Thesis for the Degree of Master of Science

Genome and Transcriptome Analyses of *Vibrio vulnificus* FORC_036 Isolated from a Surf Clam

동죽에서 분리한 패혈증 비브리오균 FORC_036에 대한 유전체 및 전사체 분석

February, 2017

So Jeong Gil

Department of Agricultural Biotechnology
College of Agriculture and Life Sciences
Seoul National University

석사학위논문

Genome and Transcriptome Analyses of *Vibrio vulnificus* FORC_036 Isolated from a Surf Clam

지도교수 최 상 호

이 논문을 석사학위논문으로 제출함

2017년 2월

서울대학교 대학원

농생명공학부

길 소 정

길소정의 석사학위논문을 인준함

2017년 2월

위원장 최 영 진 (인)

부위원장 최 상 호 (인)

위원 이 기 원 (인)

Abstract

Vibrio vulnificus is a Gram-negative, curved rod-shaped marine pathogen that causes from the gastrointestinal diseases to the primary septicemia whose mortality rate is approximately 50%. Its infection happens through the consumption of raw seafood or exposure of wound to contaminated seawater. To characterize *V. vulnificus* that isolated from a surf clam, the genome and transcriptome of this strain were analyzed. Virulence gene-specific polymerase chain reaction (PCR) screening and lactate dehydrogenase (LDH) release assay were conducted to evaluate the virulence of this strain. It showed a high level of cytotoxicity which was similar to cytotoxicity of *V. vulnificus* MO6-24/O, a clinical isolate and had many virulence genes. So, it was called as 'FORC_036' and further analyzed.

The genome of the strain FORC_036 consisted of three conitgs; two chromosomes and one plasmid. Its genome was predicted to have 5,642 open reading frames (ORFs), 203 tRNA (transfer RNA) genes and 34 rRNA (ribosomal RNA) genes. Also, *V. vulnificus* FORC_036 had a plenty of virulence genes including toxins encoding genes, secretion systems related genes, and iron uptake systems related genes. The phylogenetic analysis using 16S rRNA from *Vibrio* species was demonstrated the strain FORC_036 was grouped with other *V. vulnificus* strains. And the average nucleotide identity (ANI) analysis using the complete genome of *V. vulnificus* strains showed that a clinical isolates, CMCP6 was the closest strain to the strain FORC_036. The comparative genome analysis with *V. vulnificus* CMCP6 suggested that *V.*

vulnificus FORC_036 had additional genes; (1) L-fucose utilization related genes and (2) thermostable hemolysin VPH-delta encoding gene.

To figure out the transcriptional response when the strain FORC_036 was exposed to the model food (small octopus), RNA sequencing was conducted. The result showed that the expression of the motility related genes was down-regulated and the genes related to the adherence, galactose metabolism, and iron uptake were up-regulated. However, the expression of the toxin related genes was remained. Therefore, the strain FORC_036 recognized the small octopus as a reservoir rather than the host.

All results described above demonstrated that *V. vulnificus* FORC_036 isolated from a surf clam was possibly pathogenic to human and used the small octopus as a reservoir to survive and grow. Through this study, the information of the pathogen isolated from environment was accumulated and it helps to deal with *V. vulnificus*-related outbreaks.

Key words: *Vibrio vulnificus*, Food-borne pathogen, Genomics, Transcriptomics, Whole genome sequencing, RNA sequencing, Comparative genome analysis

Student Number: 2015-21758

Contents

List of Figures

List of Tables

I. INTRODUCTION

Vibrio vulnificus is one of marine bacteria and an opportunistic Gram-negative pathogen. Consumption of raw contaminated seafood (primarily oyster) or exposure of wound to seawater where *V. vulnificus* is present can cause a primary septicemia and wound infection (Storm and Paranjpye, 2000; Gulig *et al.*, 2005). The average mortality rate of the primary septicemia caused by *V. vulnificus* is approximately 50% (Jones and Oliver, 2009). In South Korea, during 2011-2015, 269 individuals were infected with *V. vulnificus*. Among them, 147 (54.65%) patients died (Korea Centers for Disease Control and Prevention, http://is.cdc.go.kr/dstat/jsp/stat/stat0001.jsp). This means that this pathogen is widespread and highly virulent. Actually, many outbreaks of *V. vulnificus* have been reported for many years. Therefore, it is important to understand the virulence traits of *V. vulnificus*.

