

Detection and characterization of clustered regularly interspaced short palindromic repeat-associated endoribonuclease gene variants in *Vibrio parahaemolyticus* isolated from seafoods and environment

Pallavi Baliga, Malathi Shekar and Moleyur Nagarajappa Venugopal

Department of Fisheries Microbiology, Karnataka Veterinary, Animal and Fisheries Sciences University, College of Fisheries, Mangalore, Karnataka, India.

Corresponding author: Malathi Shekar, e-mail: malathishekar@rediffmail.com

Co-authors: PB: pallavibaliga1@gmail.com, MNV: mnvenu@rediffmail.com

Received: 14-11-2018, **Accepted:** 26-03-2019, **Published online:** 21-05-2019

doi: 10.14202/vetworld.2019.689-695 **How to cite this article:** Baliga P, Shekar M, Venugopal MN (2019) Detection and characterization of clustered regularly interspaced short palindromic repeat-associated endoribonuclease gene variants in *Vibrio parahaemolyticus* isolated from seafoods and environment, *Veterinary World*, 12(5): 689-695.

Abstract

Aim: In *Vibrio parahaemolyticus*, the clustered regularly interspaced short palindromic repeat (CRISPR)-associated *cas6* endoribonuclease gene has been shown to exhibit sequence diversity and has been subtyped into four major types based on its length and composition. In this study, we aimed to detect and characterize the *cas6* gene variants prevalent among *V. parahaemolyticus* strains isolated from seafoods and environment.

Materials and Methods: Novel primers were designed for each of the *cas6* subtypes to validate their identification in *V. parahaemolyticus* by polymerase chain reaction (PCR). In total, 38 *V. parahaemolyticus* strains isolated from seafoods and environment were screened for the presence of *cas6* gene. Few representative PCR products were sequenced, and their phylogenetic relationship was established to available *cas6* gene sequences in GenBank database.

Results: Of the 38 *V. parahaemolyticus* isolates screened, only about 40% of strains harbored the *cas6* endoribonuclease gene, among which 31.6% and 7.9% of the isolates were positive for the presence of the *cas6*-a and *cas6*-d subtypes of the gene, respectively. The subtypes *cas6*-b and *cas6*-c were absent in strains studied. Sequence and phylogenetic analysis also established the *cas6* sequences in this study to match GenBank sequences for *cas6*-a and *cas6*-d subtypes.

Conclusion: In *V. parahaemolyticus*, the Cas6 endoribonuclease is an associated protein of the CRISPR-cas system. CRISPR-positive strains exhibited genotypic variation for this gene. Primers designed in this study would aid in identifying the *cas6* genotype and understanding the role of these genotypes in the CRISPR-cas immune system of the pathogen.

Keywords: *cas6* gene, clustered regularly interspaced short palindromic repeats-cas operon, endoribonuclease, type IF system, *Vibrio parahaemolyticus*.

Introduction

Vibrio parahaemolyticus is a Gram-negative, halophilic bacterium that naturally inhabits the marine, estuarine, and coastal environments. *V. parahaemolyticus* is commonly responsible for acute gastroenteritis illness and to lesser frequency wound infections and septicemia in humans [1]. Infections due to this bacterium are associated with the consumption of contaminated seafood [2,3] and exposure of open wounds to seawater [4]. Recently, *V. parahaemolyticus* has also been reported to be the causative agent of acute hepatopancreatic necrosis syndrome in penaeid shrimp taking a heavy toll on the shrimp aquaculture industry [5]. *V. parahaemolyticus* pathogenicity has been attributed to several virulence determinants including the well-known thermostable direct hemolysin (TDH)

and TDH-related hemolysin encoded by the *tdh* and *trh* genes, respectively [6]. A recent study has demonstrated a significant association between the virulence factor *tdh* and the presence of clustered regularly interspaced short palindromic repeats (CRISPR)/*cas* in *V. parahaemolyticus* [7].

The CRISPR-cas (CRISPR-associated proteins) system is a prokaryotic adaptive immune system against genomic invaders such as viruses and plasmids [8,9]. Due to its importance and widespread applications, it has been a subject for comprehensive research recently [10]. The CRISPR-cas system consists of short (~25-40 bp) direct DNA repeats separated by spacer sequences and associated variable *cas* genes [8]. The Cas proteins are a diverse group of proteins which are believed to function as nucleases, helicases, and RNA-binding proteins [11]. Defense against foreign genetic elements is achieved in three stages: Adaptation, expression, and interference. The adaptation stage involves the integration of foreign DNA fragments into the bacterial genome at the CRISPR loci. The second stage is characterized by the expression of *cas* genes and the transcription of the CRISPR loci to generate a crRNA precursor (pre-crRNA) which is later processed into mature-crRNAs.

