundarios, y con alta especificidad, para evitar la alteración de la microbiota intestinal del ser humano y garantizar una mayor efectividad de los tratamientos (Rice y Stokes, 2008).

Una de las estrategias prometedoras en el camino hacia el objetivo de encontrar antagonistas más eficaces, menos tóxicos y específicos para el control de los microorganismos causantes de infecciones, está representada por los virus que infectan exclusivamente a las bacterias, llamados bacteriófagos, y sus derivados enzimáticos (enzimbióticos). Se trata de agentes inhibidores que han sido descritos como potenciales agentes antimicrobianos con interés terapéutico (São, 2018), incluida su aplicación en el cultivo de camarón (Oliveira *et al.*, 2012; Kalatzis *et al.*, 2018).

Aunque no está claro si existe algún mecanismo de resistencia bacteriana frente a enzibióticos (fago lisinas), los estudios sobre el uso de fagos como tratamiento alternativo o control biológico muestran lo contrario, y actualmente se realizan un gran número de investigaciones en grupo de secuencias genómicas presentes en varios microorganismos procarióticos conocidos como “Repeticiones palindrómicas cortas agrupadas regularmente interespaciadas” (CRISPR), que han mostrado una distintiva característica en la mayoría de los genomas de bacterias y arqueas, que confieren resistencia contra

bacteria of clinical interest, while little is known about the dynamics of CRISPR in organisms of commercial importance (Lyons *et al.*, 2015), the objective of this study, was to perform a comparative analysis of CRISPR Systems in genomes of microorganisms with pathogenic influence on shrimp *Litopenaeus vannamei*.

## Materials and methods

### Genomic sequences and identification of CRISPR structures

Five species of *Vibrio* were studied: *Vibrio alginolyticus*, *V. anguillarum*, *V. parahaemolyticus*, *V. harveyi*, *V. vulnificus* and the species *Photobacterium damsela*, all associated with systemic infections in shrimp, such as vibriosis, of which a total of 1104 genomes were obtained according to the database of the NCBI (National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/>, Benson *et al.*, 2008). Draft genome sequences were obtained from specific web sites that are compiled in the Entrez Genome project list (<http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi>) or in the Genomes OnLine Database (<http://www.genomesonline.org/>) (Bernal *et al.*, 2001), 69 of the sequences correspond to complete genomic sequences including 14 *V. alginolyticus*, 13 *V. anguillarum*, 21 *V. parahaemolyticus*, 4 *V. harveyi*, 15 *V. vulnificus*, as well as 2 complete chromosomal sequences of *Photobacterium damsela* (Table 1).

bacteriófagos (Horvath y Barrangou, 2010; Kalatzis *et al.*, 2018).

Además de que la investigación sobre la presencia de CRISPRs se ha centrado principalmente en bacterias de interés clínico, mientras que se conoce poco sobre la dinámica de CRISPR en organismos de importancia comercial (Lyons *et al.*, 2015), el objetivo de este estudio, consistió en realizar un análisis comparativo de Sistemas CRISPR en genomas de microorganismos con influencia patógena en camarón *Litopenaeus vannamei*.

## Materiales y métodos

### Secuencias genómicas e identificación de estructuras CRISPR

Se estudiaron cinco especies de *Vibrio*: *Vibrio alginolyticus*, *V. anguillarum*, *V. parahaemolyticus*, *V. harveyi*, *V. vulnificus* y la especie *Photobacterium damsela*, todas asociadas a infecciones sistémicas en camarones, como la vibriosis, de las cuales un total de 1104 genomas se obtuvieron de acuerdo con la base de datos del NCBI (Centro Nacional de Información Biotecnológica, <http://www.ncbi.nlm.nih.gov/>, Benson *et al.*, 2008). Las secuencias preliminares del genoma se obtuvieron de sitios web específicos que se compilan en la lista de proyectos Entrez Genome (<http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi>) o en la base de datos Genomes OnLine (<http://www.genomesonline.org/>) (Bernal *et al.*, 2001), 69 de las secuencias corresponden a secuencias genómicas completas, que incluyen 14

Table 1. CRISPR/Cas systems in Vibrionaceae genomes.

Cuadro 1. Sistemas CRISPR / Cas en genomas de Vibrionaceae.

Species	Accession number GenBank	CRISPR loci							CRISPR Type	Cas gene	Ref.
		Length of CRISPR (pb)	Start	End	Spacer number	Length of Spacer (pb)	DR number	Length of DR (pb)			
<i>Vp.</i> FORC_022	CP013248.1 / P013249.1	1707	826587	828284	28	32	29	28	GTTAACCTGCCACACAGCGAGCTTAGAAA	I-F	6
	<i>Vh.</i> ATCC 43516	2287	73577	75664	37	32-33	38	29	GTTGTTCCGTTACCCACGGAGATGAACCG	I-E	8
<i>Vh.</i> YJ016	BA000057.2 / A000058.2	661	1694586	1695247	9	33-37	10	35	GTTTCAGACATGCCCGGTTTAGACGG-GATTAAGAC	III-D	7
	<i>Vh.</i> 93U204	3328	511712	515040	55	32	56	28	GTTCACTGCCGTATAGGCGAGCTTAGAAA	I-F	6
<i>Pd.</i> KC-Na-1*	CP021151.1 / P021152.1	256	2123187	2123443	3	43	4	32	GATATTTCTAACTGGGATACTTCCAAT-GTAAA		Chem <i>et al.</i> , 2003
		447	246163	246610	7	32	8	28	CITTCACCTGCGAGTAGGCGAGCTTAGAAA	I-F	6
		29	257070	257278	3	31	4	29	CITTCACCTGCGAGTAGGCGAGCTTAGAAAT		Lo <i>et al.</i> , 2014

*Vp.* FORC\_022, *Vh.* ATCC 43516, *Vh.* YJ016 y 93U204, *Pd.* KC-Na-1, are the abbreviations of *V. parahaemolyticus* FORC\_022, *V. harveyi* ATCC 43516, *V. vulnificus* YJ016 y 93U204, and *P. damselae* KC-Na-1, respectively; \*, before classified in *V. damselae*; CRISPR\_C, clustered Regularly Interspaced Short Palindromic Repeats; DR, direct repeats; Cas, CRISPR-associated (*cas*) gene cluster.