Some people in South Korea consume the raw small octopus (*Octopus minor*). In 2015, a man consuming the raw small octopus got septicemia caused by *V. vulnificus*. So, understanding the behavior of *V. vulnificus* when this pathogen is exposed to the small octopus provides important insight into figuring out the survival and pathogenesis of the *V. vulnificus* in the small octopus.

Currently, the emergence of the high throughput sequencing technologies makes the bacterial Whole Genome Sequencing (WGS) faster, cheaper and easier. It can be used to identify the species, test its properties (for example, antibiotics resistance and virulence), and monitor the emergence and spreading of bacterial pathogens (Didelot

et al., 2012). Many researches about the physiological and molecular analysis of *V. vulnificus* are carried out. On the other hand, the study about *V. vulnificus* genome is insufficient. Actually, only 8 strains had complete genome sequence and were analyzed to date (National Center for Biotechnology Information; NCBI). Therefore, more complete genome sequences of this pathogen are needed and analyzed to understand the pathogenicity and virulence of *V. vulnificus* further.

RNA sequencing is a developed approach to provide the information of transcripts in a specific condition. Compared with other methods, this technique has lots of advantages. It has high throughput and low background noise (Wang, Gerstein and Snyder, 2009). As the sensitivity of RNA sequencing increases, it is possible to detect and quantify of RNA expressed at low level (Pinto *et al.*, 2011).

In this study, to extend our knowledge about the pathogenesis of *V. vulnificus*, *V. vulnificus* FORC_036 was isolated from a surf clam (*Mactra veneriformis*) by Gyeonggi Health and Environmental Institute and Ministry of Drug and Food Safety of South Korea and its genome was completely sequenced and analyzed using various bioinformatics programs. Additionally, the comparative genome analysis between the strain FORC_036 and the other isolate whose complete genome sequence was available was performed. Also, the behavior and transcriptome profiling of *V. vulnificus* FORC_036 were revealed when this strain was exposed to the small octopus.

The results of this study provide the novel complete genome sequence and associated information of *V. vulnificus* and are useful for control of *V. vulnificus* FORC_036.

II. MATERIALS AND METHODS

Strains and growth condition

The strains used in this study are listed in Table 1. *V. vulnificus* strains were incubated at 30°C with modified Luria-Bertani (LB) medium supplemented with 1% (the final concentration is 2%) (*w/v*) sodium chloride (NaCl) for 12 h aerobically.

Genomic DNA extraction and virulence gene-specific PCR

The genomic DNA was extracted from the culture using DNeasy Blood & Tissue Kit (QIAGEN, Valencia, California, USA), according to the manufacturers' protocol. The virulence genes (*vvhA, vvpE, vcgC, rtxA, nanA* and *gyrB*) were amplified from extracted genomic DNA of *V. vulnificus* strains using various pairs of primers (listed in Table 2). The amplified DNA fragments were resolved on a 1.5% agarose gel. The result of the electrophoresis was visualized using Geldoc™ EZ Image (Bio-Rad, Richmond, California, USA).

Cytotoxicity test

The cytotoxicity of the *V. vulnificus* strains was indirectly estimated by measuring the cytoplasmic lactate dehydrogenase (LDH) activity. LDH is released from the human cell when the plasma membrane is damaged. LDH release assay was conducted using INT-407 human epithelial cells (ATCC® CCL-6™). INT-407 cells preparation and infection with the bacterial cultures were carried out in a 96-well tissue culture plate

(Nunc, Roskilde, Denmark) as described previously (Jeong *et al*., 2000). After infection, the LDH activity in supernatant was measured by using a Cytotoxicity Detection Kit (Roche, Mannheim, Germany). And the cytotoxicity was expressed using the LDH activity which was calculated using a following formula. The positive control (completely lysed) is the well added 2% of Triton X-100 and the negative control (non-lysed) is the well added the assay media instead of the bacterial culture. The experiments were repeated with biologically duplicated samples and triplicate wells were run for each sample.

$$\text{LDH activity (\%)} = \frac{\text{Experimental value - Negative control value}}{\text{Positive control value - Negative control value}} \times 100\ (\%)$$

Transmission electron microscope

V. vulnificus FORC_036 cells were negatively stained with uranyl acetate for 1 min. The stained cells were observed using TEM JEM-2100 (JEOL, Tokyo, Japan) at 200 kV.