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

During interference, the target nucleic acid is recognized and destroyed by the combined action of crRNA and Cas proteins [11,12]. The CRISPR-cas system has been identified in several bacterial and archaeal genomes [13,14] and broadly classified into five major types (I-V) and 16 subtypes (IA-F, I-U, IIA-C, IIIA-D, IV, and V), based on the organization of the CRISPR locus, *cas* gene assembly, and their interference mechanisms [15]. Studies show that different types of CRISPR-cas systems have evolved distinct mechanisms for a mature crRNA generation [16]. In type I-III systems, the mature-crRNA is generated specifically by the Cas protein Cas6 [17]. The Cas6 is a member of the RNA-binding protein superfamily called repeat-associated mysterious proteins (RAMPs) [18] which play a central role in generating guide RNAs for invader defense in prokaryotes [19]. The Cas6 proteins are also reported to be highly divergent and are thought to have coevolved with the highly variable CRISPR RNA repeat sequences [13].

In our previous study, based on computational analysis of existing *V. parahaemolyticus cas6* gene sequences in database, we have shown the existence of four subtypes for *cas6*, designated *cas6*-a, *cas6*-b, *cas6*-c, and *cas6*-d [20]. In this study, we designed novel primers for the polymerase chain reaction (PCR) detection and characterization of *cas6* gene variants in *V. parahaemolyticus*. We report the variant types harbored for the Cas6 endoribonuclease in *V. parahaemolyticus*. Few of the *cas6* sequences were sequenced to look for further variations within the variant types.

Materials and Methods

Ethical approval

Ethical approval was not applicable, as this study does not involve any human or animal studies.

Bacterial strains

V. parahaemolyticus stock cultures (n=38) maintained in Tryptic soy broth with 30% glycerol at -80°C at the Department of Fisheries Microbiology, College of Fisheries, Mangalore, were used in the experiments. The cultures were revived by growing them overnight in 5 ml of Luria-Bertani (LB) broth (HiMedia Laboratories Private Limited, Mumbai, India) at 37°C in a shaker incubator. A loopful of the culture was streaked onto thiosulfate-citrate-bile salts-sucrose agar (HiMedia Laboratories Pvt. Limited, Mumbai, India) to check for the purity of the cultures. *V. parahaemolyticus* strains used in this study are listed in Table-1.

Primer design and PCR validation of *cas6* gene variants

The primers used in this study were designed based on respective *cas6* gene sequences available in GenBank database, using the Primer 3.0 software [21]. The primer sequences are listed in Table-2.

PCR and sequencing of *cas6* gene

Crude lysate of each strain was prepared by first growing them on LB broth at 37°C. 450 µl of ×1 TE buffer was added to 50 µl of the culture and subjected to heating at 94°C for 1 min in the hot dry bath followed by snap cooling in ice. PCR for the identification of *cas6* gene type was carried out using primers designed to target the *cas6* gene types. The PCR was carried out in a 30 µl reaction mixture consisting of 3 µl of ×10 PCR buffer (Genei™, Merck Bioscience, Bengaluru), 50 µM each of the four deoxynucleotide triphosphates, 10 pmol of each primer, and 1.0 U of *Taq* DNA polymerase (Genei™, Merck Bioscience, Bengaluru). 2 µl of crude lysate was used as DNA template. The PCR assay was carried

Table-1: *Vibrio parahaemolyticus* isolates used in this study.

