Comparative Analysis of CRISPR-Cas Systems in *Vibrio* and *Photobacterium* Genomes of High Influence in Aquaculture Production

Angel E. Parra [1] ![ID], Laura Antequera [1] ![ID], Carla A. Lossada [2] ![ID], Francelys V. Fernández-Materán [2] ![ID], María M. Parra [3] ![ID], Paulo Centanaro [4] ![ID], Diego Arcos [4] ![ID], Alex Castro [4] ![ID], Luis S. Moncayo [5] ![ID], Freddy Romero [6] ![ID], José L. Paz [7] ![ID], Joan Vera-Villalobos [8] ![ID], Aleivi E. Pérez [3] ![ID], Ysaias J. Alvarado [2,*] ![ID], Lenin A. González-Paz [1,*] ![ID]

1   Laboratorio de Genética y Biología Molecular (L.G.B.M), Departamento de Biología, Facultad Experimental de Ciencias (F.E.C), Universidad del Zulia (L.U.Z), Maracaibo, Republica Bolivariana de Venezuela
2   Laboratorio de Caracterización Molecular y Biomolecular - Centro de Investigación y Tecnología de Materiales (CITeMA)- Sección de Microbiología Molecular y Biofísica, Instituto Venezolano de Investigaciones Científicas (I.V.I.C)- Zulia, Maracaibo, Venezuela
3   Laboratorio de Microbiología General (L.M.G), Departamento de Biología. Facultad Experimental de Ciencias (F.E.C), Universidad del Zulia (L.U.Z), Maracaibo, Venezuela
4   Facultad de Ciencias Agrarias, Universidad Agraria de Ecuador, Ecuador
5   Universidad Católica de Cuenca, Ecuador
6   Department of Medicine. Baylor College of Medicine, Houston, TX, USA
7   Departamento de Física, Escuela Politécnica Nacional, Ladrón de Guevara, E11-253,170517, Quito, Ecuador
8   Escuela Superior Politécnica del Litoral, Facultad de Ciencias Naturales y Matemáticas, Departamento de Ciencias Químicas y Ambientales, Laboratorio de Síntesis y Caracterización Química, Departamento de Ciencias Biológicas, Facultad de Ciencias de la Salud, Universidad Técnica de Manabí, Ecuador
*   Correspondence: lgonzalezpaz@gmail.com (L.G.); alvaradoysaias@gmail.com (Y.A.).

Abstract: Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs), together with cas genes, constitute the systems CRISPR-Cas, a novel type of innate defense system in prokaryotes for the destruction of exogenous elements (plasmids or phages). The genus *Vibrio* and *Photobacterium* represents one of the main responsible infectious agents in aquaculture and cause huge losses for the shrimp industry. The few bioinformatic investigations on CRISPR/Cas systems in the Vibrionaceae family are mainly oriented to clinical species. Therefore, the objective of this research was to make a comparative analysis of CRISPR-Cas systems in *Vibrio* and *Photobacterium* genomes of strong influence in aquaculture production. A total of 69 complete genomes across the family were studied. Seven CRISPR loci were identified in 4 species (*V. parahaemolyticus*, *V. harveyi*, *V. vulnificus,* and *Photobacterium damselae*). The CRISPR/Cas systems were found to correspond to subtypes I-E, I-F, and III-D. Only 47% of the spacers presented homology with bacteriophages. The results showed a low presence of CRISPR structures in the species of *Vibrio* and *Photobacterium* with immunity to specific phages. These results are interesting if sensitivity is considered, and it can be explored for the development of strategies such as phagotherapy in shrimp farming.

Keywords: CRISPR-Cas; bioinformatic*;* Vibrionaceae; aquaculture; phagotherapy.

1. Introduction

In shrimp culture, the family of proteobacteria Vibrionaceae, especially species grouped in the genus *Vibrio*, represents one of the main infectious agents of great influence in larval production associated with *Litopenaeus vannamei* culture. These gram-negative bacteria are ubiquitous and widely distributed in aquatic environments, from brackish water to deep seawater, in association with marine animals, algae, and detritus. Vibrios also are present in estuarine and marine aquatic ecosystems where shrimp are naturally found and/or cultivated. This genus of bacteria is considered part of the normal microbiota of shrimp; actually is part of the largest percentage of total bacteria isolated from the digestive tract, gills and cuticle, and occasionally in the hemolymph. However, under stress or unbalance conditions in the natural microbiota, these bacteria can develop infections in shrimp [1,2,3].

Among the genus *Photobacterium*, the main virulent strain for shrimp is *Photobacterium damselae* [4], which is a pathogenic bacterium for a wide variety of marine animals and has been reported as a primary pathogen causing diseases in shrimp. The geographical distribution of this bacterium is increasing, and nowadays, it constitutes an emerging pathogen in aquaculture. Its pathogenicity is attributed to the production of at least four different toxins: damselysin (Dly), phobalysin P (PhlyP), phobalysin C (PhlyC), and the phospholipase (PlpV). These four toxins are secreted via the type II secretion system [5]. High abundance and prevalence of *Photobacterium* genera have been reported in shrimp guts, and it reflects the persistence and adaptation of members from these genera to the shrimp gut environments and may potentially explain the susceptibility of shrimps to non-native pathogenic *Photobacterium* strains [6, 7].

Regarding the genus *Vibrio*, the main virulent strains for shrimp at several larval stages are *V. alginolyticus*, *V. anguillarum*, *V. parahaemolyticus*, *V. harveyi* and *V. vulnificus* [8]. They can cause diverse effects, since low levels of survival with repercussions and important losses in shrimp culture, to human gastroenteritis outbreaks given by consumption of marine products contaminated with these opportunistic pathogens [8-12]. Prevention and control of diseases in shrimp farming are based primarily on the use of antibiotics [13]. However, the growing emergence of bacterial resistant strains to antimicrobial agents and the fact of not having effective antibiotics in the near future have prompted the need to search for new antagonistic alternatives to those microorganisms that have not yet developed resistance mechanisms. These alternatives must have fewer secondary effects and high specificity to avoid the disturbance of human intestinal microbiota so that greater effectiveness of the treatments can be guaranteed [14,15].

One of the promising strategies towards the objective of finding more effective 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 (enzybiotics). These are inhibitory agents that have been described as potential antimicrobials with therapeutic interest [16], and they have important applications in shrimp farming [8,15]. Although it is not clear if there is any mechanism of bacterial resistance against enzybiotics (phage lysins), a large number of investigations conducted on a group of prokaryotic genomic sequences known as "clustered regularly interspaced short palindromic repeats" (or CRISPR) have shown a distinctive feature of most bacterial and archaeal genomes, which endows them with resistance against bacteriophages [8,17,18].

Recently factors such as pathogenicity, stress response, tolerance to physicochemical factors, and even metabolic aspects related to biological cycles in prokaryotes have been associated with the presence of these CRISPR sequences, which along with Cas nucleases, make up the CRISPR/Cas complex, a system discovered as a specific immune defense mechanism present in bacteria and archaea. This allows them to gain exogenous DNA fragments from their natural pathogens (bacteriophages) and incorporate them between repeated palindromic sequences. Subsequently, generate an RNA called crRNA (CRISPR-RNA) when transcribed, so that in case of a second infection, the crRNA coupled with the Cas proteins recognize the transcript of the pathogen and thus degrades mRNA in a manner analogous to RNA interference (siRNA) [19,20].

In the case of *Vibrio*, important advances have been made [8, 21-26]. However, it is possible this information needs to be updated since it is known that the genomes of environmental species may exhibit a greater number of undetected CRISPR structures. In part, because many of these microbial genomes in databases have only been partially studied in the search for CRISPR structures. The use of tools that do not lose even the smallest CRISPRs structures (containing a single spacer) or those CRISPRs that differ slightly from canonical structures is required [23]. These repeated sequences continue to be a subject of debate and intense interest in molecular microbiology and biophysics, and their study has been the focus of comparative analysis using diverse bioinformatics tools [21,25-34]. Additionally, research on the presence of CRISPRs has focused mainly on bacteria of clinical interest, while less is known about CRISPR dynamics in environmental species [35]. Therefore, the objective of this research was to study, through comparative analysis, CRISPR-Cas systems in Vibrio and Photobacterium genomes of high influence in aquaculture production.