*Vp.* FORC\_022, *Vh.* ATCC 43516, *Vh.* YJ016 y 93U204, *Pd.* KC-Na-1, son las abreviaturas de *V. parahaemolyticus* FORC\_022, *V. harveyi* ATCC 43516, *V. vulnificus* YJ016 y 93U204, y *P. damselae* KC-Na-1, respectivamente; \*, antes clasificado en *V. damselae*; CRISPR, repeticiones palindrómicas cortas agrupadas regularmente interespaciadas; DR, repeticiones directas; Cas, grupo de genes asociados a CRISPR (*cas*).



For published genome sequences, CRISPR loci were obtained from the CRISPRdb database (Grissa *et al.*, 2007a). Alternatively, the CRISPR loci in the genomes were identified by CRISPRFinder (Grissa *et al.*, 2007b). For this, two search criteria were followed. One of them was the screening of possible CRISPR locations by detecting maximum repetitions (repetition with the maximum possible extension to the right or left without incurring a mismatch) by using the VMatch package, which is the REPuter update (Kurtz and Schleiermacher, 1999) based on an efficient implementation of improved suffix arrays (Abouelhoda, 2004). The default parameters used were the following: a repetition length of 23 to 55 bp, a gap size between repetitions of 25 to 60 bp, a 20% mismatch of nucleotides between repetitions. The other search criterion was based on the CRISPR function determination, for which filters are added to help validate a CRISPR, such as the search for non-identical spacers with a size that should be from 0,6\* to 2,5\* the size of the repetition. This filter is configured to eliminate tandem repeats. The comparison of the spacers was made by aligning them (using the default parameters of the Muscle program). The percentage of similarity of the spacers is calculated with the percentage\_identity function of the (Bio)perl interface (AlignIO methods, Muscle interface), a parameter set to 60 %.

To discriminate between confirmed CRISPR structures from those questionable, small structures similar to CRISPR, that is, having only two

*V. alginolyticus*, 13 *V. anguillarum*, 21 *V. parahaemolyticus*, 4 *V. harveyi*, 15 *V. vulnificus*, así como 2 secuencias cromosómicas completas de *Photobacterium damsela* (Cuadro 1).

Las secuencias genómicas públicas de los loci CRISPR se obtuvieron de la base de datos CRISPRdb (Grissa *et al.*, 2007a). Alternativamente, los loci CRISPR en los genomas fueron identificados por CRISPRFinder (Grissa *et al.*, 2007b). Para ello se siguieron dos criterios de búsqueda. Uno de ellos fue el cribado de posibles ubicaciones de CRISPR mediante la detección de repeticiones máximas (repetición con la máxima extensión posible a la derecha o izquierda sin incurrir en un desajuste) mediante el uso del paquete VMatch, que es la actualización de REPuter (Kurtz y Schleiermacher, 1999) basada en una implementación eficiente de arreglos de sufijos mejorados (Abouelhoda, 2004). Los parámetros predeterminados utilizados fueron los siguientes: una longitud de repetición de 23 a 55 pb, un tamaño de espacio entre repeticiones de 25 a 60 pb, un 20% de desajuste de nucleótidos entre repeticiones. El otro criterio de búsqueda se basó en la determinación de la función CRISPR, para lo cual se agregan filtros para ayudar a validar un CRISPR, como la búsqueda de espaciadores no idénticos con un tamaño que debe ser de 0,6 \* a 2,5 \* el tamaño de la repetición. Este filtro está configurado para eliminar repeticiones en tándem. La comparación de los espaciadores se realizó alineándolos (utilizando los parámetros predeterminados del



or three repeated sequences or direct repeats (DR), are classified using a level of Evidence, rated from 1 to 4, where 1 includes small CRISPR (with 3 or fewer spacers) and 2 to 4 are classified based on the repetition and similarity of the spacer. This is because questionable CRISPR structures often tend to be byproducts of the CRISPRFinder identification that are not true CRISPR. Further, the tests to verify the internal preservation of the candidate repetitions and the divergence of the candidate spacers offered by CRISPRFinder are considered.

### **Bioinformatic analysis and statistical results**

For the search of genes *cas*, the first step consisted in the identification of open reading frames (ORF) with Prodigal (Hyatt, 2010). Then these ORFs were analyzed by the MacSyFinder program by the models for the search of HMM genes (Hidden Markov Models) in a library of known Cas proteins (Abby, 2014). BLAST was used to identify the *cas* genes in the upstream and downstream sequences of the CRISPR and TIGRFAM loci (Horvath *et al.*, 2009). The *cas* type and subtype were found through the analysis of *cas* conglomerates, thanks to the CRISPRCas-Finder program (Grissa *et al.*, 2007b). The unique spacer sequences as well as their origin were identified by NCBI's multiple sequence alignment program with predetermined arguments. All spacers were compared with the GenBank database to find homologous sequences with matches of  $\geq 85\%$  (minimum of 28/33 coincident nucleotides) (Horvath *et al.*, 2009; Shen *et al.*, 2017).

programa Muscle). El porcentaje de similitud de los espaciadores se calcula con la función percent\_identity de la interfaz (Bio) perl (métodos AlignIO, interfaz Muscle), un parámetro establecido en 60%.

Para discriminar entre estructuras CRISPR confirmadas de aquellas cuestionables, las estructuras pequeñas similares a CRISPR, es decir, que tienen solo dos o tres secuencias repetidas o repeticiones directas (DR), se clasifican utilizando un nivel de evidencia, calificado de 1 a 4, donde 1 incluye pequeños CRISPR (con 3 o menos espaciadores) y 2 a 4 se clasifican en función de la repetición y similitud del espaciador. Esto se debe a que las estructuras CRISPR cuestionables a menudo tienden a ser subproductos de la identificación de CRISPRFinder que no son CRISPR verdadero. Además, se consideran las pruebas para verificar la preservación interna de las repeticiones candidatas y la divergencia de los espaciadores candidatos ofrecidos por CRISPRFinder.