Genome sequencing and annotation

The whole genome sequencing (WGS) and assembly were performed at ChunLab Incorporation (Seoul, South Korea). The sequences of *V. vulnificus* FORC_036 were acquired utilizing two different platforms. One was Illumina MiSeq platform (Illumina, San Diego, California, USA), and the other was PacBio platform (Pacific Biosciences, Menlo Park, California, USA). The obtained sequence reads were also

assembled with two assembly programs. The raw sequences from Illumina MiSeq platform were assembled by CLC Genomics workbench 7.5.1 (CLC Bio, Aarhus, Denmark), and the raw sequences from PacBio platform were assembled by PacBio SMRT Analysis 2.3.0 (Pacific Biosciences), respectively.

The prediction of ORFs and their annotation were conducted using the GeneMarkS program (Besemer, Lomsadze, and Borodovsky, 2001) and the rapid annotations using subsytems technology (RAST) server (Aziz *et al*., 2008). The ribosomal binding sites (RBSs) were predicted using the RBS finder (J. Craig Venter Institute, Rockville, Maryland, USA). Additionally, the prediction of the function of ORFs and their conserved protein domains were performed using two programs; InterProScan 5 (Jones *et al*., 2014) and GAMOLA (Altermann and Klaenhammer, 2003). Also, the prediction and characterization of putative virulence factors of this strain were conducted using the basic local alignment search tool (BLAST) and the virulence factor database (VFDB; http://www.mgc.ac/VFs/). Lastly, according to the acquired result, the circular genome maps of the strain FORC_036 were drawn utilizing the cluster of orthologous group (COG)-based WebMGA program and the GenVision program (DNASTAR, Madison, Wisconsin, USA). The final results were submitted to GenBank and received the accession numbers.

Phylogenetic analysis and comparative genome analysis

Through 16S rRNA sequencing, the strain FORC_036 was identified as *V. vulnificus*. And phylogenetic analysis using MEGA7 (Kumar, Strecher, and Tamura, 2016) was

performed using the 16S rRNA sequences of *Vibrio* species and *Shewanella baltica* OS678 as an out-group. The phylogenetic tree was created from the aligned sequences. The neighbor joining method was used and bootstrap replicates was 1000. For characterization the evolutionary relationship in *V. vulnificus*, average nucleotide identity (ANI) analysis was also conducted. The completed genome sequences of 6 strains including MO6-24/O, CMCP6, 93U204, YJ016, FORC_009, and FORC_036 were used. The genome sequences were obtained from the NCBI. The ANI values were calculated using Jspecies program (Richter and Móra, 2009) comparing the 1020 bp-fragmented sequences between two full genomes. And the result was visualized as a phylogenetic tree by using the R program. With the closest strain, the comparative genome analysis between two strains was carried out using the Artemis comparison tool (ACT) (Carver *et al*., 2005).

RNA extraction

V. vulnificus FORC_036, an environmental isolate, was grown to mid-log phase (OD_{600} of 0.8). The culture was transferred to 50 ml tubes (SPL, Gyeonggi, South Korea) and centrifuged at 1,500 x g for 10 min. The supernatant was removed and the pellet was washed with 1X phosphate buffered saline (PBS). And then, the solution was centrifuged at 5,000 x g for 10 min. This step was repeated two or three times. After final wash step, the pellet was resuspended with pre-warmed *Vibrio fisheri* minimal medium containing glycerol (VFMG) medium. Prepared *V. vulnificus* FORC_036 was incubated in the VFMG medium with the presence or absence of the

small octopus at 30°C for 4 h. After incubation, the culture was filtrated with syringe, sterilized gauze, and vacuum filter with Whatman no. 1 filter paper (Whatman International Ltd, Maidstone, England). The filtrated culture was transferred to 50 ml tubes (SPL) and centrifuged at 5,000 x g at 4°C for 10 min. The pellet was resuspended with 0.5 ml of cold-diethylpyrocarbonate (DEPC) treated PBS and the solution was quickly mixed with 1 ml of RNAprotect® Bacteria Reagent (RBR) (QIAGEN).

Total RNAs were extracted from RBR treated samples using miRNeasy Mini Kit (QIAGEN) according to the manufacturers' protocol. DNAs were eliminated by TURBO DNase (AMbion, Austin, Texas, USA) to protect contamination. And then, RNAs were cleaned up using RNeasy MinElute Cleanup Kit (QIAGEN). The quality of RNAs were confirmed by ChunLab Incorporation using Agilent 2100 Bioanalyzer and Agilent RNA 6000 Nano reagents (Agilent Technologies, Waldbronn, Germany).