2. Materials and Methods

2.1. Genomic sequences and identification of CRISPR structures.

Five species of Vibrio and one of Photobacterium were studied: *V. alginolyticus*, *V. anguillarum*, *V. parahaemolyticus*, *V. harveyi*, *V. vulnificus* and *Photobacterium damselae,* all associated with systemic infections in shrimp, such as vibriosis. Genome sequences were obtained from specific web sites compiled in the Entrez Genome project list (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi) or the Genomes OnLine Database (http://www.genomesonline.org/) [36]. Sixty-nine (69) complete genomic sequences were studied. These included fourteen from *V. alginolyticus*, thirteen from *V. anguillarum*, twenty-one from *V. parahaemolyticus*, four from *V. harveyi*, fifteen from *V. vulnificus*, as well as two complete chromosomal sequences from *Photobacterium damselae* (Table 1). CRISPR loci were obtained from the CRISPRdb database [37]. Alternatively, the CRISPR loci in the genomes were identified by CRISPRFinder [38], using two search criteria.

First, the screening of possible CRISPR locations by detecting maximum repetitions (with the maximum possible extension to the right or the left without incurring a mismatch) by means of utilizing the VMatch package [39,40]. 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. Second, a criterion based on the CRISPR function determination, for which filters are added to help validate CRISPRs, such as the search for non-identical spacers. 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). To discriminate between the confirmed CRISPR structures and those questionable - small structures similar to CRISPR having only two or three repeated sequences or direct repeats (DR), the sequences were 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 based on the repetition and similarity of the spacers. This is because questionable CRISPR structures often tend to be byproducts of the CRISPRFinder identification process, and they 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 were considered. They allow more accurate identification of true CRISPR. To identify the potential orientation of CRISPR, two tests have been implemented: prediction of the structure orientation by comparing to a set of data of consensus repetitions contained in the CRISPRFinder database, by means of the CRISPRDirection program, and the calculation of the AT% of 100 bp that flank the matrix on both sides, taking into account that the region with the highest percentage of AT is considered a leader (Figure 1).

2.2. Bioinformatic analysis and statistical results.

For the search of cas genes, the first step consisted in the identification of open reading frames (ORF) with Prodigal [41]. Then, these ORFs were analyzed with the MacSyFinder program, implementing the HMM method (Hidden Markov Models) in a library of known Cas proteins [42]. Alternatively, BLAST was used to identify the cas genes in the upstream and downstream sequences of the CRISPR and TIGRFAM loci [43]. The Cas type and subtype could be found through the analysis of Cas conglomerates, thanks to the CRISPRCas-Finder program [40]. 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 ≥85% (minimum of 28/33 coincident nucleotides) [44,45]. The 20 bps upstream and downstream sequence of the protospacers was used to identify the possible motifs adjacent to the protospacer (PAM) visualized by WebLogo [46]. The prediction of the secondary structure of CRISPR-type RNA, as well as the minimum free energy (MFE) formation, were performed by RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) with default arguments [47].

The organisms, as well as the MFE of their CRISPR, are shown in Table 1. The conservation of direct repeats (DR) was represented by WebLogo. They were compared with the proposed criteria for the prediction of secondary structures of prokaryotic siRNA precursors (psiRNA) transcribed by CRISPR loci based on DR [21,48] or complete CRISPR structures [37]. Phylogenetic trees were generated by using the unweighted pair group method (UPGMA) for the core protein Cas, specifically Cas1 employing the MUSCLE algorithm [48]. In parallel, multiple sequence alignments and phylogenetic analyses were performed using ClustalX; and dendrograms were visualized with the accompanying application NJ Plot [49]. The conservation of Cas proteins was evaluated by multiple sequence alignments with Geneious global alignment (Needleman-Wunsch) with predetermined arguments [47]. A t-test was used as a statistical model to establish differences between the comparative parameters analyzed (including the thermodynamic findings in the elements and CRISPR structures).

3. Results

3.1. Organization and diversity of CRISPR/Cas systems in vibrionaceae genome.

A total of 7 CRISPR structures were found in 5 of the genomes of the Vibrionaceae family, 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, whose characteristics are described in Table 1.

Table 1. CRISPR/Cas systems in Vibrionaceae genomes.

Sp	String	CRISPR Access ID	Start-End (pb)	DR$_c$	DR consensus	CRISPR Subtypes	Cas genes
Vp	FORC_022	PO13249.1	826587-828294	29	GTTAACTGCCACACAGGCAGC TTAGAAA	I-F	6
Vh	ATCC 43316	P014039.2	73377-75664	38	GTGTTCTCCGTACCCACGGAG ATGAACCG	I-5	8
Vv	YJ016	A000038.2	1694586-1695247	10	GTTTCAGACATGCCCGGTTTA GACGGGATTAAGAC	III-D	7
	93U204	P009262.1	511712-515040	56	GTTCACTGCCGTATAGGCAGC TTAGAAA	I-F	6
			2123187-2123443	4	GATATTTCTAACTGGGATACT TCCAATGTAAA		
*Pd	KC-Na-1	P021152.1	246163-246610	8	CTTCACTGCCGAGTAGGCAGC TTAGAAA	I-F	6
			257070-257278	4	CTTCACTGCCGAGTAGGCAGC TTAGAAAT		

Vp. FORC_022 (CP013248.1), *Vh*. ATCC43516 (CP014038.2), *Vv*. YJ016 (BA000037.2) and 93U204 (CP0092621.1), *Pd*. KC-Na-1 (CP021151.1), are the acronyms of *V. parahaemolyticus* FORC_022, *V. harveyi* ATCC 43516, *V. vulnificus* YJ016 and 93U204, and *P. damselae* KC-Na-1 respectively Species (Sp) found by their respectively accesion number at the GenBank; CRISPR access ID proyect at the CRISPRdatabase, location (Start-End), and characteristic such as Direct Repeat sequences counting (DR$_c$), DR consensus sequence, subtype and CRISPR-Associated (*cas*) genes clusters counting is also specified. Besides, spacers quantity were counted, was found 28 spacers (32 pb length each one) belonging to *Vp*. FORC_022 CRISPR, 37 spacers (32-33pb) to *Vh*. ATCC43516, and 9 (33-37pb) to *Vv*. YJ016; two distanced CRISPR were found at *Vv*. 93U204 and *Pd*. KC-Na-1 genome, the longest Vv CRISPR (3328pb) owns 55 spacers 32pb length, and the shortest (256pb) 3 of 43pb; to *Pd* longer, CRISPR (447pb) belongs 7 spacers of 32pb against 3 of 31pb long from the second and smaller CRISPR (29pb) analyzed for this bacteria. *before classified in *V. damselae*
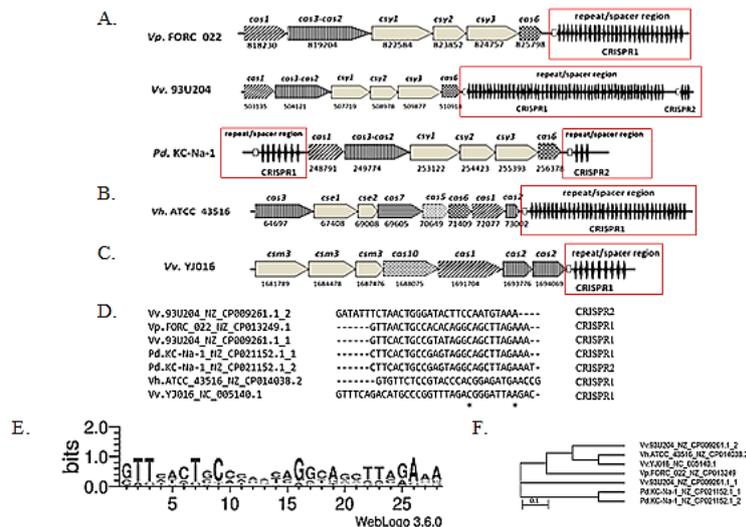

Figure 1. Architecture loci, gene organization, and direct repetitions (DR) conservation of CRISPR systems. (A), subtypes I-F;, subtypes I-E; C, subtypes III-D. The black diamonds represent the CRISPR loci, and the white arrows (with lines in different directions) represent the different cas genes found. The relative position on the chromosome is indicated below each gene. (E) the logo of the sequence was created by WebLogo 3.6.0, and the comparative analysis of the CRISPR repeats (B and C). Seven repeated CRISPR sequences were aligned using ClustalX [28] or generated using the MUSCLE algorithm in MEGA6 (D). *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. damselae* KC-Na-1, respectively.
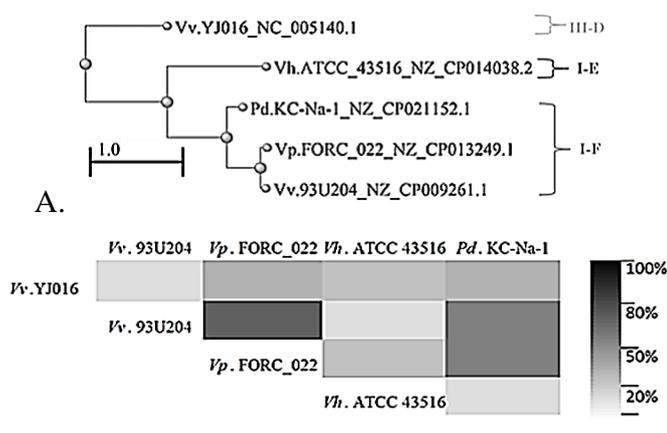
Figure 2. Phylogenetic tree for Cas1 proteins in the family Vibrionaceae.The UPGMA tree of the Cas1 protein was generated using the MUSCLE algorithm in MEGA6. The Cas1 proteins representative of the subtypes found were selected (I-F, I-E, and III-D) (A). Percentage identity of Cas1 proteins in the Vibrionaceae 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. damselae* KC-Na-1, respectively.