### **Análisis bioinformático y resultados estadísticos**

Para la búsqueda de genes *cas*, el primer paso consistió en la identificación de marcos de lectura abiertos (ORF) con Prodigal (Hyatt, 2010). Luego, estos ORF fueron analizados por el programa MacSyFinder mediante los modelos para la búsqueda de genes HMM (Hidden Markov Models) en una biblioteca de proteínas Cas conocidas (Abby, 2014). BLAST se usó para identificar los genes *cas* en las secuencias aguas arriba y aguas

Phylogenetic trees were generated based on the unweighted pair group method (UPGMA) for the core protein Cas, specifically Cas1 using the MUSCLE algorithm (Tamura *et al.*, 2013). In parallel, multiple sequence alignments and phylogenetic analyzes were performed using Clustal X, and dendograms were visualized with the accompanying application NJ Plot (Larkin *et al.*, 2007). The distance matrix was calculated using the Jaccard coefficient. The structure of Cas proteins was evaluated by multiple sequence alignments with Geneious global alignment (Needleman-Wunsch) with predetermined arguments (Shen *et al.*, 2017). The PubMLST database was used for the determination of the multilocus sequences and the sequence typing data were used for the goeBURST analysis (Shen *et al.*, 2017). The one-factor analysis of variance (ANOVA) was used as a statistical model to establish differences between the comparative parameters analyzed.

## Results and discussion

The present study aims to offer an analysis of CRISPR/Cas gene sequences in microorganisms of high influence in aquaculture production associated with shrimp infections, specifically in the species of commercial interest *L. vannamei*, intending to offer alternatives for the control of pathogens of bacterial origin, through phagotherapy, a promising technology for shrimp farming based on the management of specific lytic phages of bacteria or their lytic enzymes (enzymotics).

abajo de los loci CRISPR y TIGRFAM (Horvath *et al.*, 2009). El tipo y subtipo *cas* se encontraron mediante el análisis de conglomerados *cas*, gracias al programa CRISPRCas-Finder (Grissa *et al.*, 2007b). Las secuencias espaciadoras únicas, así como su origen, fueron identificadas por el programa de alineación de secuencias múltiples del NCBI con argumentos predeterminados. Todos los espaciadores se compararon con la base de datos GenBank para encontrar secuencias homólogas con coincidencias de  $\geq 85\%$  (mínimo de 28/33 nucleótidos coincidentes) (Horvath *et al.*, 2009; Shen *et al.*, 2017).

Los árboles filogenéticos se generaron con base en el método de grupos de pares no ponderados (UPGMA) para la proteína central Cas, específicamente Cas1, utilizando el algoritmo MUSCLE (Tamura *et al.*, 2013). En paralelo, se realizaron múltiples alineamientos de secuencia y análisis filogenéticos utilizando Clustal X, y los dendogramas se visualizaron con la aplicación adjunta NJ Plot (Larkin *et al.*, 2007). La matriz de distancias se calculó mediante el coeficiente de Jaccard. La estructura de las proteínas Cas se evaluó mediante múltiples alineamientos de secuencia con alineamiento global Geneious (Needleman-Wunsch) con argumentos predeterminados (Shen *et al.*, 2017). La base de datos PubMLST se utilizó para la determinación de las secuencias multilocus y los datos de tipificación de secuencias se utilizaron para el análisis goeBURST (Shen *et al.*, 2017). El análisis de varianza de

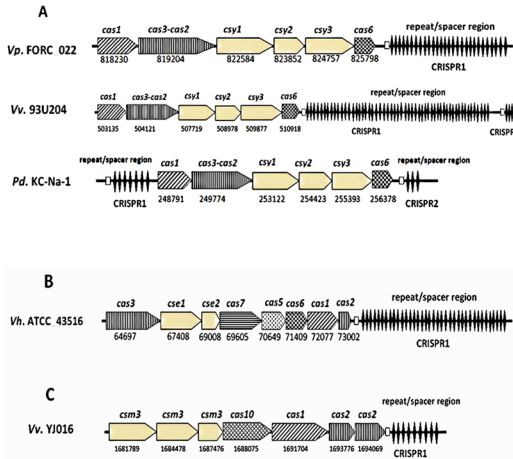
**Organization and diversity of CRISPR/Cas systems in Vibrionaceae genome**

Seven CRISPR structures were found in 5 genomes within the total (1104 genomes) of the family of proteobacteria Vibrionaceae studied, distributed in 3 *Vibrio* species (*V. parahaemolyticus* FORC\_022, *V. harveyi* ATCC 43516, *V. vulnificus* YJ016 and 93U204) and in the related species *Photobacterium damselae* KC-Na-1 (formerly classified as *V. damselae*) whose characteristics are described in Table 1. Three different CRISPR/Cas arrays in *Vibrio* that have different locations on the chromosomes were detected (Figure 1).

un factor (ANOVA) se utilizó como modelo estadístico para establecer diferencias entre los parámetros comparativos analizados.

**Resultados y discusión**

El presente estudio tiene como objetivo ofrecer un análisis de secuencias de genes CRISPR/Cas en microorganismos de alta influencia en la producción acuícola asociados a infecciones del camarón, específicamente en la especie de interés comercial *L.vannamei*, con la intención de ofrecer alternativas para el control de patógenos de origen



**Figure 1. Architecture loci and gene organization of CRISPR systems in three different regions.** A, type I-F; B, type I-E; C, type III-D. The black diamonds represent the CRISPR loci, and the white arrows (with lines in different directions) represent the different cas genes found. The relative position on the chromosome is indicated below each gene.

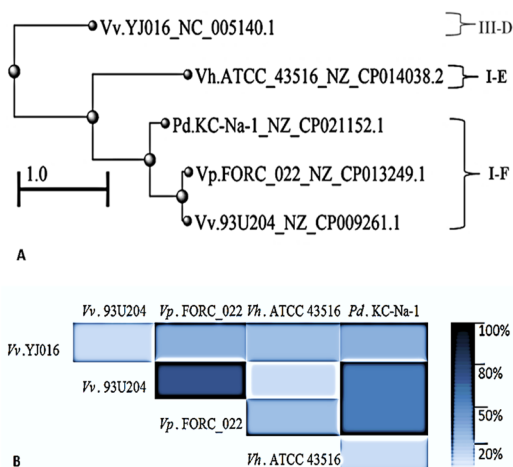
**Figura 1. Loci de la Arquitectura y Organización genética de sistemas CRISPR en tres regiones diferentes.** A, tipo I-F; B, tipo I-E; C, tipo III-D. Los diamantes negros representan los loci CRISPR y las flechas blancas (con líneas en diferentes direcciones) representan los diferentes genes cas encontrados. La posición relativa en el cromosoma se indica debajo de cada gen.