S.No	Isolate	Isolation region	Isolation source	Presence of <i>cas6</i> /subtype
1	VP 1	Karwar	Shrimp	+/d
2	VPh 2	Karwar	Shrimp	-
3	VPh 3	Karwar	Shrimp farm sediment	+/d
4	VP 15	Sasthan	Shrimp farm water	+/a
5	VPh 4	Karwar	Shrimp farm water	-
6	VP 5	Karwar	Shrimp farm sediment	+/d
7	VPh 6	Karwar	Shrimp farm water	-
8	VPh 8	Kundapur	Shrimp	-
9	VPh 9	Kundapur	Shrimp farm water	-
10	VP 4	Mulki	Clam	-
11	VPh 10	Kundapur	Shrimp farm water	-
12	VPh 11	Ankola	Shrimp	-
13	VPh 7	Kundapur	Shrimp	-
14	VPh 12	Ankola	Shrimp	-
15	VPh 13	Ankola	Fish	-
16	VP 6	Mulki	Oyster	+/a
17	MR 32	Sasthan	clams	+/a
18	MR 34	Sasthan	clams	-
19	125	Gangolli	Shrimp farm water	-
20	IIWVp	Katpadi	Shrimp farm water	-
21	VPh 1	Sasthan	Oysters	-
22	VP 3	Sasthan	Oyster	-
23	VP 9	Sasthan	Oyster	+/a
24	VP 10	Sasthan	Clam	+/a
25	VP 11	Sasthan	Oyster	-
26	VP 7	Mulki	Clam	+/a
27	VP 8	Sasthan	Oyster	+/a
28	VP 14	Mulki	Oyster	+/a
29	VP 18	Mulki	Fish	-
30	VP 19	Sasthan	Oyster	-
31	VP 20	Mulki	Oyster	-
32	VP 21	Mulki	Clam	-
33	VP 22	Sasthan	Clam	-
34	VP 23	Mulki	Oyster	+/a
35	VP 25	Mulki	Clam	+/a
36	VP 27	Sasthan	Clam	+/a
37	VP 17	Sasthan	Clam	+/a
38	VPh 5	Mulki	Oyster	-

out in a Programmable Thermocycler (PTC 200, Bio-Rad, CA) with the program as follows: Initial denaturation at 94°C for 5 min, 30 cycles of 94°C for 1 min, annealing temperature as listed in Table-2 for 1 min, 72°C for 1 min, and a final extension at 72°C for 10 min. The PCR products were resolved in 1.5% agarose gel containing 0.5 mg/ml ethidium bromide, and the size of amplicons was determined by comparison with a 100 bp DNA ladder (Genei™, Merck Bioscience, Bengaluru). The bands were visualized using a Gel documentation system (Bio-Rad, USA). The generated PCR products were sequenced (Bioserve Biotechnologies Limited, Hyderabad). The *V. parahaemolyticus* strains VP49 (which harbors the cas6-a genotype) and AQ4037 (reference strain) were used as positive and negative controls, respectively.

In silico analysis

The *V. parahaemolyticus* cas6 sequences were downloaded from the National Centre for

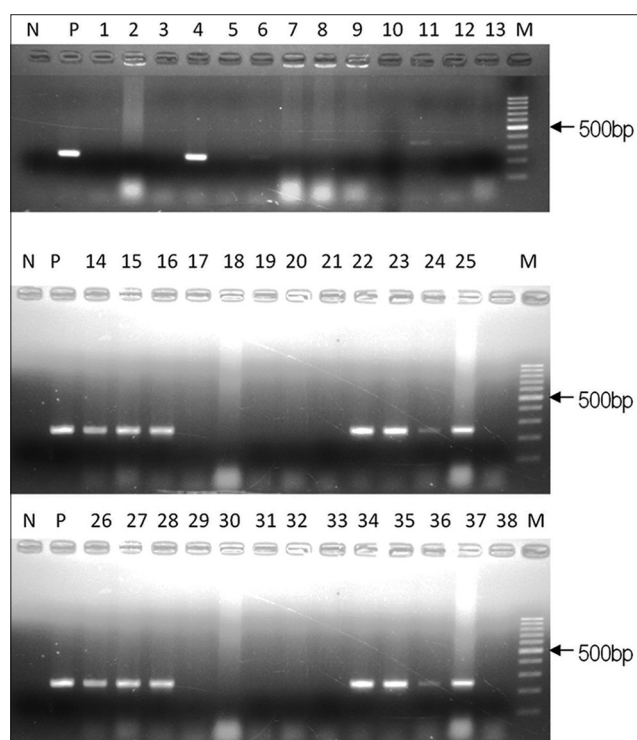


Figure-1: Polymerase chain reaction amplification of cas6-a subtype in *Vibrio parahaemolyticus*. N: Negative control; P: Positive control; M: 100 bp DNA marker; Lane 1-38: Environmental isolates of *V. parahaemolyticus*.

Biotechnology Information (<https://www.ncbi.nlm.nih.gov/>) database. The sequences obtained from our study were aligned with these sequences using the program MultAlin[22] and grouped based on similarity. The phylogenetic tree was generated using MEGA ver.5.0 [23].