Three different arrays of CRISPR-Cas systems were detected with different locations on the chromosomes (Figure 1). The subtypes of CRISPR/Cas systems were found to correspond to I-E, I-F, and III-D. The distinction between the subtypes of CRISPR found by the analysis of direct repeat (DR) conservation was also evaluated (Figure 2). 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 and Photobacterium, or the members with CRISPR matrices in the Vibrionaceae family, a phylogenetic tree was constructed based on the homology of the Cas1 proteins of each species (Figure 2).

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 DRs were found to be interspaced, with a minimum of 3 to a maximum of 55 spacer sequences of a similar length (31 - 43 bp). Besides the fact that most of the spacer sequences presented size of 32 bp, the analysis of the variance indicates that there is no statistically significant difference ($p > 0.01$). However, there is a very low correlation between the number and size of these sequences (Figure 3).
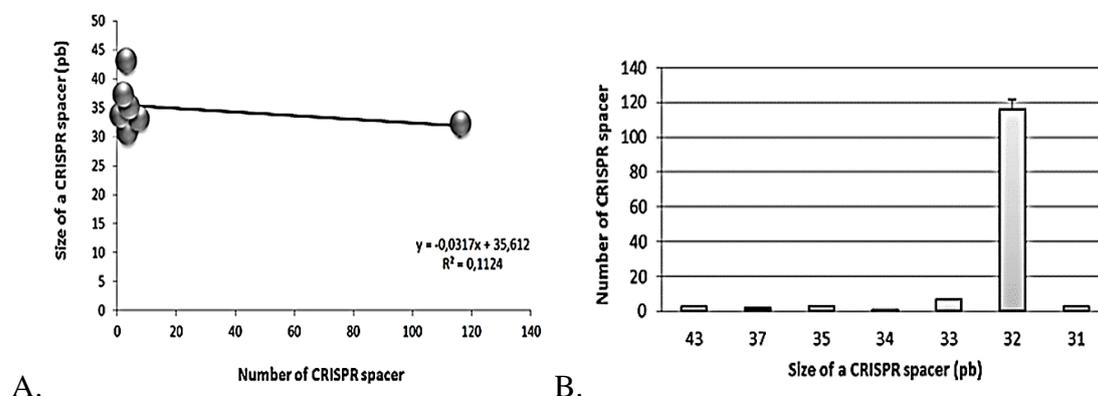

Figure 3. Correlation of the number of spacers in the CRISPR loci and the size of the CRISPR (bp) spacers found. A very low negative correlation was found ($\leq 30\%$) (A). Variability of the size of the CRISPR spacers and their number ($p > 0.05$). The x-axis represents the size of a CRISPR spacer, in nucleotides. The y-axis represents the number of CRISPR spacing sequences of a given size (B).

The total number of spacers detected per subtypes of CRISPR was 96/142 for subtypes I-F (68%), 37/142 for subtypes I-E (26%), and 9/142 for subtypes III-D (6%). CRISPR structures exhibited between 6 to 8 genes associated with the CRISPR sequences (*cas*) (Table 1). The repeat sequences could form a putative secondary RNA structure with a free energy of the thermodynamic ensemble of -8.20 kcal/mol for CRISPR1 in *V. parahaemolyticus* FORC_022, -9.90 kcal/mol for CRISPR1 in *V. harveyi* ATCC 43516, -6.80 kcal/mol for CRISPR1 in *V. vulnificus* YJ016, -8.60 kcal/mol for CRISPR1 and -1.90 kcal/mol for CRISPR2 in *V. vulnificus* 93U204, -8.60 kcal/mol for CRISPR1 and -8.60 kcal/mol for CRISPR2 in *P. damselae* KC-Na-1 (Figure 4).
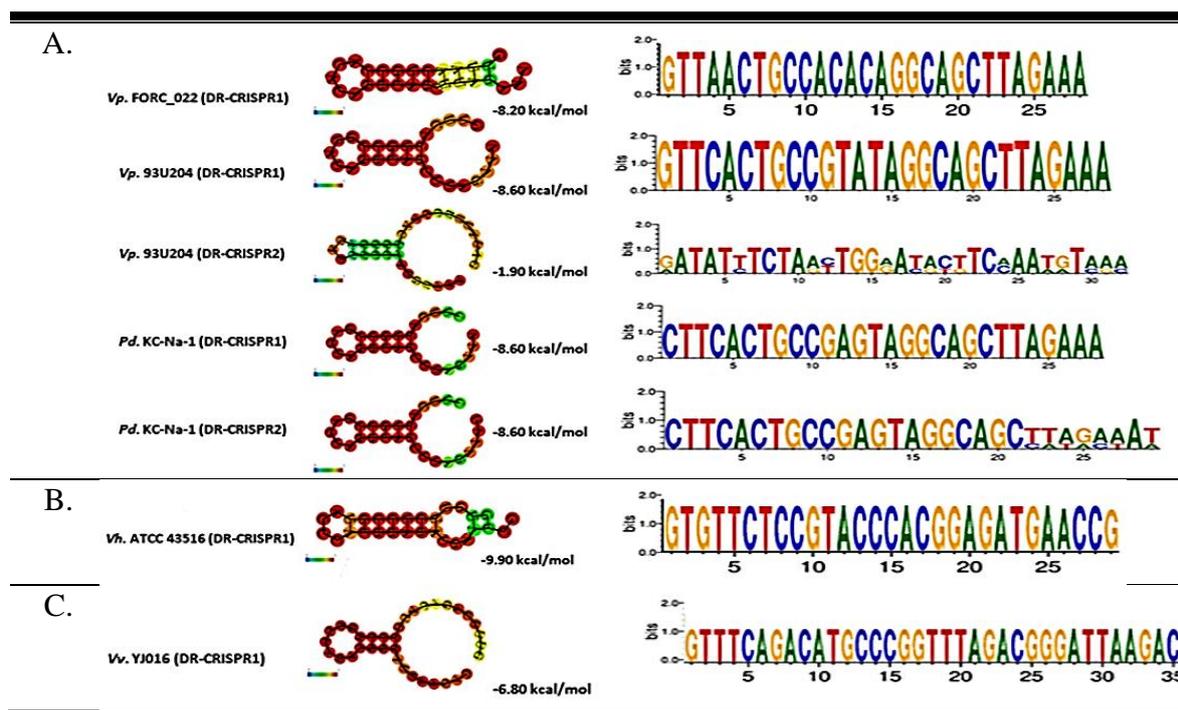

Figure 4. Secondary structures of RNA and the minimum free energy of formation (MFE) of the direct repeats (DR) of each subtype of CRISPR found as well as the conservation of the DR. Represented with WebLogo 3.6.0. Typical stable stem-loop structures predicted consistently for DR by RNAfold are shown. The darkest base pairs represent the highest probability of mating. A, subtypes I-F; B, subtypes I-E; C, subtypes III-D. *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. damselae* KC-Na-1, respectively

3.2. Identification of CRISPR repeats is locus-specific.

CRISPRs are typically defined by the sequence of the repeat. In line with this, CRISPR matrices were found with many typical direct repeats (DR) with a minimum of four to a maximum of 56 DR, which had a length between 28-35 bp. The observed CRISPR subtypes I-F loci presented a minimum of 4 DR and a maximum of 56, while the only loci subtypes I-E and III-D detected exhibited 38 and 10 DR, respectively. Interestingly, subtypes III-D presented the longer DR with about 35 bp, as compared to the subtypes I-E and I-F, of 29 and 28 bp, respectively (Table 1). CRISPR repeats were compared using multiple sequence alignment, where only one repeat sequence per CRISPR locus was included. The dendrogram presented in Figure 1F illustrates how the various typical repeat sequences of the Vibrionaceae family split into CRISPR repeat families. The comparative analysis between the DR showed a high degree of conservation of the repeated sequences despite the existence of non-consensus repetitions characterized by the presence of point mutations in different regions, mainly in the

terminal region of the typical sequences that trigger the presence of non-consensus DR (Figure 1D).