Strand-specific cDNA library construction and RNA sequencing

The process for strand-specific complementary DNA (cDNA) library construction and RNA sequencing were performed by ChunLab Incorporation. First, mRNA was selectively enriched by depleting rRNAs using Ribo-Zero™ rRNA Removal Kit (Epicentre, Madison, Wisconsin, USA). Enriched mRNA was subject to construction of the cDNA library. TruSeq Stranded mRNA Sample Preparation kit (Illumina) was used in this process following the manufacturers' procedure. The quality of cDNA libraries was checked as described above for the quality confirmation of total RNA,

except that Agilent DNA 1000 reagents was used (Agilent Technologies). Strand-specific paired-ended 100 nucleotide reads from each cDNA library were obtained using Illumina HiSeq2500. Two cDNA libraries were constructed for biological replication. These libraries were sequenced from RNAs isolated from two independent filtered culture supernatants of *V. vulnificus* FORC_036.

Transcriptomic data analysis

The reads obtained from RNA sequencing were mapped to the *V. vulnificus* FORC_036 reference genome using CLC Genomics Workbench 5.5.1 (CLC Bio). The GenBank (http://www.ncbi.nlm.nih.gov) accession numbers of the reference genome are CP015512 (Chromosome I), CP015513 (Chromosome II), and CP015514 (Plasmid). The relative transcript abundance was measured by the reads per kilobase (kb) of transcript per million mapped sequence read (RPKM) (Mortazavi *et al*., 2008). As the fold changes of RPKM values and their significance were assigned, the genes were considered to be differentially expressed in samples when the fold change was 2 or greater and *P*-value was < 0.05. The volcano plot and the heat map were created to visualize the RNA sequencing result using CLC Genomics Workbench 5.5.1 (CLC Bio) and Gitools (Biomedical Research Park, Barcelona, Spain; Perez-Llamas and Lopez-Bigas, 2011), respectively.

RNA purification and transcript analysis

Total cellular RNAs from the strain FORC_036 were isolated using RNeasy Mini Kit

(QIAGEN). The culture condition was the same as described previously (See the 'RNA extraction' part). And cDNA was synthesized using iScripit™ cDNA Synthesis Kit (Bio-Rad). The quantitative real-time PCR (qRT-PCR) of the synthesized cDNA was conducted using iQ™ SYBR® Green Supermix (Bio-rad) and the Chromo 4 Real-Time PCR detection system (Bio-rad) as described previously (Lim and Choi, 2014). The represented genes and their specific primers used in this process are listed in Table 2. The relative expression levels of the specific transcripts were calculated by using the 16S rRNA expression level as the internal reference for normalization. The experiments were repeated with biologically duplicated samples and technically triplicate wells were run for each sample.

Growth kinetics of *V. vulnificus* FORC_036

The growth condition and bacteria preparation step was the same as described above (See the 'RNA extraction' part), except the incubation times; 0 h, 2 h, 4 h, and 6 h. After filtration step, the culture was transferred to 50 ml tubes (SPL). The filtrated culture was serially diluted by 1X PBS and plated on LB agar supplemented with 1% (final concentration is 2%) (*w/v*) sodium chloride (NaCl). The plates were incubated in 30°C for 12 h and the colonies were counted. The experiments were repeated with biologically duplicated samples and technically triplicate were conducted for each sample.

Table 1. Bacterial strains used in this study

Bacterial strain	Relevant characteristics	Reference or Source
V. vulnificus		
MO6-24/O	Clinical isolate, virulent	Laboratory collection
1001957 [a]	Environmental isolate	Gyeonggi Health and Environmental Institute

[a] This strain 1001957 is called as 'FORC_036' based on further results.