Results

PCR for cas6 gene

In this study, the primers designed were used to evaluate and detect the cas6 gene subtypes in seafood and environmental isolates of *V. parahaemolyticus*. For initial identification of the type cas6-a, in *V. parahaemolyticus*, the PCR primers cas6-aF and cas6-aR (Table-2) were used, targeting an internal fragment of the gene. Among the 38 isolates studied, 12 (31.58%) amplified the expected product size of 245 bp (Figure-1) indicating the presence of cas6-a subtype in these strains. Similarly, three strains (VP1, VP5, and VP3) were found to be positive for the type cas6-d variant and amplified a product size of 600 bp (Figure-2). However, cas6-b and cas6-c were not detectable in our strains. The remaining strains showed no amplification with any of the primers and hence considered negative for the presence of cas6 gene. The PCR products of six isolates (Vp6, Vp8, Vp9, Vp10, Vp14, and Vp17) positive for the cas6-a gene variant were sequenced using the primers IF-a-F (forward) and IF-a-R (reverse) to generate a sequence of 597bp (Table-2). Two (Vp1 and Vp5) of the three strains positive for type cas6-d were also sequenced.

In silico analysis of V. parahaemolyticus cas6 sequences

The phylogenetic tree generated based on Cas6 amino acid sequences was seen to subgroup into four major clusters (cas6-a-cas6-d, Figure-3). Pairwise alignment of representative Cas6 protein sequences from the different groups showed the sequences to be diverse with a homology of 21-47% between sequences. It was also seen that there exist residue differences within sequences within a given cluster (Figures-4a and b). Sequence analysis of the protein identified in this study showed proteins with accession IDs AUD40493, AUD40494, AUD40495, AUD40496, AUD 40497, and AUD40498 to be

Table-2: Primers designed in this study for the detection of cas6 gene variants.

Gene	Primer	Sequence (5'-3')	Annealing temperature	Expected product size
cas6-a (internal)	cas6-a F	CCAAGAAACGGTGGGACGTA	60°C	245 bp
	cas6-a R	CGCGTTCTAAAGCTCTTCGC		
cas6-a (full)	IF-a-F	GAACCATCACATTTTACCTGAA	50°C	597 bp
	IF-a-R	CAATGGAACAACCTGCAATG		
cas6-b	IF-b-F	TTATTGGCGGGTTCGCTGTAT	60°C	506 bp
	IF-b-R	TTCGTTGCGAGTCCGTAAC		
cas6-c	IF-c-F	GAAGCATTAAATCGGGCACTG	51°C	500 bp
	IF-c-R	GTCCATAACTTGAGAATGCC		
cas6-d	IF-d-F	TGGACTACTACCAAGAAATTAC	46°C	600 bp
	IF-d-R	TAAAATTGTGGGACAGTC		

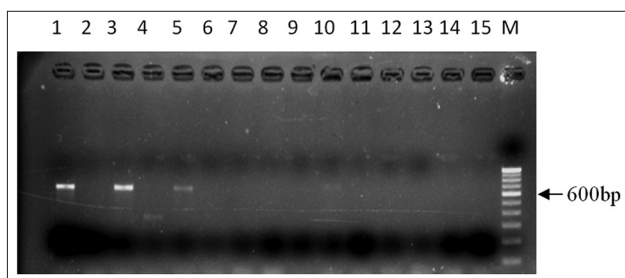


Figure-2: Polymerase chain reaction (PCR) amplification of *cas6-d* subtype in *Vibrio parahaemolyticus*. Only samples positive for PCR shown. Lanes 1-15: Environmental isolates of *V. parahaemolyticus*. M: 100bp DNA marker.