The conservation of the DR regrouped the species in three large groups (Figure 1D,E,F), locating at a third level the species *V. vulnificus* 93U204, *V. harveyi* ATCC 43516, and *V. vulnificus* YJ016 as the most distant from each other and from the rest of the species, but more related to *V. parahaemolyticus* FORC_02. These findings are interesting considering that in the case of *V. vulnificus* 93U204 and YJ016, they belong to the same species and that the phylogenetic relationship between this species and *V. harveyi* ATCC 43516 is not as close as it would be expected, taking into account the presence of homologous spacer sequences in *V. harveyi* ATCC 43516 that provide immunity against plasmid p48/10 present in related *V. vulnificus* species. Interestingly, the DR found in the CRISPR1 matrix present in *V. vulnificus* 93U204 is less related to its CRISPR2 structure and therefore, the dendrogram regroups it with a ratio moderately distant from the rest of the strains of the same species compared before, showing the differentiation between CRISPR matrices in the same genome. On the other hand, the first well-defined group based on the conservation of DRs is represented by the two repeated sequences of the species *P. damselae* KC-Na-1, which reflects both a high conservation of such structures as well as a clear phylogenetic distance with the other genera of the family of Vibrionaceae, despite the existence of spacer sequences that give this species resistance against the plasmid pYJ016 present in *V. vulnificus* YJ016 or against typical bacteriophages of the genus *Vibrio*.

3.3. The origin of CRISPR spacers.

Among the 130/142 (92%) unique spacer sequences found in the CRISPR loci, it is important to note that all of them had homology with some sequence contained in the GenBank database, as well as great diversity in the homology to related or not sequences of Vibrionaceae family. About 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 5A). These results indicate a clear immunity function against foreign genetic material (plasmids or phages), and high specificity of the different CRISPR/Cas systems.
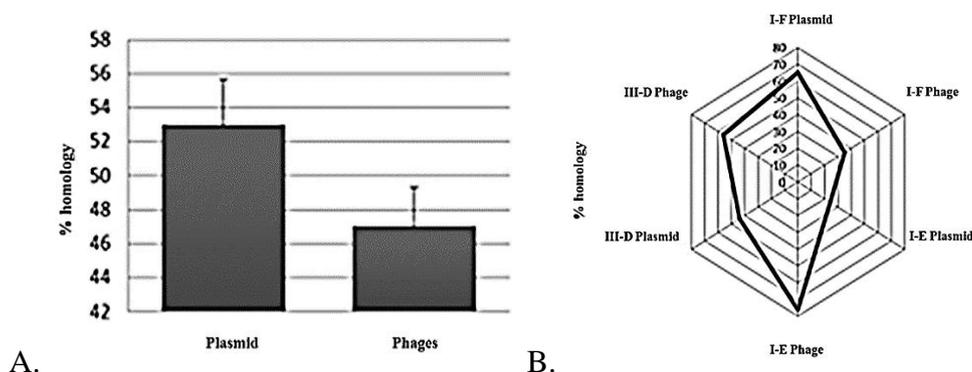

Figure 5. The origin of the CRISPR spacers. The overall distribution of CRISPR spacers in plasmids and phages (A) and for type (B) was determined by sequence identity. All spacers were compared against the GenBank database to find a sequence of homology.

The subtypes of CRISPR that gathered the greatest number of spacer sequences were the I-F subtypes present in the species *P. damselae* 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%)

respectively. In subtypes, I-E and III-D of species *V. harveyi* ATCC 43516 and *V. vulnificus* YJ016, respectively, were observed 37 and 9 corresponding spacer sequences. It is important to note that the spacer sequences found in the CRISPR subtypes I-E were the ones that presented a greater homology to bacteriophage sequences with 76% (28/37) followed by subtypes III-D with 56% (5/9) and in smaller proportion, the I-F subtypes with 35% (34/96). Interestingly, and in a contrary way, the spacers found in subtypes I-F exhibited the highest homology to plasmid sequences with 65% (62/96) (Figure 5B).

The homologies evidenced by the spacer sequences indicated a high diversity in terms of the recognition targets of CRISPR immunity systems, represented by the homology to DNA sequences related or not to representatives of the Vibrionaceae family. Specifically, there is little immunity to typical *Vibrio* and *Photobacterium* infective bacteriophages (Figure 5A) such as phage VP2, phage douglas 12A4 and phage VSK (the latter is also a host of *P. damselae* KC-Na-1) and to plasmid sequences also present in strains of the genus *Vibrio* and *Photobacterium* (Figure 5A), 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 subtypes I-F), or against phage of *Vibrio* 11895-B1 and plasmid p48/10 of *V. vulnificus* 48/10 (in the CRISPR system subtypes I-E), or against phage of *V. eugene* 12A10 and plasmid pSNUTY1 of *V. coralliilyticus* strain SNUTY-1 (in the CRISPR system subtypes III-D). This contrasts with the immunity found in greater proportion against a great variety of infective phages of *Bacillus*, enterobacteria, *Pseudomonas,* and cyanobacteria (results not shown). These results show a low occurrence of CRISPR structures in these Vibrionaceae species conferring immunity to lytic-specific phages, especially of the genus *Vibrio*. This is important because this bacterial weakness in the presence of bacteriophages can be useful in the adequacy of the phagotherapy to the shrimp industry.

3.4. Search for CRISPR polymorphisms.

A total of 130/142 (92%) unique spacer sequences were found in the 7 CRISPR loci observed, 9/130 (7%) for subtypes III-D, 37/130 (28%) for subtypes I-E and 84/130 (65%) for subtypes I-F, distributed in 120/130 (92%) spacer sequences unique for representatives of the genus *Vibrio* and 10/130 (8%) for *P. damselae* (Figure 6).
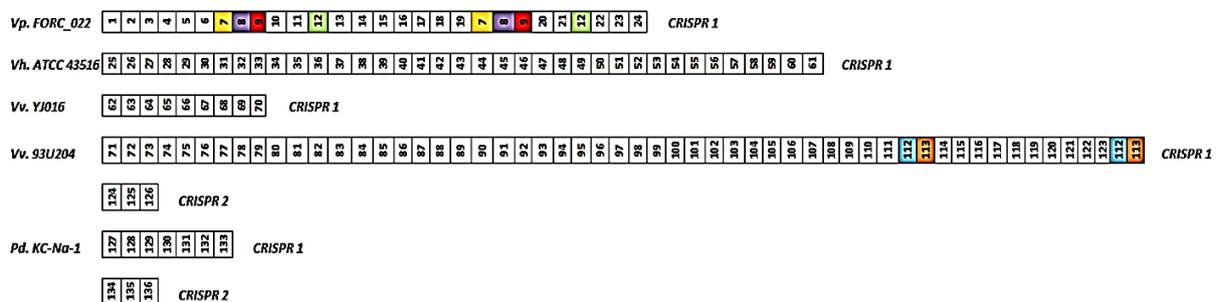

Figure 6. CRISPR polymorphisms. The spacers are represented by boxes; direct repeats (DR) are not included. The same number and color represent the same spacer. The white color represents a unique spacer. Spacers with self-orientation are indicated by dots at the bottom of the spacer. *Vp*. FORC_022, *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. damselae* KC-Na-1, respectively.

The CRISPR loci found in both *Vibrio* and *P. damselae* vary in length and content of spacers: the longest CRISPR locus contains 58 spacers and is found in *V. vulnificus* 93U204,

while only three spacers were found both second CRISPR loci of *V. vulnificus* 93U204 and in one of the two loci of *P. damselae* KC-Na-1. Marked conservation of the spacer sequences was evident, with very few spacer sequences identical between the different subtypes of CRISPR systems exhibited by the different species of the Vibrionaceae family (Figure 6).