Table 2. Oligonucleotides used in this study

Oligonucleotide	Sequence (5` → 3`)	Use	Predicted size (bp)
For virulence gene-specific PCR screening [a]			
vvhA_F	GACTATCGCATCAACAACCG	PCR analysis of *vvhA*	704
vvhA_R	AGGTAGCGAGTATTACTGCC		
vvpE_F	AGGCAGTGTTGACTGGATCG	PCR analysis of *vvpE*	120
vvpE_R	ATGGAGCGGCCATCTTTTGA		
vcgC_F	CGCCTTTGTCAGTGTTGCA	PCR analysis of *vcgC*	223
vcgC_R	TAACGCGAGTAGTGAGCCG		
rtxA5_F	AACGCCAATATTCACGTGGG	PCR analysis of *rtxA*	736
rtxA5_R	AAGCAGAGGCATCACCAGAC		
rtxA6_F	GCCAATATTCACGTGGGCGA	PCR analysis of *rtxA*	734
rtxA6_R	CAAGCAGAGGCATCACCAGA		
nanA_F	TTATCGCCGCTCCCCATACA	PCR analysis of *nanA*	745
nanA_R	GCAACGCCACCGTATTCAAC		
gyrB_F	GTCCGCAGTGGAATCCTTCA	PCR analysis of *gyrB*	285

Oligonucleotide	Sequence (5` → 3`)	Use	Predicted size (bp)
gyrB_R	TGGTTCTTACGGTTACGGCC	PCR analysis of *gyrB*	285
For qRT-PCR [b]			
FORC36_0520_qRT_F	TCTAATTGCTGACGTTAAACACG	qRT-PCR of FOR36_0520	198
FORC36_0520_qRT_R	TGCTTCTGGTGGAAGTTCAG		
FORC36_0727_qRT_F	GGTGGTCTAGAGCCCATTG	qRT-PCR of FOR36_0727	114
FORC36_0727_qRT_R	TAGCATGGATTGGCGCACT		
FORC36_0734_qRT_F	GCGAGTGCTCTACAAGACGA	qRT-PCR of FOR36_0734	192
FORC36_0734_qRT_R	TTCAGACTGGAAGGTCGTGC		
FORC36_0938_qRT_F	CAACGCAGCATTCGGTTCAT	qRT-PCR of FOR36_0938	144
FORC36_0938_qRT_R	GGTGTTGCCGCATTTGAAGT		
FORC36_1250_qRT_F	GGGACCAAAAGCTCGCTACT	qRT-PCR of FOR36_1250	151
FORC36_1250_qRT_R	TGAAACCATTTCGCCGATGC		
FORC36_1288_qRT_F	ATGCAGCAGGTAACGAAGTCA	qRT-PCR of FOR36_1288	194
FORC36_1288_qRT_R	CACCAAGGTCACAATGCCAG		
FORC36_1691_qRT_F	GCGACAAGGCAAGTTGTTGA	qRT-PCR of FOR36_1691	123
FORC36_1691_qRT_R	TCCACCAGCGGAGTTGATTC		

Oligonucleotide	Sequence (5` → 3`)	Use	Predicted size (bp)
For qRT-PCR [b]			
FORC36_2260_qRT_F	ACCTCGAAGGCACTGAAACC	qRT-PCR of FOR36_2260	200
FORC36_2260_qRT_R	CAGCACCACCTTTTTCACCG		
FORC36_2358_qRT_F	CAGCAATGGCTGCTGAAGAG	qRT-PCR of FOR36_2358	148
FORC36_2358_qRT_R	TCTTGCGCGAGTTTTTCTGC		
FORC36_3267_qRT_F	CGCTCAGTGTGGGCATCTAT	qRT-PCR of FOR36_3267	120
FORC36_3267_qRT_R	AGCCGTTACTGTGTCCATCG		
FORC36_3328_qRT_F	CCGCTTCTCAAGGAGTCAGT	qRT-PCR of FOR36_3328	199
FORC36_3328_qRT_R	TCAAGCATTTTTGCGCCACT		
FORC36_3773_qRT_F	GCGCTGTTTTCGGTTTACGG	qRT-PCR of FOR36_3773	112
FORC36_3773_qRT_R	TAACGTGTGCTTCCGCTTCA		
FORC36_3861_qRT_F	GTCTTGCAGACCGTGCTACT	qRT-PCR of FOR36_3861	131
FORC36_3861_qRT_R	AGCAACGTACTGAGCTGCTT		
FORC36_4377_qRT_F	GCCTACCTTCCCCAACATCC	qRT-PCR of FOR36_4377	133
FORC36_4377_qRT_R	TGGCTTCATCGACAATCGCA		
FORC36_4386_qRT_F	CGTATGCAAGCGTCTGGTTG	qRT-PCR of FOR36_4386	130