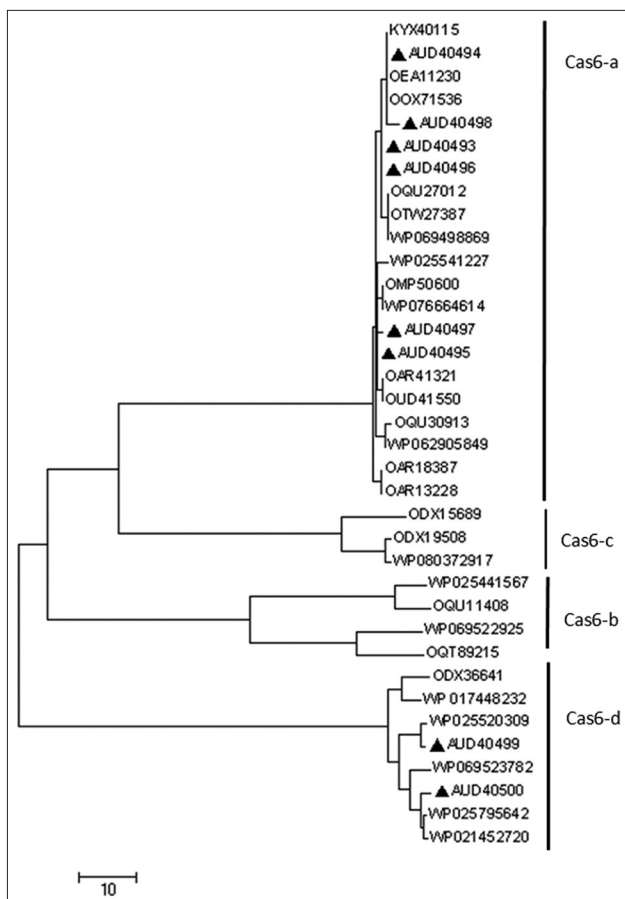


Figure-3: Phylogenetic tree of representative *cas6* protein sequences. Represents Cas6-a (AUD40493, AUD40494, AUD40495, AUD40496, AUD40497, and AUD40498) and Cas6-d (AUD40499 and AUD40500) subtypes identified in this study.

grouped along with the *cas6-a* and AUD40499 and AUD40500 grouped with Cas6-d proteins.

Discussion

V. parahaemolyticus is known to harbor the type IF CRISPR-cas system [7]. *V. parahaemolyticus* subtype IF is associated with CRISPR-associated genes *cas1*, *cas3*, *cas8*, *cas5*, *cas7*, and *cas6*. However, a search for *cas* genes in this study showed that the majority (>90%) of *V. parahaemolyticus* strains contain a minimalistic type IF system containing core genes *cas5*, *cas7*, and *cas6*. Among

these, the peripheral Cas6 domain belongs to the RAMP family of RNases functioning as an endonuclease that interacts with CRISPR RNAs to generate crRNAs [24,25]. Cas5 and Cas7 also belong to the RAMP superfamily and are implicated in interference and stabilization stages of crRNA generation [26]. Bioinformatics-based analysis of the *V. parahaemolyticus cas6* sequence showed the existence of four different sequence subtypes (*cas6-a-cas6-d*), with a majority of strains analyzed seen associated with *cas6-a* subtype [20]. The subtypes *cas6-b* and *cas6-c* were not detected in the strains studied which probably is due to the under-representation of environmental strains taken for analysis. A BLAST analysis for the *V. parahaemolyticus cas6-b*, *cas6-c*, and *cas6-d* subtypes showed homologous *cas6* to be present in *Vibrio anguillarum*, *Vibrio cholerae*, *Vibrio fluvialis*, *Vibrio vulnificus*, and several other *Vibrio* spp. There is much evidence that horizontal transfer of CRISPR and *cas* genes can occasionally occur between different strains, species, and even distant genera [26], which probably explains the diversity seen for *cas6* gene in *V. parahaemolyticus*. Cas6 belonging to the CRISPR-cas type IF system functions as an endonuclease that cleaves CRISPR RNAs to generate pre-crRNA [16]. Recent studies have shown that *cas6* (subtype IF) in *Pseudomonas* spp. recognizes its pre-crRNA substrate with high affinity, recognition of which is mediated by sequence and structure-specific interactions [17,27]. *V. parahaemolyticus* also contains the type IF system, but studies on its Cas6-mediated cleavage and the process to generate a pre-crRNA are lacking. Therefore, based on the diversity observed for *cas6* gene sequences in this study, we probably presume that, in *V. parahaemolyticus*, the Cas6 endonuclease could be functioning in more ways than one in recognizing the primary transcript and in bringing about cleavage. Further, in this study, an alignment of subtypes of *cas6-a* and *cas6-d* sequences revealed substitutions in several amino acid loci in the Cas6 protein (Figure-4a and b). This could be of significance, as studies show that mutations in any of the catalytic residues reduce the endonuclease activity of Cas6 resulting in non-cleaved pre-crRNA [27,28]. Whether these substitutions have any influence on the endonuclease activity, bringing about altered antiviral immunity, therefore, remains to be seen.

The gene coding for *cas6* endonuclease plays a critical role in the CRISPR cas-based adaptive immunity [29]. Studying the prevalence of the genes encoding these systems in seafood and/or environmental isolates of *V. parahaemolyticus* is expected to open new avenues in understanding the dynamics of the CRISPR-based immunity in this pathogen. Our study has elucidated the diversity of *cas6* gene in CRISPR-cas operon harboring *V. parahaemolyticus*. Thus, the PCR primers designed in this study could

help in identifying and distinguishing the presence of *cas6* endoribonuclease variants in this pathogenic bacterium.