3.5. Determination of PAM sequences and self-targeting spacers.

This research study found six spacers (4.2%) that coincide with their own sequences within the CRISPR matrices, which could be indicative of the presence of possible self-targeting spacers. Specifically, 4/6 of them belonged to the species *V. parahaemolyticus* FORC_022 and 2/6 were located in the species *V. vulnificus* 93U204; in both cases, the CRISPR systems were of subtypes I-F. These results indicate a low proportion of possible self-targeting spacers in environmental *Vibrio* species and exhibit the fact that in their CRISPR systems, immunity against mobile genetic elements such as plasmids predominates. No PAM sequences were found adjacent to the potential self-directed proto-spacer, indicating that these structures are not associated with a phenomenon of autoimmunity, although, in the CRISPR1 structure of the species *V. vulnificus* YJ016, all potential proto-spacers have a PAM sequence represented by the triplet "AAG". However, there is no coincidence between such spacers with their own sequences, this shows the high specificity of the different CRISPR/Cas systems studied. No identical CRISPR matrices were found between the different strains examined despite being them closely related (Figure 6). In fact, strains *V. vulnificus* YJ016 and *V. vulnificus* 93U204 were isolated in Taiwan while *V. parahaemolyticus* FORC_022 and *P. damselae* KC-Na-1 were isolated from the Republic of South Korea (NCBI, http://www.ncbi.nlm.nih.gov/, [20]). Besides, the absence of identical CRISPR matrices is important for their use as potential targets for the detection, typing, and epidemiological surveillance of these species of high clinical-environmental interest, and also for the shrimp industry.

3.6. Identification and comparison of CRISPR associated genes and cas proteins.

A total of thirteen putative *cas* genes were identified upstream of the CRISPR region in the Vibrionaceae genome. In all the genomes analyzed with defined CRISPR structures, genes associated with CRISPR or *cas* were found, including all the genes belonging to the *cas* (*cas*1, *cas*2, and *cas*3, except *cas*4) nucleus as well as various subtype genes characteristic of the *csy* family, *cse* and *csm*. All the *cas* genes detected belonged to the class I of the architecture of the genomic loci for the subtypes of CRISPR/Cas systems, and were distributed in seven typical genes of the CRISPR matrices of the I-F subtype (*cas*1, *cas*2, *cas*3, *cas*6, *csy*1, *csy*2, and *csy*3) in *P. damselae* KC-Na-1, *V. parahaemolyticus* FORC_022 and *V. vulnificus* 93U204, eight of subtype I-E (*cas*1, *cas*2, *cas*3, *cas*5, *cas*6, *cse*1, and *cse*2) in *V. harveyi* ATCC 43516 and five corresponding to arrangements of subtype III-D (*cas*1, *cas*2, *cas*10, *csm*3) in *V. vulnificus* YJ016 (Figure 1A,B,C).

It is important to note that subtype III-D presented three sequences of the characteristic gene *csm*3 with chromosomal positions defined upstream of the CRISPR region in the *V. vulnificus* YJ016 genome. Among the three *csm*3 genes found, the variant located in the chromosome region 1681789-1683951 pb shares a certain degree of homology with subtype III-A according to the bioinformatic analysis performed (results not shown). The phylogenetic relationship is given by the homology with the Cas1 protein groups the species *P. damselae* KC-Na-1, *V. parahaemolyticus* FORC_022, and *V. vulnificus* 93U204 in subtype I-F, *V.*

harveyi ATCC 43516 in subtype I-E, and *V. vulnificus* YJ016 in subtype III-D (Figure 2A). The highest identity was found between the Cas1 of *V. vulnificus* 93U204 and *V. parahaemolyticus* FORC_022 (> 80%), followed by an identity> 50% between *V. vulnificus* 93U204 and *P. damselae* KC-Na-1, as well as the observed between *V. parahaemolyticus* FORC_022 and *P. damselae* KC-Na-1. These results correspond to the subtypes of CRISPR matrix to which they belong, all grouped in subtype I-F (Figure 2B). Interestingly, the lowest homology was found between Cas1 of the same species of *V. vulnificus* (93U204 and YJ016) and between *V. harveyi* ATCC 43516 and *V. vulnificus* 93U204 (Figure 7).
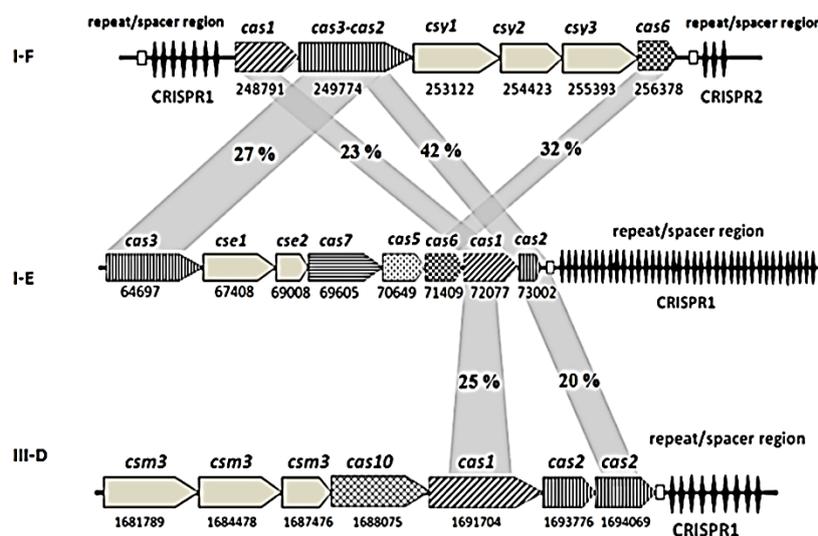

Figure 7. Description of CRISPR subtypes s I-F, I-E, and III-D found in the genomes Vibrionaceae. The black diamonds represent the CRISPR loci, and the white arrows (with lines in different directions) represent the different *cas* genes found. The genomic number of ORF is indicated below each gene. The numbers in the gray shading indicate the percentage identity between the homologous sequences of the Cas protein. Cas protein sequences that do not share significant similarities were not connected with gray shading.

Regarding the homology between the different Cas proteins present in the three subtypes s of CRISPR found, it was observed that there is a greater kinship between subtypes I-F and I-E, represented by a putative protein homology of 42% for Cas2, 32% for Cas6, 27% for Cas3 and 23% for Cas1. This coincides with the description in the literature, where the kinship between these two *cas* genes is seen, placing them in the same class and subtypes of CRISPR. A lower homology exists between subtype I-E and III-D, evidenced only for *cas*1 (25%) and *cas*2 (20%) (Figure 7). Interestingly, the comparative analysis between the DR showed a high degree of conservation of the direct repeats and regrouped the species into three large groups (Figure 1 D, E, F) that are closely related to the clade derived from the homology with the Cas1 proteins.

3.7. Predicting the secondary structure of psiRNA (prokaryotic siRNA) and CRISPR structures.

After predicting the secondary structure of prokaryotic precursors of the RNA species that are key in the mechanism of adaptive immunity, it was found that the DR could form a putative RNA secondary structure with the free energy of the thermodynamic ensemble or minimum free energy (MFE) of formation of -8.50 kcal/mol for CRISPR1 in *V. parahaemolyticus* FORC_022, -10.70 kcal/mol for CRISPR1 in *V. harveyi* ATCC 43516, -7.20 kcal/mol for CRISPR1 in *V. vulnificus* YJ016, -8.50 kcal/mol for CRISPR1 and -3.00 kcal/mol for CRISPR2 in *V. vulnificus* 93U204, -8.40 kcal/mol for CRISPR1 and -8.40 kcal/mol for CRISPR2 in *P. damselae* KC-Na-1 (Figure 4). While the energy of formation of the transcribed

CRISPR-type structures obtained by calculating the MFE from the analysis of the complete structure (DR plus spacers) showed that the complete CRISPR could form a putative RNA secondary structure with the free energy of ensemble of -532.90 kcal/mol for CRISPR1 in *V. parahaemolyticus* FORC_022, -870.30 kcal/mol for CRISPR1 in *V. harveyi* ATCC 43516, -237.80 kcal/mol for CRISPR1 in *V. vulnificus* YJ016, -1189.20 kcal/mol for CRISPR1 and -106.50 kcal/mol for CRISPR2 in *V. vulnificus* 93U204, -173.80 kcal/mol for CRISPR1 and -102.30 kcal/mol for CRISPR2 in *P. damselae*KC-Na-1 (Figure 4).

The non-consensus DR that presented point mutations distributed along the typical sequences were also evaluated to predict the secondary structure of these DR variants, and it was found that the MFE of the non-consensus DR of CRISPR1 in *V. parahaemolyticus* FORC_022 remained at a mean -8.95 kcal/mol and -10.70 kcal/mol for CRISPR1 in *V. harveyi* ATCC 43516. The free energy of ensemble for DR non-consensus of CRISPR1 and CRISPR2 in *V. vulnificus* 93U204 presented means of -8.50 kcal/mol and -4.36 kcal/mol, respectively, while the mean thermodynamic of the non-consensus DR for CRISPR2 in *P. damselae* KC-Na-1 was -8.40 kcal/mol. The DR of the matrices CRISPR1 in *V. vulnificus* YJ016 and *P. damselae* KC-Na-1 did not present repeated sequences variants (Table 2). The analysis of these results indicates that there is no statistically significant difference ($p > 0.01$) so that it cannot be emphasized that the point mutations found in different regions of the repeated sequences cause important differences between the thermodynamic mean of the MFE of formation of the typical DR and its variants.