The types of CRISPR/Cas systems found correspond to types I-E, I-F and III-D. According to the relative location on the chromosome, the CRISPR loci were designated CRISPR1 and CRISPR2. To better distinguish the CRISPR/Cas system in the genus *Vibrio*, or in the members with CRISPR matrices found in the Vibrionaceae family, a phylogenetic tree was constructed based on the homology of the Cas1 proteins of each species (Figure 2).

bacteriano., a través de la fagoterapia, una tecnología prometedora para el cultivo de camarón basada en el manejo de fagos líticos específicos de bacterias o sus enzimas líticas (enzimbióticos).

### Organización y diversidad de sistemas CRISPR / Cas en el genoma de Vibrionaceae

Se encontraron siete estructuras CRISPR en 5 genomas dentro del total (1104 genomas) de la familia



**Figure 2. Phylogenetic tree for Cas1 proteins in the family Vibrionaceae.** The UPGMA tree of the Cas1 protein was generated using the MUSCLE algorithm in MEGA6. The Cas1 proteins representative of the subtypes found were selected (I-F, I-E and III-D) (A). Percentage identity of Cas1 proteins in the Vibrionaceae genomes studied. (B). Vp. FORC\_022, Vh. ATCC 43516, Vv. YJ016 and 93U204, Pd. KC-Na-1, are the abbreviations of *V. parahaemolyticus* FORC\_022, *V. harveyi* ATCC 43516, *V. vulnificus* YJ016 and 93U204, and *P. damsela* KC-Na-1, respectively.

**Figura 2. Árbol filogenético de las proteínas Cas1 de la familia Vibrionaceae.** El árbol UPGMA de la proteína Cas1 se generó utilizando el algoritmo MUSCLE en MEGA6. Se seleccionaron las proteínas Cas1 representativas de los subtipos encontrados (I-F, I-E y III-D) (A). Porcentaje de identidad de las proteínas Cas1 en los genomas de Vibrionaceae estudiados. (B). Vp. FORC\_022, Vh. ATCC 43516, Vv. YJ016 y 93U204, Pd. KC-Na-1, son las abreviaturas de *V. parahaemolyticus* FORC\_022, *V. harveyi* ATCC 43516, *V. vulnificus* YJ016 y 93U204, y *P. damsela* KC-Na-1, respectivamente.

However, it is important to note that since few confirmed CRISPR sequences have been obtained in the genomes of interest to date, this may make it difficult to establish a more adequate phylogenetic relationship, so our results represent only a limited approximation of the relationship based only on the conservation of Cas1, and can be seen as an internal control of the previous classification performed by the CRISPRdb database and the CRISPRFinder (Grissa *et al.*, 2007a, b). We recommend performing a phylogenetic analysis with a greater number of genomic sequences (including partial genomes) and that include CRISPR systems with different levels of confidence to obtain a better discriminatory phylogenetic relationship. The distinction between the type of CRISPR found by the analysis of direct repeat (DR) sequence was also evaluated (Figure 3). Typical CRISPR matrices were observed with a minimum of 4 to a maximum of 56 DR, which had a length between 28-35 bp. The DR were therefore found to be interspaced, with a minimum of 3 to a maximum of 55 spacer sequences of a similar length between 31-43 bp. Despite the fact that most of the spacer sequences presented a size of 32 bp, the analysis of variance indicates that there is no statistically significant difference ( $p > 0.01$ ) that allows to point out that the obtained means corresponding to the number and size of the spacer sequences are different from each other.

de proteobacterias Vibrionaceae estudiadas, distribuidas en 3 especies de *Vibrio* (*V. parahaemolyticus* FORC\_022, *V. harveyi* ATCC 43516, *V. vulnificus* YJ016 y 93U204) y en la especie relacionada *Photobacterium damsela* KC-Na-1 (anteriormente clasificada como *V. damsela*) cuyas características se describen en la Tabla 1. Se detectaron tres matrices CRISPR/Cas diferentes en *Vibrio* que tienen diferentes ubicaciones en los cromosomas (Figura 1).

Los tipos de sistemas CRISPR/Cas encontrados corresponden a los tipos I-E, I-F y III-D. Según la ubicación relativa en el cromosoma, los loci CRISPR se denominaron CRISPR1 y CRISPR2. Para distinguir mejor el sistema CRISPR/Cas en el género *Vibrio*, o en los miembros con matrices CRISPR encontradas en la familia Vibrionaceae, se construyó un árbol filogenético basado en la homología de las proteínas Cas1 de cada especie (Figura 2).

Sin embargo, es importante señalar que dado que hasta la fecha se han obtenido pocas secuencias CRISPR confirmadas en los genomas de interés, esto puede dificultar el establecimiento de una relación filogenética más adecuada, por lo que nuestros resultados representan solo una aproximación limitada de la relación basada únicamente sobre la conservación de Cas1, y puede verse como un control interno de la clasificación anterior realizada por la base de datos CRISPRdb y el CRISPRFinder (Grissa *et al.*, 2007a, b). Recomendamos realizar un análisis filogenético con un mayor