In the CRISPR-mediated immunity system, the *cas* genes are functionally paired with CRISPR repeats [30]. Thus, the absence of *cas6* gene in several of our environmental isolates indicates the presence of a sub-population of *V. parahaemolyticus* that probably lacks the CRISPR-cas system. To ascertain this, our future work would

focus on studying CRISPR loci and finding an association with their presence/absence to *cas* genes in *V. parahaemolyticus*. Representative *cas6* genes pertaining to types *cas6*-a and *cas6*-d have been sequenced and deposited in GenBank having accession numbers MG417090 (strain Vp8); MG417091 (StrainVp9); MG417092 (StrainVp10); MG417093 (StrainVp14); MG417094 (StrainVp6); MG417095 (StrainVp17); MG417096 (StrainVp1); and MG417097 (StrainVp5).

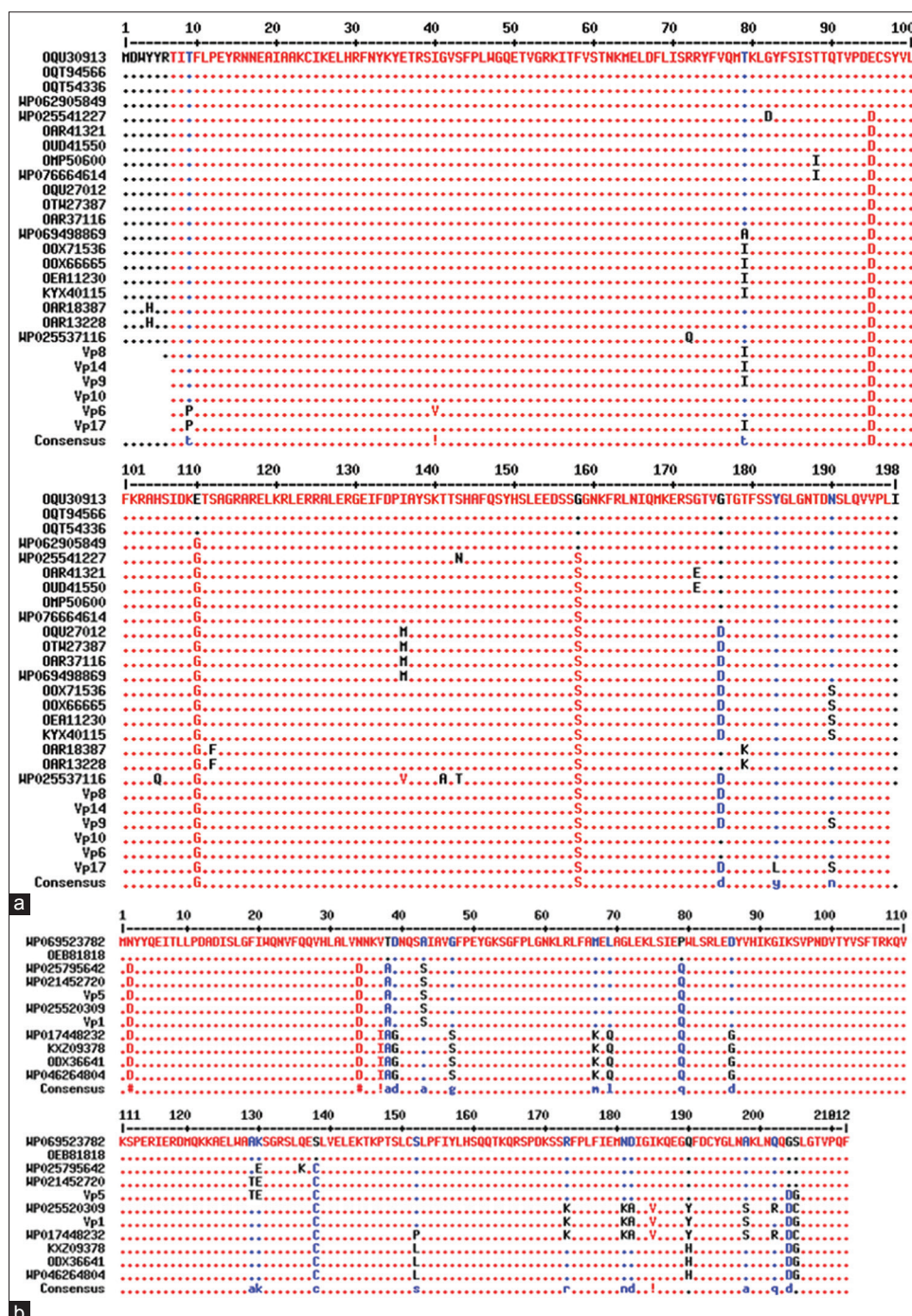


Figure-4: (a) Alignment of Cas6-a protein sequences of *Vibrio parahaemolyticus*. Vp6, Vp8-10, Vp14, and Vp17 represent environmental strains from this study. Amino acid variations among sequences are shown. (b) Alignment of Cas6-d protein sequences of *V. parahaemolyticus*. Vp1 and Vp5 represent environmental strains from this study. Amino acid variations among sequences are shown.

Conclusion

In prokaryotes harboring the CRISPR-cas type I-F system, the Cas6-associated protein functions as an endoribonuclease bringing about cleavage of the CRISPR RNAs and generation of pre-crRNA [16]. In *V. parahaemolyticus*, a previous study with *cas6* showed the gene sequences to be diverse and broadly classified into four genotypes [20]. The primers designed in this study could be used in the identification of the *cas6* genotypes in *V. parahaemolyticus*. Further, our studies with seafood and environmental samples show the *cas6*-a variant to be most prevalent. Altered Cas6 could probably impact endoribonuclease activity. Thus, the proper identification of *cas6* genotypes in strains of *V. parahaemolyticus* is needed, which could contribute in understanding further the impact of such altered genotypes on the CRISPR-Cas immune system of this pathogenic bacterium.

Authors' Contributions

MS and MNV designed the study. MS designed the primers and suggested necessary steps involved in the research throughout the study. PB carried out the experimental work. PB and MS drafted the manuscript. MNV corrected the manuscript. All authors read and approved the final manuscript.

Acknowledgments

The financial support from the Department of Biotechnology, Government of India, under the Bioinformatics Center program (BT/BI/04/049/99) is gratefully acknowledged.

Competing Interests

The authors declare that they have no competing interests.

Publisher's Note

Veterinary World remains neutral with regard to jurisdictional claims in published institutional affiliation.