Table 2. Minimum free energy of formation (MFE) of typical repetitions and variants in the CRISPR/Cas systems found.

Species/DR	Type	Repeat sequence (5'-3')	MFE (kcal/mol)
Vp. FORC_022_{DR_CRISPR1}	Typical repeat	GTTAACTGCCCACACAGGCAGCTTAGAAA	-8.20
	Repeat variants	**T**TTAACTGCCCACACAGGCAGCTTAGAAA	-9.50
	Repeat variants	GTTAACTGCCCACACAGGCAGCTTAGA**G**A	-8.20
Vh. ATCC 43516_{DR_CRISPR1}	Typical repeat	GTGTTCTCCGTACCCACGGAGATGAACCG	-9.90
	Repeat variants	GTGTTCTCCGTACCCACGGAGATGA**GCCA**	-10.20
Vv.YJ016_{DR_CRISPR1}	Typical repeat	GTTTCAGACATGCCCGGTTTAGACGGGATTAAGAC	-6.80
Vv. 93U204_{DR_CRISPR1}	Typical repeat	GTTCACTGCCGTATAGGCAGCTTAGAAA	-8.60
	Repeat variants	GTTCACTGCCGTATAGGCAGC**G**TAGAAA	-8.60
Vv. 93U204_{DR_CRISPR2}	Typical repeat	GATATTTCTAATTGGAA**C**AC**G**TCAAAT**A**TAA**C**	-1.90
	Repeat variants	**A**ATATTTCTA**G**TTGGAAT**GT**TTCAAA**A**GT**CG**A	-4.10
	Repeat variants	GATATTTCTAACTGGGATACTTCCAATGTAAA	-2.80
	Repeat variants	GATAT**C**TCTAACTGGGATACTTCCAATGTAAA	-4.90
Pd. KC-Na-1_{DR_CRISPR1}	Typical repeat	CTTCACTGCCGAGTAGGCAGCTTAGAAA	-8.60
Pd. KC-Na-1_{DR_CRISPR2}	Typical repeat	CTTCACTGCCGAGTAGGCAGCTTAGAAAT	-8.60
	Repeat variants	CTTCACTGCCGAGTAGGCAGC**CATACTAA**	-8.60

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. damselae* KC-Na-1, respectively; DR_CRISPR1 and DR_CRISPR2, direct repeats of relative location on the chromosome of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) designated CRISPR1 and CRISPR2.

3.8.. Discussion.

Subtypes I-F of CRISPR/Cas systems isolated from *V. parahaemolyticus* species found in seafood, environment [50,51], as well as in clinical samples, has been described [40]. Likewise, the presence of I-E subtypes has been reported in *V. cholerae* O395, in which

3 *cas* core genes (*cas*1, *cas*2, and *cas*3) and 5 specific genes of the I-E subtype (*cse*1, *cse*2, *cse*3, *cse*4, *cas*5e) have been seen [21,37]. On the other hand, it has been reported that CRISPR subtypes I-E systems always have two loci pairs, at least that is what has been evidenced in Enterobacteria genomes [49]. However, our results differ from this, since in the species *V. harveyi* ATCC 43516 only a confirmed CRISPR structure was detected, like that described in strains like *V. cholerae* O395 [52]. Interestingly, a new form of disposition of *cas* genes was found for the subtypes III-D detected in *V. vulnificus* YJ016. Although, according to the current classification model [53], it should still be a CRISPR/Cas subtypes III-D system because the *cas* genes found are within the arrangement III-D. However, a pair of *cas*2 genes were detected as well as one *cas*1 gene at chromosomal locations 1693776-1694069 bp and 1691704 bp, respectively. This does not correspond to the genes of subtype III-D [54], and rather, they remind us of the arrangement of the subtype III-A architectures, which possess both the *cas*10 and *csm*3 genes of subtypes III-D, and a *cas* nucleus represented by the *cas*1 and *cas*2 [55].

Therefore, it is between two very closely related subtypes, with a possible common origin. These results could explain why the comparative analysis of one of the three *csm*3 genes found in the region chromosome 1681789-1683951 pb shares a certain degree of homology with subtype III-A (results not shown). The comparative analysis between direct repeats (DR) showed an important degree of conservation for the repeated sequences exhibited by the evaluated species, despite the existence of non-consensus repeats, characterized by presenting SNPs mainly in the terminal region of the typical sequences. The entire detected DRs could form secondary RNA structures with minimum free energies (MFE) of assembly that turn out to be thermodynamically highly spontaneous. Several authors predict the thermodynamic stability of transcribed CRISPR-type structures by calculating the MFE either from the DR [21,45] or from the complete CRISPR structure (DR plus spacers) [51].

However, few studies consider the thermodynamic analysis of complete CRISPR structures. The present study observed after the thermodynamic calculation carried out as recently reported for the comparative analysis between the MFE of the CRISPR and the consequent establishment of the thermodynamically more stable structures based only on the DR, that it shows magnitudes with few differences and with an average in the thermodynamic divergence of,

$$\Delta G_d \ = \ \mid \Delta G_{(DR1)} \ - \ \Delta G_{(DR2)} \mid = -5 \ \text{kcal/mol},$$

being able to reach zero, in the case of genomes with two or more CRISPR matrices, such as *V. vulnificus* 93U204 and *P. damselae* KC-Na-1, respectively, which renders questionable thermodynamic discrimination between stable CRISPR structures and those that are not, at least under the conditions of this study. However, when applying the strategy that considers the complete CRISPR structure, MFE had clear differences from the thermodynamic point of view between the ΔG calculated, with a thermodynamic mean of,

$$\Delta G_D \ = \ \mid \Delta G_{(CRISPR1)} \ - \ \Delta G_{(CRISPR2)} \mid = -1082 \ \text{kcal/mol},$$

and,

$$\Delta G_D \ = \ \mid \Delta G_{(CRISPR1)} \ - \ \Delta G_{(CRISPR2)} \mid = -72 \ \text{kcal/mol},$$

for *V. vulnificus* 93U204 and *P. damselae* KC-Na-1, respectively. It has been reported that structures with longer stem tend to have lower MFEs and more stable structures [44]. And this stability may be influenced by the length of the "stem", that is, the longer the stem, the greater the stability of the secondary structure, a characteristic shared by the CRISPR found in this study, which theoretically exhibited a more spontaneous assembly.

Ninety-two percent 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, a fact that indicates a clear immunity function against foreign genetic material and high specificity of the different CRISPR/Cas systems studied. On the other hand, although PAM sequences were not found adjacent to the potential self-directed proto-spacer, which is indicative of CRISPR structures that are not associated with a phenomenon of autoimmunity, interestingly, in the structure CRISPR1 present in *V. vulnificus* YJ016, all the possible proto-spacers have a well-defined PAM sequence conformed by the triplet "AAG", as already described in the interference machinery of *Escherichia coli* in which both *cas*3 nuclease and cascade complex participate [53].

It has been described 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 hence the cell can mediate the immune reaction in a specific way by recognizing the homologous sequence to such spacer [18]. Therefore, the profile of spacers can be a reflection of the bacterial lifestyle or their habitat [31]. In this study, it is important to point out that most of the spacer sequences studied do not compromise the integrity of the CRISPR systems, and no clear indications of autoimmunity have been found; rather, the systems present in *Vibrio* and *P. damselae* confer a marked immunity against mobile genetic elements like plasmids (53%), just like what has been reported [25]. Likewise, spacer sequences were found against a wide variety of infective phages from *Bacillus*, Enterobacteria, *Pseudomonas,* and cyanobacteria. These are promising results because they show a low occurrence of CRISPR structures with immunity to lytic-specific phages in *Vibrio* and *P. damselae*, which is interesting if we consider that this sensitivity can be explored for the development of strategies such as phagotherapy in aquaculture, a technology based on the use of enzymes or phages destined to control typical bacterial pathogens [56].