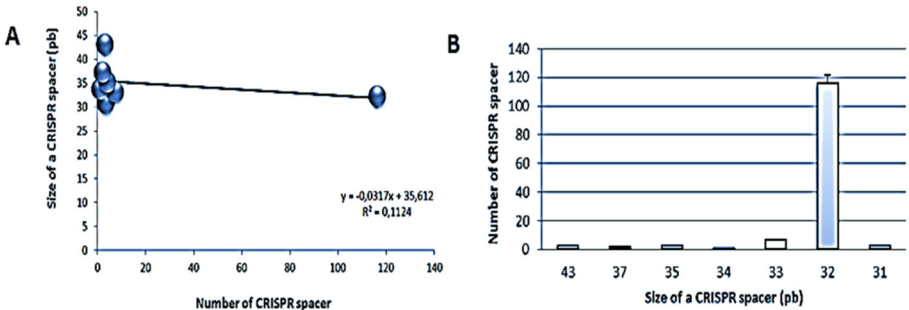




the genes encoding the Cas1, Cas3 proteins have been identified and Cas6, and based on the sequences of proteins and their organization, have been grouped in CRISPR/Cas systems of type I-F (Labbate *et al.*, 2016), as well as in clinical isolates of the species *V. parahaemolyticus* in which the I-F CRISPR/Cas subtype has been described (Sun *et al.*, 2015). Likewise, the presence of I-E subtypes has been reported in strains of the genus such as *V. cholerae* O395, in which 3 *cas* core genes (*cas1*, *cas2* and *cas3*) and 5 specific genes of the I-E subtype (*cse1*, *cse2*, *cse3*, *cse4*, *cas5e*) have been identified (Chakraborty *et al.*, 2009; Box *et al.*, 2016).

similar entre 31 y 43 pb. A pesar de que la mayoría de las secuencias espaciadoras presentaron un tamaño de 32 pb, el análisis de varianza indica que no existe una diferencia estadísticamente significativa ( $p > 0.01$ ) que permita señalar que las medias obtenidas correspondientes al número y tamaño de las secuencias espaciadoras son diferentes entre sí.

Sin embargo, existe una correlación muy baja entre el número y el tamaño que presentan estas secuencias (Figura 4). El número total de espaciadores detectados por tipo de CRISPR fue 96/142 para los tipos I-F (68%), 37/142 para el tipo I-E (26%) y 9/142 para el tipo III-D (6%).



**Figure 4. Correlation of the number of spacers in the CRISPR loci and the size of the CRISPR (bp) spacers found.** A very low negative correlation was found ( $\leq 30\%$ ) (A). Variability of the size of the CRISPR spacers and their number ( $p > 0.05$ ). The x-axis represents the size of a CRISPR spacer, in nucleotides. The y-axis represents the number of CRISPR spacing sequences of a given size (B).

**Figura 4. Correlación del número de espaciadores en los loci CRISPR y el tamaño de los espaciadores CRISPR (bp) encontrados.** Se encontró una correlación negativa muy baja ( $\leq 30\%$ ) (A). Variabilidad del tamaño de los espaciadores CRISPR y su número ( $p > 0,05$ ). El eje X representa el tamaño de un espaciador CRISPR, en nucleótidos. El eje Y representa el número de secuencias de espaciado CRISPR de un tamaño dado (B).



On the other hand, it has been reported that CRISPR Type I-E systems always have two loci pairs, at least that is what has been evidenced in enterobacteria genomes (Shen *et al.*, 2017). However, our results differ from the aforementioned, since in the species *V. harveyi* ATCC 43516 only, a confirmed CRISPR structure was detected, such as that described in strains such as *V. cholerae* O395 (Chakraborty *et al.*, 2009).

Interestingly, a new form of disposition of *cas* genes was found for the Type III-D subtype detected in *V. vulnificus* YJ016. Although according to the current classification model (Makarova *et al.*, 2015), it should still be a CRISPR/Cas type III-D system because the *cas* genes found are within the subtype arrangement III-D, however, a pair of *cas2* genes was detected as well as the *cas1* gene at chromosomal location 1693776-1694069 bp and 1691704 bp, respectively. This does not correspond to the genes of subtype III-D, and rather, they remind us of the arrangement of the subtype III-A architectures, which possess both the *cas10* and *cas3* genes of type III-D, and a *cas* nucleus represented by the *cas1* and *cas2* (Makarova *et al.*, 2015), therefore it is in the presence of two very closely related subtypes, and with a possible common origin.

### **The origin of CRISPR spacers**

Of the 130/142 (92 %) unique spacer sequences found in the CRISPR loci detected, it is important to note that all had homology with some sequence contained in the GenBank database, as well as a great diversity

Las estructuras CRISPR observadas exhibieron entre 6 y 8 genes asociados con las secuencias CRISPR (*cas*) (Cuadro 1).

La presencia de subtipos I-F corresponde a lo descrito en la literatura en cepas del género como *V. cholerae* O1/O139, en las que, tras análisis bioinformático, se han estudiado módulos CRISPR/Cas que contienen los genes que codifican las proteínas Cas1, Cas3, Cas4, Cas5, Cas6, y en base a las secuencias de proteínas y su organización, se han agrupado en sistemas CRISPR/Cas de tipo IF (Labbate *et al.*, 2016), como en aislados clínicos de la especie *V. parahaemolyticus* en la que el subtipo CRISPR/Cas I-F si se ha descrito (Sun *et al.*, 2015). Asimismo, se ha reportado la presencia de subtipos de I-E en cepas del género como *V. cholerae* O395, en el que 3 genes del núcleo *cas* (*cas1*, *cas2* y *cas3*) y 5 genes específicos del subtipo I-E (*cse1*, *cse2*, *cse3*, *cse4*, *cas5e*) han sido identificados (Chakraborty *et al.*, 2009; Box *et al.*, 2016).

Por otro lado, se ha reportado que los sistemas CRISPR tipo I-E siempre tienen dos pares de loci, al menos eso es lo que se ha evidenciado en genomas de enterobacterias (Shen *et al.*, 2017). Sin embargo, nuestros resultados difieren de los anteriores, ya que en la especie *V. harveyi* ATCC 43516 únicamente se detectó una estructura CRISPR confirmada, como la descrita en cepas como *V. cholerae* O395 (Chakraborty *et al.*, 2009).

Curiosamente, se encontró una nueva forma de disposición de genes *cas* para el subtipo III-D detectado

in relation to the presented homology to related or not sequences with the family of proteobacteria Vibrionaceae. 53 % of the spacers (75/142) presented homology with extrachromosomal genetic material, while the remaining 47 % of the spacer sequences (67/142) exhibited homology with bacteriophages (Figure 5). Results that indicate a clear immunity function against foreign genetic material (plasmids or phages) and a high specificity of the different CRISPR/Cas systems studied. The type of CRISPR that gathered the greatest number of spacer sequences was the I-F type present in the species *P. damsela* KC-Na-1, *V. parahaemolyticus* FORC\_022 and *V. vulnificus* 93U204, who exhibited 10/96 (10.4 %), 28/96 (29.2 %) and 58/96 (60.4 %) spacer sequences respectively. In types I-E and III-D found in species *V. harveyi* ATCC 43516 and *V. vulnificus* YJ016 respectively, were observed 37 and 9 corresponding spacer sequences (Figure 5).