References

- Morris, J. and Black, R. (1985) Cholera and other vibrioses in the United States. *N. Engl. J. Med.*, 312(6): 343-350.
- Li, H., Tang, R., Lou, Y., Cui, Z., Chen, W., Hong, Q., Zhang, Z., Malakar, P.K., Pan, Y. and Zhao, Y. (2017) A comprehensive epidemiological research for clinical *Vibrio parahaemolyticus* in Shanghai. *Front. Microbiol.*, 8: 1043.
- Ghenem, L., Elhadi, N., Alzahrani, F. and Nishibuchi, M. (2017) *Vibrio parahaemolyticus*: A review on distribution, pathogenesis, virulence determinants and epidemiology. *Saudi J. Med. Sci.*, 5(2): 93-103.
- Paranjpye, R., Hamel, O.S., Stojanovski, A. and Liermann, M. (2012) Genetic diversity of clinical and environmental *Vibrio parahaemolyticus* strains from the Pacific Northwest. *Appl. Environ. Microbiol.*, 78(24): 8631-8638.
- Wang, R., Zhong, Y., Gu, X., Yuan, J., Saeed, A.F. and Wang, S. (2015) The pathogenesis, detection, and prevention of *Vibrio parahaemolyticus*. *Front. Microbiol.*, 6: 144.
- Raghunath, P. (2015) Roles of thermostable direct hemolysin (TDH) and TDH-related hemolysin (TRH) in *Vibrio parahaemolyticus*. *Front. Microbiol.*, 5: 805.
- Sun, H., Li, Y., Shi, X., Lin, Y., Qiu, Y., Zhang, J., Liu, Y., Jiang, M., Zhang, Z., Chen, Q., Sun, Q. and Hu, Q. (2015) Association of CRISPR/Cas evolution with *Vibrio parahaemolyticus* virulence factors and genotypes. *Food Borne Pathog. Dis.*, 12(1): 68-73.
- Carte, J., Pfister, N., Compton, M.M., Terns, R.M. and Terns, M.P. (2010) Binding and cleavage of CRISPR RNA by *cas6*. *RNA*, 16(11): 2181-2188.
- Sokolowski, R.D., Graham, S. and White, M.F. (2014) *Cas6* specificity and CRISPR RNA loading in a complex CRISPR-Cas system. *Nucleic Acids Res.*, 42(10): 6532-6541.
- Hille, F. and Charpentier, E. (2016) CRISPR-Cas: Biology, mechanisms and relevance. *Philos. Trans. R. Soc. B.* 371(1707): 20150496.
- Rath, D., Amlinger, L., Rath, A. and Lundgren, M. (2015) The CRISPR-cas immune system: Biology, mechanisms and applications. *Biochimie*, 117: 119-128.
- Wakefield, N., Rajan, R. and Sontheimer, E.J. (2015) Primary processing of CRISPR RNA by the endonuclease *cas6* in *Staphylococcus epidermidis*. *FEBS Lett.*, 589(20): 3197-3204.
- Brendel, J., Stoll, B., Lange, S.J., Sharma, K., Lenz, C., Stachler, A.E., Maier, L.K., Richter, H., Nickel, L., Schmitz, R.A., Randau, L., Allers, T., Urlaub, H., Backofen, R. and Marchfelder, A. (2014) A complex of cas proteins 5, 6, and 7 is required for the biogenesis and stability of clustered regularly interspaced short palindromic repeats (CRISPR)-derived RNAs (crRNAs) in *Haloflex volcanii*. *J. Biol. Chem.*, 289(10): 7164-7177.
- Horvath, P. and Barrangou, R. (2010) CRISPR/Cas the immune system of bacteria and archaea. *Science*, 327(5962): 167-170.
- Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A. and Saunders, S.J. (2015) An updated evolutionary classification of CRISPR-cas systems. *Nat. Rev. Microbiol.*, 13(11): 722-736.
- Charpentier, E., Richter, H., van der Oost, J. and White, M.F. (2015) Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol. Rev.*, 39(3): 428-441.
- Sternberg, S.H., Haurwitz, R.E. and Doudna, J.A. (2012) Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA*, 18(4): 661-672.
- Wang, R. and Li, H. (2012) The mysterious RAMP proteins and their roles in small RNA-based immunity. *Protein Sci.*, 21(4): 463-470.
- Carte, J., Christopher, R.T., Smith, J.T., Olson, S., Barrangou, R., Moineau, S., Glover, C.V. 3rd, Graveley, B.R., Terns, R.M. and Terns, M.P. (2014) The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Mol. Microbiol.*, 93(1): 98-112.
- Baliga, P., Shekar, M. and Venugopal, M.N. (2018) Investigation of direct repeats, spacers and proteins associated with clustered regularly interspaced short palindromic repeat (CRISPR) system of *Vibrio parahaemolyticus*. *Mol. Genet. Genomics*, 294(1):253-262.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. (2012) Primer3-new capabilities and interfaces. *Nucleic Acids Res.*, 40(15): e115.
- Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.*, 16(22): 10881-10890.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, 28(10): 2731-2739.
- Gleditsch, D., Müller-Esparza, H., Pausch, P., Sharma, K., Dwarakanath, S., Urlaub, H., Bange, G. and Randau, L.

- (2016) Modulating the cascade architecture of a minimal Type I-F CRISPR-Cas system. *Nucleic Acids Res.*, 44(12): 5872-5882.
25. Wang, R., Preamplume, G., Terns, M.P., Terns, R.M. and Li, H. (2011) Interaction of the *cas6* ribonuclease with CRISPR RNAs: Recognition and cleavage. *Structure*, 19(2): 257-264.
26. Pourcel, C. and Drevet, R.C. (2013) Occurrence, diversity of CRISPR-Cas systems and genotyping implications. In: Barrangou, R. and van der Oost, J., editors. CRISPR-Cas Systems. Springer-Verlag, Berlin Heidelberg. p33-59.
27. Haurwitz, R.E., Jinek, M., Wiedenhef, B., Zhou, K. and Doudna, J.A. (2010) Sequence-and structure-specific RNA processing by a CRISPR endonuclease. *Science*, 329(5997): 1355-1358.
28. Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin E.V. and van der Oost J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, 321(5891): 960-964.
29. Niewoehner, O., Jinek, M. and Doudna, J.A. (2014) Evolution of CRISPR RNA recognition and processing by *cas6* endonucleases. *Nucleic Acids Res.*, 42(2): 1341-1353.
30. Takeuchi, N., Wolf, Y.I., Makarova, K.S. and Koonin, E.V. (2012) Nature and intensity of selection pressure on CRISPR-associated genes. *J. Bacteriol.*, 194(5): 1216-1225.