Oligonucleotide	Sequence (5` → 3`)	Use	Predicted size (bp)
For qRT-PCR [b]			
FORC36_4386_qRT_R	TGACAAGGAACGGGTGAGTG	qRT-PCR of FOR36_4386	130
FORC36_4406_qRT_F	CGCCTTTGTGCCTAGCG	qRT-PCR of FOR36_4406	105
FORC36_4406_qRT_R	CCATTGTAGACAGAGCCAGC		
FORC36_5121_qRT_F	CAGGCTGACAGTGAACACCA	qRT-PCR of FOR36_5121	145
FORC36_5121_qRT_R	AGTTACGGTGTCACCCGTTT		

[a] The oligonucleotides were designed using *V. vulnificus* MO6-24/O genome sequence (GenBank™ accession number CP002469 and CP002470).

[b] The oligonucleotide were designed using *V. vulnificus* FORC_36 genome sequence (GenBank™ accession number CP015512, CP015513 and CP015514).

III. RESULT

Virulence gene-specific PCR screening

To check the virulent possibility of this environmental strain, various virulence genes of *V. vulnificus* were amplified using PCR. *V. vulnificus* MO6-24/O was used as a positive control of this experiment. The target virulence genes and their expected sizes were listed in Table 2. All 6 genes were detected in both strains (Fig. 1).

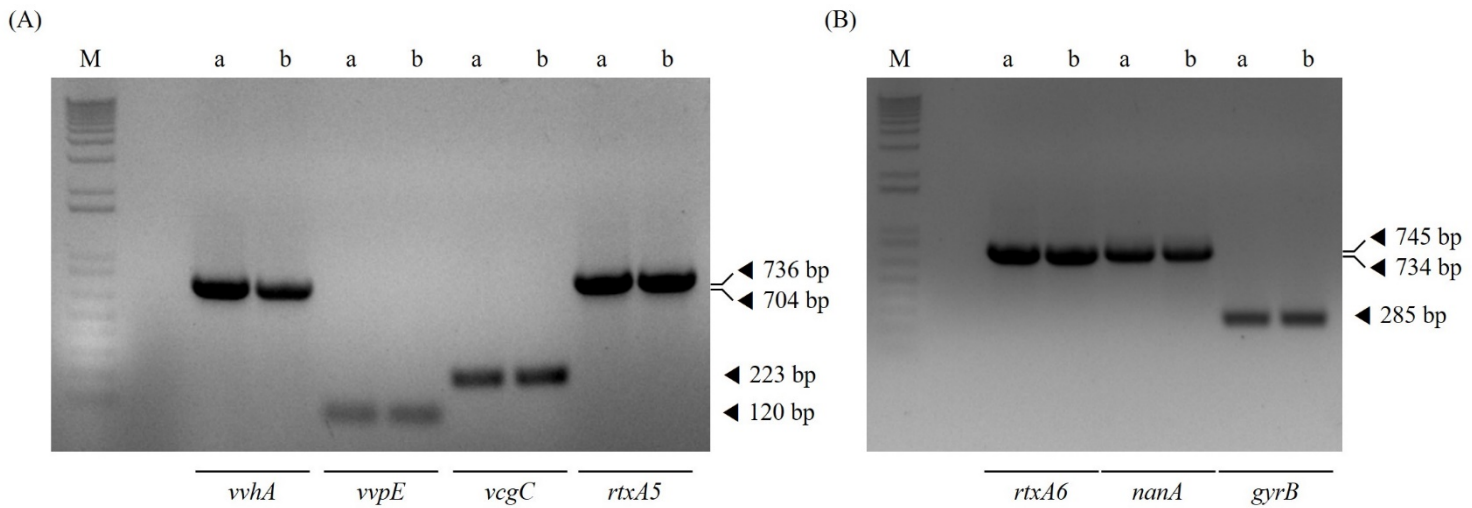

Fig 1. Virulence gene-specific PCR screening. 6 virulence genes (7 pairs of primers, 2 primers for *rtxA*) of *V. vulnificus* were amplified using PCR. The size of each band was written beside the figure. M, Molecular weight marker; a, *V. vulnificus* MO6-24/O, a positive control; b, *V. vulnificus* 1001957; *vvhA*, hemolysin encoding gene; *vvpE*, elastase encoding gene; *vcgC*, virulence correlated gene; *rtxA*, repeated toxin protein A encoding gene; *nanA*, *N*-acetylneuraminate lyase encoding gene; *gyrB*, gyrase B encoding gene.