It is important to note that the spacer sequences found in the CRISPR type I-E were the ones that presented a greater homology to bacteriophage sequences with 76 % (28/37) followed by type III-D with 56 % (5/9) and in smaller proportion, the I-F type with 35 % (34/96). Interestingly, on the contrary, the spacers found in type I-F exhibited the highest homology to plasmid sequences with 65 % (62/96) (Figure 5). The homologies evidenced by the spacer sequences studied indicated a high diversity in terms of the recognition targets of the CRISPR

en *V. vulnificus* YJ016. Aunque de acuerdo con el modelo de clasificación actual (Makarova *et al.*, 2015), aún debería ser un sistema CRISPR/Cas tipo III-D porque los genes *cas* encontrados están dentro del arreglo de subtipos III-D, sin embargo, un par de genes *cas2* se detectó, así como el gen *cas1* en la ubicación cromosómica 1693776-1694069 bp y 1691704 bp, respectivamente. Esto no se corresponde con los genes del subtipo III-D, sino que nos recuerdan la disposición de las arquitecturas del subtipo III-A, que poseen los genes *cas10* y *csm3* del subtipo III-D, y un núcleo *cas* representado por el *cas1* y *cas2* (Makarova *et al.*, 2015), por lo que se encuentra en presencia de dos subtipos muy relacionados, y con un posible origen común.

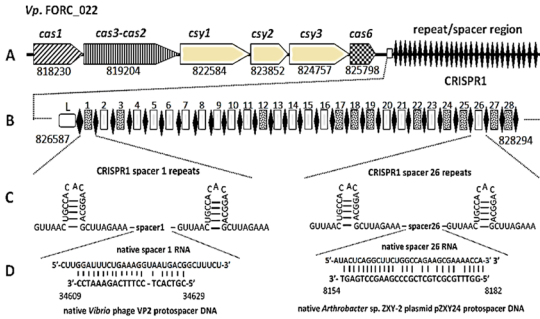
### El origen de los espaciadores CRISPR

De las 130/142 (92%) secuencias espaciadoras únicas encontradas en los loci CRISPR detectados, es importante señalar que todas tenían homología con alguna secuencia contenida en la base de datos GenBank, así como una gran diversidad en relación con la homología presentada con secuencias relacionadas o no con la familia de proteobacterias Vibrionaceae. El 53% de los espaciadores (75/142) presentaron homología con material genético extracromosómico, mientras que el 47% restante de las secuencias espaciadoras (67/142) exhibieron homología con bacteriófagos (Figura 5). Resultados que indican una clara función de inmunidad frente a material genético extraño (plásmidos o fagos) y una alta especificidad de

immunity systems examined, represented by the homology to DNA sequences related or not to representatives of the Vibrionaceae family. Specifically, there is little immunity to typical *Vibrio* infective bacteriophages (Figure 5) such as phage VP2, phage Douglas 12A4 and phage VSK (the latter is also a host of *P. damsela* KC-Na-1) and to plasmid sequences also present in strains of the genus *Vibrio* (Figure 5) such as p380 of *V. coralliilyticus* RE98, plasmid QT6D1 of *V. shilonii* strain, plasmid pMJ100 of *V. fischeri* MJ11 and plasmid pYJ016 of *V. vulnificus* YJ016 (spacers found at some of the CRISPR loci type I-F observed), or against phage of *Vibrio* 11895-B1 and plasmid p48/10 of *V. vulnificus* 48/10 (in the CRISPR system type I-E detected), or against phage of *Vibrio eugene* 12A10 and plasmid pSNUTY1 of *V. coralliilyticus* strain SNUTY-1 (in the CRISPR system type III-D studied). Like the immunity found in higher proportion against a great variety of infective phages of *Bacillus*, *Enterobacteria*, *Pseudomonas* and *Cyanobacteria* (results not shown). The results obtained show a low occurrence of CRISPR structures in these Vibrionaceae species, that provide them with immunity to lytic specific phages, especially of the *Vibrio* genus, which is important, considering that this sensitivity can be explored for the development of analytical tools such as phage therapy in shrimp culture, a technology based on the use of enzymes or phages for the control of typical shrimp pathogens.

los diferentes sistemas CRISPR/Cas estudiados. El tipo de CRISPR que reunió el mayor número de secuencias espaciadoras fue el tipo I-F presente en las especies *P. damsela* KC-Na-1, *V. parahaemolyticus* FORC\_022 y *V. vulnificus* 93U204, quienes exhibieron 10/96 (10,4%), 28/96 (29,2%) y 58/96 (60,4%) secuencias espaciadoras respectivamente. En los tipos I-E y III-D encontrados en las especies *V. harveyi* ATCC 43516 y *V. vulnificus* YJ016 respectivamente, se observaron 37 y 9 secuencias espaciadoras correspondientes (Figura 5).

Es importante señalar que las secuencias espaciadoras encontradas en el CRISPR tipo I-E fueron las que presentaron una mayor homología con las secuencias de bacteriófagos con 76% (28/37) seguidas del tipo III-D con 56% (5/9) y en menor proporción, el tipo I-F con 35% (34/96). Curiosamente, por el contrario, los espaciadores encontrados en el tipo I-F exhibieron la mayor homología con las secuencias de plásmido con 65% (62/96) (Figura 5). Las homologías evidenciadas por las secuencias espaciadoras estudiadas indicaron una alta diversidad en términos de las dianas de reconocimiento de los sistemas de inmunidad CRISPR examinados, representada por la homología con secuencias de DNA relacionadas o no con representantes de la familia Vibrionaceae. Específicamente, hay poca inmunidad a los bacteriófagos infecciosos típicos de *Vibrio* (Figura 5) como el fago VP2, el fago Douglas 12A4 y el fago VSK (este último también es un hospedador de *P. damsela* KC-Na-1) y a las secuencias de plásmidos



**Figure 5. The origin of CRISPR spacers.** General description of the CRISPR locus (A), Spacers acquired in phage or plasmid resistance (B). The CRISPR locus of *V. parahaemolyticus* FORC\_022 (Type I-F) and *V. vulnificus* YJ016 (Type III-D) is in the upper part. The region of spacers and repeats are in the middle: repeats (black diamonds), spacers (gray or dotted squares numbered, gray, for sequences with homology for plasmids and dotted, for homology with phage), leader (L, white box). (Below) the content of the spacer is detailed (C) and the sequences with homology to phages or plasmids (D). The global distribution of CRISPR spacers in plasmids and phages (E) and for the type (F) was determined by sequence identity. All spacers were compared against the GenBank database to find a sequence of homology.