4. Conclusions

This research represents the first report of CRISPR systems in the species of strong influence in aquaculture production *P. damselae*. And although CRISPR sites can be found in several *Vibrio* species, only four of the genomes analyzed in the genus *Vibrio*, specifically associated with vibriosis, harbored these RNA-based immunity matrices. These loci include all the expected typical characteristics of CRISPR matrices, mainly highly conserved direct repeats, variable spacers, and *cas* genes. The bioinformatic analysis and comparison of the CRISPR/Cas system can help to understand the environmental adaptability of each evolutionary lineage of species of agro-industrial interest, providing data support for bacterial typing, traceability, analysis, and non-conventional CRISPR exploration. On the other hand, the conservation of the aforementioned CRISPR structures was observed at different levels, which allows the use of CRISPR in epidemiology, typing, and evolution studies of this type of pathogen of clinical-environmental interest. Besides, all direct repeats can form stable secondary RNA structures, and it has been shown that spacer sequences originate

predominantly from exogenous plasmids. The low occurrence of CRISPR structures in these species of the family Vibrionaceae studied that confer immunity to lytic-specific phages is a factor that can be used to ponder the implementation of alternative initiatives such as phagotherapy in shrimp farming or aquaculture in general.

Funding

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Kalatzis, P.; Castillo, D.; Katharios, P.; Middelboe, M. Bacteriophage interactions with marine pathogenic Vibrios: Implications for phage therapy. *Antibiot* **2018**, *7*, https://doi.org/10.3390/antibiotics7010015.
2. Kaneshamoorthy, S. Efficacy of Natural Extracts Against *Vibrio Parahaemolyticus* Causing Acute Hepatopancreatic Necrosis Disease (AHPND). In: *Pacific White Shrimp Penaeus Vannamei*. Master's Theses. The University of Arizona, Tucson, United Stated Of America, **2020**.
3. Yu, L.H.; Teh, C.S.J.; Yap, K.P.; Ung, E.H.; Thong, K.L. Comparative genomic provides insight into the virulence and genetic diversity of Vibrio parahaemolyticus associated with shrimp acute hepatopancreatic necrosis disease. *Infection, Genetics and Evolution* **2020**, *83*, https://doi.org/10.1016/j.meegid.2020.104347.
4. Liu, F.; Liu, G.; Li, F. Characterization of two pathogenic Photobacterium strains isolated from Exopalaemon carinicauda causing mortality of shrimp. *Aquaculture* **2016**, *464*, 129-135, https://doi.org/10.1016/j.aquaculture.2016.06.019.
5. Terceti, M.; Vences, A.; Matanza, X.; Dalsgaard, I.; Pedersen, K.; Osorio, C. Molecular Epidemiology of *Photobacterium damselae* subsp. Damselae Outbreaks in Marine Rainbow Trout Farms Reveals Extensive Horizontal Gene Transfer and High Genetic Diversity. *Front. Microbiol* **2018**, *9*, https://doi.org/10.3389/fmicb.2018.02155.
6. Garibay-Valdez, E.; Martínez-Córdova, L.R.; Vargas-Albores, F.; Gollas-Galván, T.; Lago-Leston, A.; Calderón, K.; Martínez-Porchas, M. Biofilm consumption shapes the intestinal microbiota of shrimp (Penaeus vannamei). *Aquaculture Nutrition* **2019**, *25*, 427-435, https://doi.org/10.1111/anu.12868.
7. Moi, I.M.; Roslan, N.N.; Leow, A.T.C.; Ali, M.S.M.; Rahman, R.N.Z.R.A.; Rahimpour, A.; Sabri, S. The biology and the importance of Photobacterium species. *Applied Microbiology and Biotechnology* **2017**, *101*, 4371-4385, https://doi.org/10.1007/s00253-017-8300-y.
8. MdZoqratt, M.; Han Eng, W.; Thanh Thai, B.; Austin, C.; Ming Gan, H. Microbiome analysis of Pacific white shrimp gut and rearing water from Malaysia and Vietnam: implications for aquaculture research and management. *PeerJ* **2018**, *6*, https://doi.org/10.7717/peerj.5826.
9. Won, K.M.; Park, S.I. Pathogenicity of Vibrio harveyi to cultured marine fishes in Korea. *Aquaculture* **2008**, *285*, 8-13, https://doi.org/10.1016/j.aquaculture.2008.08.013.
10. Xie, T.; Yu, Q.; Tang, X.; Zhao, J.; He, X. Prevalence, antibiotic susceptibility and characterization of Vibrio parahaemolyticus isolates in China. *FEMS Microbiology Letters* **2020**, *367*, https://doi.org/10.1093/femsle/fnaa136.
11. Kongrueng, J.; Srinitiwarawong, K.; Nishibuchi, M.; Mittraparp-arthorn, P.; Vuddhakul, V. Characterization and CRISPR-based genotyping of clinical trh-positive Vibrio parahaemolyticus. *Gut Pathogens* **2018**, *10*, https://doi.org/10.1186/s13099-018-0275-4.
12. Zhang, M.; Liu, C.; Shi, Y.; Wu, J.; Wu, J.; Chen, H. Selective endpoint visualized detection of Vibrio parahaemolyticus with CRISPR/Cas12a assisted PCR using thermal cycler for on-site application. *Talanta* **2020**, *214*, https://doi.org/10.1016/j.talanta.2020.120818.
13. Alagappan, K.M.; Deivasigamani, B.; Somasundaram, S.T.; Kumaran, S. Occurrence of Vibrio parahaemolyticus and Its Specific Phages from Shrimp Ponds in East Coast of India. *Current Microbiology* **2010**, *61*, 235-240, https://doi.org/10.1007/s00284-010-9599-0.

14. Lomelí-Ortega, C.O.; Martínez-Díaz, S.F. Phage therapy against Vibrio parahaemolyticus infection in the whiteleg shrimp (Litopenaeus vannamei) larvae. *Aquaculture* **2014**, *434*, 208-211, https://doi.org/10.1016/j.aquaculture.2014.08.018.

15. Rice, L.B. Federal Funding for the Study of Antimicrobial Resistance in Nosocomial Pathogens: No ESKAPE. *The Journal of Infectious Diseases* **2008**, *197*, 1079-1081, https://doi.org/10.1086/533452.

16. São-José, C. Engineering of Phage-Derived Lytic Enzymes: Improving Their Potential as Antimicrobials. *Antibiotics* **2018**, *7*, https://doi.org/10.3390/antibiotics7020029.

17. Oliveira, J.; Castilho, F.; Cunha, A.; Pereira, M.J. Bacteriophage therapy as a bacterial control strategy in aquaculture. *Aquaculture International* **2012**, *20*, 879-910, https://doi.org/10.1007/s10499-012-9515-7.

18. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D.A.; Horvath, P. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. **2007**, *315*, 1709-1712, https://doi.org/10.1126/science.1138140.

19. Horvath, P.; Barrangou, R. CRISPR/Cas, the Immune System of Bacteria and Archaea. **2010**, *327*, 167-170, https://doi.org/10.1126/science.1179555.

20. Lyons, C.; Raustad, N.; Bustos, M.A.; Shiaris, M. Incidence of Type II CRISPR1-Cas Systems in Enterococcus Is Species-Dependent. *PLOS ONE* **2015**, *10*, https://doi.org/10.1371/journal.pone.0143544.

21. Rodrigues, M.; Hullahalli, K.; Palmer, K. CRISPR-mediated removal of antibiotic resistance genes in Enterococcus faecalis populations. *The FASEB Journal* **2017**, *31*, 903-909.

22. Chakraborty, S.; Waise, T.M.; Hassan, F.; Kabir, Y.; Smith, M.A.; Arif, M. Assessment of the evolutionary origin and possibility of CRISPR-Cas (CASS) mediated RNA interference pathway in Vibrio cholerae O395. *In silico biology* **2009**, *9*, 245-254.

23. Sun, H.; Li, Y.; Shi, X.; Lin, Y.; Qiu, Y.; Zhang, J.; Liu, Y.; Jiang, M.; Zhang, Z.; Chen, Q.; Sun, Q.; Hu, Q. Association of CRISPR/Cas Evolution with Vibrio parahaemolyticus Virulence Factors and Genotypes. *Foodborne Pathogens and Disease* **2014**, *12*, 68-73, https://doi.org/10.1089/fpd.2014.1792.

24. Labbate, M.; Orata, F.D.; Petty, N.K.; Jayatilleke, N.D.; King, W.L.; Kirchberger, P.C.; Allen, C.; Mann, G.; Mutreja, A.; Thomson, N.R.; Boucher, Y.; Charles, I.G. A genomic island in Vibrio cholerae with VPI-1 site-specific recombination characteristics contains CRISPR-Cas and type VI secretion modules. *Scientific Reports* **2016**, *6*, https://doi.org/10.1038/srep36891.

25. Baliga, P.; Shekar, M.; Venugopal, M.N. Investigation of direct repeats, spacers and proteins associated with clustered regularly interspaced short palindromic repeat (CRISPR) system of Vibrio parahaemolyticus. *Molecular Genetics and Genomics* **2019**, *294*, 253-262, https://doi.org/10.1007/s00438-018-1504-8.