Cytotoxicity analysis of *V. vulnificus*

Because this strain was isolated from environment, the experiment which evaluates the virulence of this stain was needed. Therefore, LDH release assay was performed using INT-407 cell lines. The host cells were infected with this strain and *V. vulnificus* MO6-24/O (as a positive control; a clinical isolate) at various multiplicity of infections (MOIs) for 2 h. And the LDH activity released from the host cells which were infected with *V. vulnificus* 1001957 was compared with a positive control. The result showed that the LDH activity of the experimental strain was as high as that of *V. vulnificus* MO6-24/O (Fig. 2). From now on, this strain (1001957) is called as 'FORC_036'.

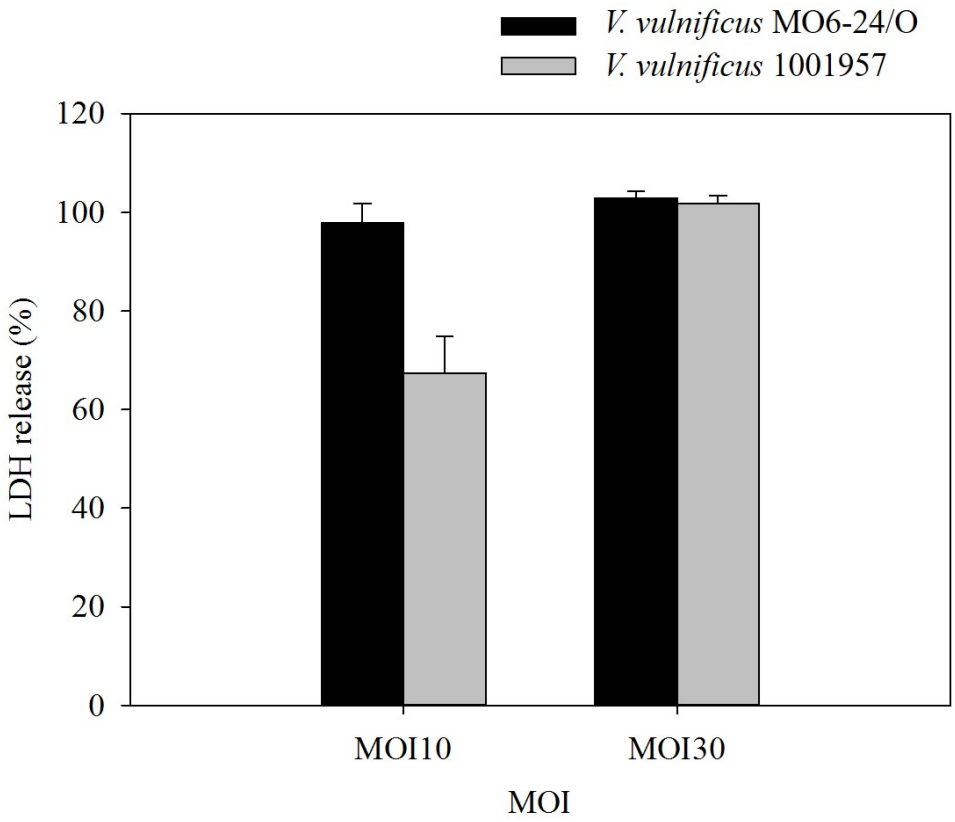

Fig 2. Cytotoxicity analysis of *V. vulnificus* strains. INT-407 cells were infected with *V. vulnificus* stains at various MOIs for 2 h. The cytotoxicity was determined by the LDH release activity. The total LDH (100%) is the amount of LDH released from the cells which were completely lysed by 2% Triton X-100. The experiments were repeated with biologically duplicated samples and triplicate wells were run for each sample. The error bars represented the standard deviations (SD).

Transmission electron microscopy image

The morphology of *V. vulnificus* FORC_036 was checked through the transmission electron microscope (TEM). The strain FORC_036 is a curved and rod-shaped bacterium with a single polar flagellum. The length was 1.16μm and the width was 0.76μm (Fig. 3).