**Figura 5. El origen de los espaciadores CRISPR.** Descripción general del locus CRISPR (A), Espaciadores adquiridos en fagos o resistencia a plásmidos (B). El locus CRISPR de *V. parahaemolyticus* FORC\_022 (Tipo I-F) y *V. vulnificus* YJ016 (Tipo III-D) se encuentra en la parte superior. La región de espaciadores y repeticiones está en el medio: repeticiones (rombos negros), espaciadores (cuadrados grises o punteados numerados, gris, para secuencias con homología para plásmidos y punteados, para homología con fagos), líder (L, recuadro blanco). (A continuación) se detalla el contenido del espaciador (C) y las secuencias con homología a fagos o plásmidos (D). La distribución global de espaciadores CRISPR en plásmidos y fagos (E) y para el tipo (F) se determinó por identidad de secuencia. Todos los espaciadores se compararon con la base de datos GenBank para encontrar una secuencia de homología.

The 92 % of the unique spacer sequences found in the CRISPR loci presented homology with some sequence contained in the GenBank database against extrachromosomal genetic material (plasmids) or bacteriophages, which is indicative of a clear immunity function against foreign genetic material and of a high specificity of the different CRISPR/Cas systems studied. It has been described

también presentes en cepas del género *Vibrio* (Figura 5) tales como p380 de *V. coralliilyticus* RE98, plásmido QT6D1 de la cepa de *V. shilonii*, plásmido pMJ100 de *V. fischeri* MJ11 y plásmido pYJ016 de *V. vulnificus* YJ016 (espaciadores encontrados en algunos de los loci CRISPR tipo I-F observado), o contra el fago de *Vibrio* 11895-B1 y el plásmido p48/10 de *V. vulnificus* 48/10 (en el sistema

that the CRISPR/Cas system provides immunity against the virus in prokaryotes, thanks to the fact that the spacers are obtained from invading elements and in this way the cell can mediate the immune reaction in a specific way by recognizing the homologous sequence of such a spacer (Barrangou and Horvath, 2017). Therefore, the profile of spacers can be a reflection of the bacterial lifestyle or of their habitat (Horvath *et al.*, 2009). In this study, it is important to note that most of the spacer sequences studied confer immunity to plasmids (53 %). Likewise, spacer sequences were found against a wide variety of infective phages from *Bacillus*, Enterobacteria, *Pseudomonas* and Cyanobacteria.

## Conclusion

The results obtained in this research are very promising because they show a low occurrence of CRISPR structures in these species Vibrionaceae with immunity to lytic-specific phages of the genus *Vibrio*, which is important if we consider that this sensitivity can be explored for the development of strategies such as phagotherapy in shrimp farming, a technology based on the use of enzymes or phages for the control of typical shrimp pathogens, reducing in this way the losses caused by infections of shrimp farms.

CRISPR el tipo I-E detectado), o contra el fago de *V. eugene* 12A10 y el plásmido pSNUTY1 de la cepa de *V. coralliilyticus* SNUTY-1 (en el sistema CRISPR tipo III-D estudiado). Como la inmunidad encontrada en mayor proporción contra una gran variedad de fagos infecciosos de *Bacillus*, Enterobacteria, *Pseudomonas* y Cyanobacteria (resultados no mostrados). Los resultados obtenidos muestran una baja ocurrencia de estructuras CRISPR en estas especies de Vibrionaceae, que las dotan de inmunidad frente a fagos líticos específicos, especialmente del género *Vibrio*, lo cual es importante, considerando que esta sensibilidad puede ser explorada para el desarrollo de herramientas analíticas como terapia de fagos en cultivo de camarón, una tecnología basada en el uso de enzimas o fagos para el control de patógenos típicos del camarón.

El 92% de las secuencias espaciadoras únicas encontradas en los loci CRISPR presentaron homología con alguna secuencia contenida en la base de datos GenBank contra material genético extracromosómico (plásmidos) o bacteriófagos, lo que es indicativo de una clara función de inmunidad contra material genético extraño y de una alta especificidad de los diferentes sistemas CRISPR/Cas estudiados. Se ha descrito que el sistema CRISPR/Cas proporciona inmunidad frente al virus en procariontas, gracias a que los espaciadores se obtienen a partir de elementos invasores y de esta forma la célula puede mediar la reacción inmune de forma específica al reconocer la secuencia homóloga. de

*End of English Version*

---

tal espaciador (Barrangou y Horvath, 2017). Por lo tanto, el perfil de los espaciadores puede ser un reflejo del estilo de vida bacteriano o de su hábitat (Horvath *et al.*, 2009). En este estudio, es importante señalar que la mayoría de las secuencias espaciadoras estudiadas confieren inmunidad a los plásmidos (53%). Asimismo, se encontraron secuencias espaciadoras contra una amplia variedad de fagos infecciosos de *Bacillus*, Enterobacteria, *Pseudomonas* y Cyanobacteria.

## Conclusión

Los resultados obtenidos en esta investigación son muy prometedores porque muestran una baja ocurrencia de estructuras CRISPR en estas especies Vibrionaceae con inmunidad a fagos líticos específicos del género *Vibrio*, lo cual es importante si consideramos que esta sensibilidad puede ser explorada para el desarrollo de estrategias como la fagoterapia en el cultivo de camarón, una tecnología basada en el uso de enzimas o fagos para el control de patógenos típicos del camarón, reduciendo de esta manera las pérdidas ocasionadas por infecciones de las granjas camaroneras.