26. Cázares, G. *Evaluación antagonista in vitro de bacterias aisladas de ecosistemas marinos en contra de Vibrio parahaemolyticus AHPND como patógeno de importancia acuícola*. Doctoral dissertation, Universidad Autónoma De Nuevo León, Monterrey, Mexico, **2020**.

27. McDonald, N.D.; Regmi, A.; Morreale, D.P.; Borowski, J.D.; Boyd, E.F. CRISPR-Cas systems are present predominantly on mobile genetic elements in Vibrio species. *BMC Genomics* **2019**, *20*, 1-12, https://doi.org/10.1186/s12864-019-5439-1.

28. Shen, J.P.; Zhao, D.; Sasik, R.; Luebeck, J.; Birmingham, A.; Bojorquez-Gomez, A.; Licon, K.; Klepper, K.; Pekin, D.; Beckett, A.N.; Sanchez, K.S.; Thomas, A.; Kuo, C.-C.; Du, D.; Roguev, A.; Lewis, N.E.; Chang, A.N.; Kreisberg, J.F.; Krogan, N.; Qi, L.; Ideker, T.; Mali, P. Combinatorial CRISPR–Cas9 screens for de novo mapping of genetic interactions. *Nature Methods* **2017**, *14*, 573-576, https://doi.org/10.1038/nmeth.4225.

29. Rahmatabadi, S.S.; Nezafat, N.; Negahdaripour, M.; Hajighahramani, N.; Morowvat, M.H.; Ghasemi, Y. Studying the features of 57 confirmed CRISPR loci in 29 strains of Escherichia coli. *Journal of Basic Microbiology* **2016**, *56*, 645-653, https://doi.org/10.1002/jobm.201500707.

30. Hidalgo-Cantabrana, C.; Crawley, A.B.; Sanchez, B.; Barrangou, R. Characterization and Exploitation of CRISPR Loci in Bifidobacterium longum. **2017**, *8*, https://doi.org/10.3389/fmicb.2017.01851.

31. Negahdaripour, M.; Nezafat, N.; Hajighahramani, N.; Rahmatabadi, S.S.; Ghasemi, Y. Investigating CRISPR-Cas systems in Clostridium botulinum via bioinformatics tools. *Infection, Genetics and Evolution* **2017**, *54*, 355-373, https://doi.org/10.1016/j.meegid.2017.06.027.

32. Zhang, J.; Li, X.; Deng, Z.; Ou, H.-Y. Comparative Analysis of CRISPR Loci Found in Streptomyces Genome Sequences. *Interdisciplinary Sciences: Computational Life Sciences* **2018**, *10*, 848-853, https://doi.org/10.1007/s12539-018-0301-8.

33. Qun Wang, Zhe Li, Lin Zhao, Zijian Wang, Hongqun Zhao, Biao Kan and Bo Pang. Research on detection method of Vibrio parahaemolyticus based on CRISPR/Cas protein[J]. *Disease Surveillance* **2020**.

34. Zhao, X.; Yu, Z.; Xu, Z. Study the Features of 57 Confirmed CRISPR Loci in 38 Strains of Staphylococcus aureus. **2018**, *9*, 1-7, https://doi.org/10.3389/fmicb.2018.01591.

35. Hao, M.; Cui, Y.; Qu, X. Analysis of CRISPR-Cas System in Streptococcus thermophilus and Its Application. **2018**, *9*, 257-259, https://doi.org/10.3389/fmicb.2018.00257.

36. Grissa, I.; Vergnaud, G.; Pourcel, C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **2007**, *8*, 172-178, https://doi.org/10.1186/1471-2105-8-172.

37. Bernal, A.; Ear, U.; Kyrpides, N. Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. *Nucleic Acids Res* **2001**, *29*, 126-127, https://doi.org/10.1093/nar/29.1.126.

38. Benson, D.; Karsch, I.; Lipman, D.; Ostell, J. GenBank. *Nucleic Acids Res* **2008**, *36,* 25-28,https://doi.org/10.1093/nar/gks1195.

39. Grissa, I.; Vergnaud, G.; Pourcel, C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* **2007**, *35*, W52-W57, https://doi.org/10.1093/nar/gkm360.

40. Kurtz, S.; Schleiermacher, C. REPuter: fast computation of maximal repeats in complete genomes. *Bioinformatics* **1999**, *15,* 426-427, https://doi.org/10.1093/bioinformatics/15.5.426.

41. Abouelhoda, M.I.; Kurtz, S.; Ohlebusch, E. Replacing suffix trees with enhanced suffix arrays. *Journal of Discrete Algorithms* **2004**, *2*, 53-86, https://doi.org/10.1016/S1570-8667(03)00065-0.

42. Hyatt, D.; Chen, G.-L.; LoCascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **2010**, *11*, 119-123, https://doi.org/10.1186/1471-2105-11-119.

43. Abby, S.S.; Néron, B.; Ménager, H.; Touchon, M.; Rocha, E.P.C. MacSyFinder: A Program to Mine Genomes for Molecular Systems with an Application to CRISPR-Cas Systems. *PLOS ONE* **2014**, *9*, 110-116, https://doi.org/10.1371/journal.pone.0110726.

44. Horvath, P.; Coûté-Monvoisin, A.-C.; Romero, D.A.; Boyaval, P.; Fremaux, C.; Barrangou, R. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *International Journal of Food Microbiology* **2009**, *131*, 62-70, https://doi.org/10.1016/j.ijfoodmicro.2008.05.030.

45. Crooks, G.; Hon, G.; Chandonia, J.; Brenner, S. WebLogo: a sequence logo generator. *Genome Res* **2004**, *14*, 1188-1190, https://doi.org/10.1101/gr.849004.

46. Hofacker, I.L. Vienna RNA secondary structure server. *Nucleic Acids Res* **2003**, *31*, 3429-3431, https://doi.org/10.1093/nar/gkg599.

47. Li, W.; Bian, X.; Evivie, S.; Huo, G. Comparative Analysis of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) of *Streptococcus thermophilus* St-I and its Bacteriophage-Insensitive Mutants (BIM) Derivatives. *Curr. Microbiol* **2016**, *73*, 393-400, https://doi.org/10.1007/s00284-016-1076-y.

48. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **2013**, *30*, 2725-2729, https://doi.org/10.1093/molbev/mst197.

49. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947-2948, https://doi.org/10.1093/bioinformatics/btm404.

50. Baliga, P.; Shekar, M.; Venugopal, M.N. Detection and characterization of clustered regularly interspaced short palindromic repeat-associated endoribonuclease gene variants in *Vibrio parahaemolyticus* isolated from seafoods and environment. *Veterinary World* **2019**, *12*, https://doi.org/10.14202/vetworld.2019.689-695.

51. Makarova, K.S.; Wolf, Y.I.; Alkhnbashi, O.S.; Costa, F.; Shah, S.A.; Saunders, S.J.; Barrangou, R.; Brouns, S.J.J.; Charpentier, E.; Haft, D.H.; Horvath, P.; Moineau, S.; Mojica, F.J.M.; Terns, R.M.; Terns, M.P.; White, M.F.; Yakunin, A.F.; Garrett, R.A.; van der Oost, J.; Backofen, R.; Koonin, E.V. An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology* **2015**, *13*, 722-736, https://doi.org/10.1038/nrmicro3569.

52. Stern, A.; Keren, L.; Wurtzel, O.; Amitai, G.; Sorek, R. Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends in Genetics* **2010**, *26*, 335-340, https://doi.org/10.1016/j.tig.2010.05.008.

53. Nuñez, J.K.; Kranzusch, P.J.; Noeske, J.; Wright, A.V.; Davies, C.W.; Doudna, J.A. Cas1–Cas2 complex formation mediates spacer acquisition during CRISPR–Cas adaptive immunity. *Nature Structural & Molecular Biology* **2014**, *21*, 528-534, https://doi.org/10.1038/nsmb.2820.

54. González-Delgado, A.; Mestre, M.R.; Martínez-Abarca, F.; Toro, N. Spacer acquisition from RNA mediated by a natural reverse transcriptase-Cas1 fusion protein associated with a type III-D CRISPR–Cas system in Vibrio vulnificus. *Nucleic Acids Res* **2019**, *47*, 10202-10211, https://doi.org/10.1093/nar/gkz746.

55. Marraffini, L.A.; Sontheimer, E.J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **2010**, *463*, 568-571, https://doi.org/10.1038/nature08703.

56. Hoikkala, V.; Almeida, G.M.F.; Laanto, E.; Sundberg, L.-R. Aquaculture as a source of empirical evidence for coevolution between CRISPR-Cas and phage. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2019**, *374*, https://doi.org/10.1098/rstb.2018.0100.