## Cited literature

Abby, S., B. Néron., H. Ménager., M. Touchon and E. Rocha. 2014. MacSyFinder: a program to mine genomes for molecular systems with an application to CRISPR-Cas systems. *PloS one*. 9(10): e110726.

Abouelhoda, M., S. Kurtz and E. Ohlebusch. 2004. Replacing suffix trees with enhanced suffix arrays. *J. Discrete Algorithms*. 2(1): 53-86.

Aguirre, G., J. Sánchez, R. Pérez, A. Palacios, T. Trujillo and N. de la Cruz. 2010. Pathogenicity and infection route of *Vibrio parahaemolyticus* in american white shrimp, *Litopenaeus vannamei*. *J. World Aquac. Soc.* 41(3): 464-470.

Alagappan, K., D. Deivasigamani, S. Somasundaram and S. Kumaran. 2010. Occurrence of *Vibrio parahaemolyticus* and its specific phages from shrimp ponds in east coast of India. *Curr. Microbiol.* 61(4): 235-240.

Barrangou, R. and P. Horvath. 2017. A decade of discovery: CRISPR functions and applications. *Nat. Microbiol.* 2(7): 17092.

Benson, D., I. Karsch, D. Lipman., J. Ostell and D. Wheeler. 2008. GenBank. *Nucleic Acids Res.* 36(Database issue): D25.

Bernal, A., U. Ear and N. Kyrpides. 2001. Genomes OnLine Database (GOLD): a monitor of genome projects worldwide. *Nucleic Acids Res.* 29(1): 126-127.

Box, A., M. McGuffie., B. O'Hara and K. Seed. 2016. Functional analysis of bacteriophage immunity through a type I-E CRISPR/Cas system in *Vibrio cholerae* and its application in bacteriophage genome engineering. *J. Bacteriol.* 198(3): 578-590.

Chakraborty, S., T. Waise., F. Hassan., Y. Kabir., M. Smith and M. Arif. 2009. Assessment of the Evolutionary Origin and Possibility of CRISPR-Cas (CASS) Interference Pathway in *Vibrio cholerae* O395. *In silico biology*. 9(4): 245-254.

Cheng, L., J. Huang, C. Shi, K. Thompson, B. Mackey and J. Cai. 2008. *Vibrio parahaemolyticus* associated with mass mortality of postlarval abalone, *Haliotis diversicolor supertexta* (L.), in Sanya, China. *J. World Aquac. Soc.* 39(6): 746-757.

Grissa, I., G. Vergnaud and C. Pourcel. 2007a. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35 (suppl 2): W52-W57.



- Grissa, I., G. Vergnaud and C. Pourcel. 2007b. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*. 8(1):172.
- Horvath, P., A. Côté, D. Romero, P. Boyaval., C. Fremaux and R. Barrangou. 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int. J. Food. Microbiol.* 131(1): 62-70.
- Horvath, P. and R. Barrangou. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science*. 327(5962): 167-170.
- Hyatt, D., G. Chen., P. LoCascio., M. Land., F. Larimer and L. Hauser. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*. 11(1): 119.
- Kalatzis, P., D. Castillo, P. Katharios and M. Middelboe. 2018. Bacteriophage interactions with marine pathogenic *Vibrios*: Implications for phage therapy. *Antibiot.* 7(1): 15.
- Kurtz, S and C. Schleiermacher. 1999. REPuter: fast computation of maximal repeats in complete genomes. *Bioinformatics (Oxford, England)*, 15(5): 426-427.
- Labbate, M., F. Orata., N. Petty., N. Jayatilake., W. King., P. Kirchberger and Y. Boucher. 2016. A genomic island in *Vibrio cholerae* with VPI-1 site-specific recombination characteristics contains CRISPR-Cas and type VI secretion modules. *Sci. Rep.* 6: 36891.
- Larkin, M., G. Blackshields., N. Brown., R. Chenna., P. McGettigan., H. McWilliam and J. Thompson, J. D. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*. 23(21): 2947-2948.
- Lomelí, C. and S. Martínez. 2014. Phage therapy against *Vibrio parahaemolyticus* infection in the whiteleg shrimp (*Litopenaeus vannamei*) larvae. *Aquaculture*. 434: 208-211.
- Lyons, C., N. Raustad., M. Bustos and M. Shiaris. 2015. Incidence of Type II CRISPR1-Cas Systems in *Enterococcus* Is Species-Dependent. *PLoS one*. 10(11): e0143544.
- Makarova, K., Y. Wolf., O. Alkhnbashi., F. Costa., S. Shah., S. Saunders and P. Horvath. 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13(11): 722.
- Oliveira, J., F. Castilho., A. Cunha and M. Pereira. 2012. Bacteriophage therapy as a bacterial control strategy in aquaculture. *Aquacult. Int.* 20(5): 879-910.
- Rice, L. and Stokes, L. 2008. Federal Funding for the study of antimicrobial resistance in nosocomial pathogens: No ESKAPE. *National Institute of Allergy and Infectious Diseases*. 197(8): 1079-1081.
- São, C. 2018. Engineering of phage-derived lytic enzymes: Improving their potential as antimicrobials. *Antibiot.* 7(2): 29.
- Shen, J., L. Lv., X. Wang., Z. Xiu and G. Chen. 2017. Comparative analysis of CRISPR Cas systems in *Klebsiella* genomes. *J. Basic Microbiol.* 57(4): 325-336.
- Sun, H., Y. Li., X. Shi., Y. Lin., Y. Qiu., J. Zhang and Q. Sun. 2015. Association of CRISPR/Cas evolution with *Vibrio parahaemolyticus* virulence factors and genotypes. *Foodborne Pathog Dis.* 12(1): 68-73.
- Tamura, K., G. Stecher., D. Peterson, A. Filipski and S. Kumar. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30(12): 2725-2729.
- Yooseph, S., K. Nealson., D. Rusch., J. McCrow., C. Dupont., M. Kim., J. Johnson., R. Montgomery., S. Ferreira and K. Beeson. 2010. Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature*. 468(7320): 60-66.
- Won, K. and S. Park. 2008. Pathogenicity of *Vibrio harveyi* to cultured marine fishes in Korea. *Aquaculture*. 285(1-4): 8-13.