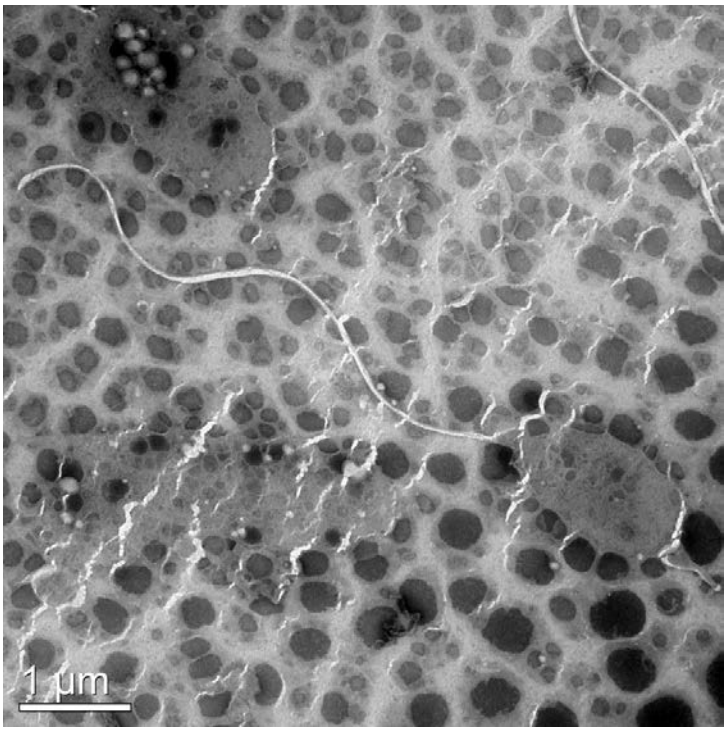

Fig 3. Transmission electron microscopy image of *V. vulnificus* FORC_036. The cells were negatively stained with 2% uranyl acetate. Bars, 1 μm.

Genome properties of *V. vulnificus* FORC_036

The information of the WGS was summarized in Table 3. The genome of *V. vulnificus* FORC_036 had two circular chromosomes and one large plasmid. Chromosome I consists of 3,260,504 bp with a GC content of 46.63% containing 2,922 predicted ORFs, 105 tRNA genes and 31 rRNA genes. Among these predicted ORFs, 2,388 ORFs (81.72%) were predicted to encode the functional protein and 534 ORFs (18.28%) were expected to encode hypothetical proteins. Chromosome II consists of 1,882,082 bp with a GC content of 47.08% containing 1,609 predicted ORFs, 14 tRNA genes, and 3 rRNA genes. Among them, 1,266 ORFs (78.68%) were predicted to be functional and 343 ORFs (21.32%) were expected to encode hypothetical proteins. The large plasmid, called pFORC36, consists of 925,374 bp with a GC content of 37.35% containing 1,111 predicted ORFs and 84 tRNA genes. Only 209 ORFs (18.81%) were predicted to encode the functional protein and remaining 902 ORFs (81.19%) were expected to encode hypothetical proteins. All the results are summarized in Table 4.

Based on the bioinformatics analysis of two chromosomes and a plasmid, circular genome maps were drawn (Fig. 4) and genes with the specialized function such as putative virulence factors were expressed in each genome map.

Table 3. Summary of *V. vulnificus* FORC_036 genome sequencing

Property	Term
Finishing quality	Finished
Libraries used	Illumina 300 base pair paired-end library
	PacBio SMRTbell™ library (> 10 kb) for draft assembly
Sequencing platform	Illumina MiSeq
	PacBio
Assembler	CLC Genomics Workbench 7.5.1
	PacBio SMRT Analysis 2.3.0
Gene calling method	GeneMarkS
	RAST server
Average genome coverage	371.5 X
Contig length (bp)	3,260,504 (Chromosome I)
	1,882,082 (Chromosome II)
	925,374 (Plasmid)
Contig No.	3
Scaffold No.	3
N50	3,260,504
Locus tag	FORC36
GenBank accession No.	CP015512, CP015513, CP015514
GenBank release date	2017-04-28
BioProject No.	PRJNA320042
Source material identifier	FORC_036
Project relevance	Agricultural

Table 4. Chromosomal properties of *V. vulnificus* FORC_036

Property	Term		
Sample	*V. vulnificus* FORC_036		
Classification	Chromosome I	Chromosome II	Plasmid
GenBank accession No.	CP015512	CP015513	CP015514
Genome size (bp)	3,260,504	1,882,082	925,374
Protein coding genes	2922	1609	1111
Annotated genes	2388	126	209
Hypothetical genes	534	343	902
tRNA	105	14	84
rRNA	31	3	0

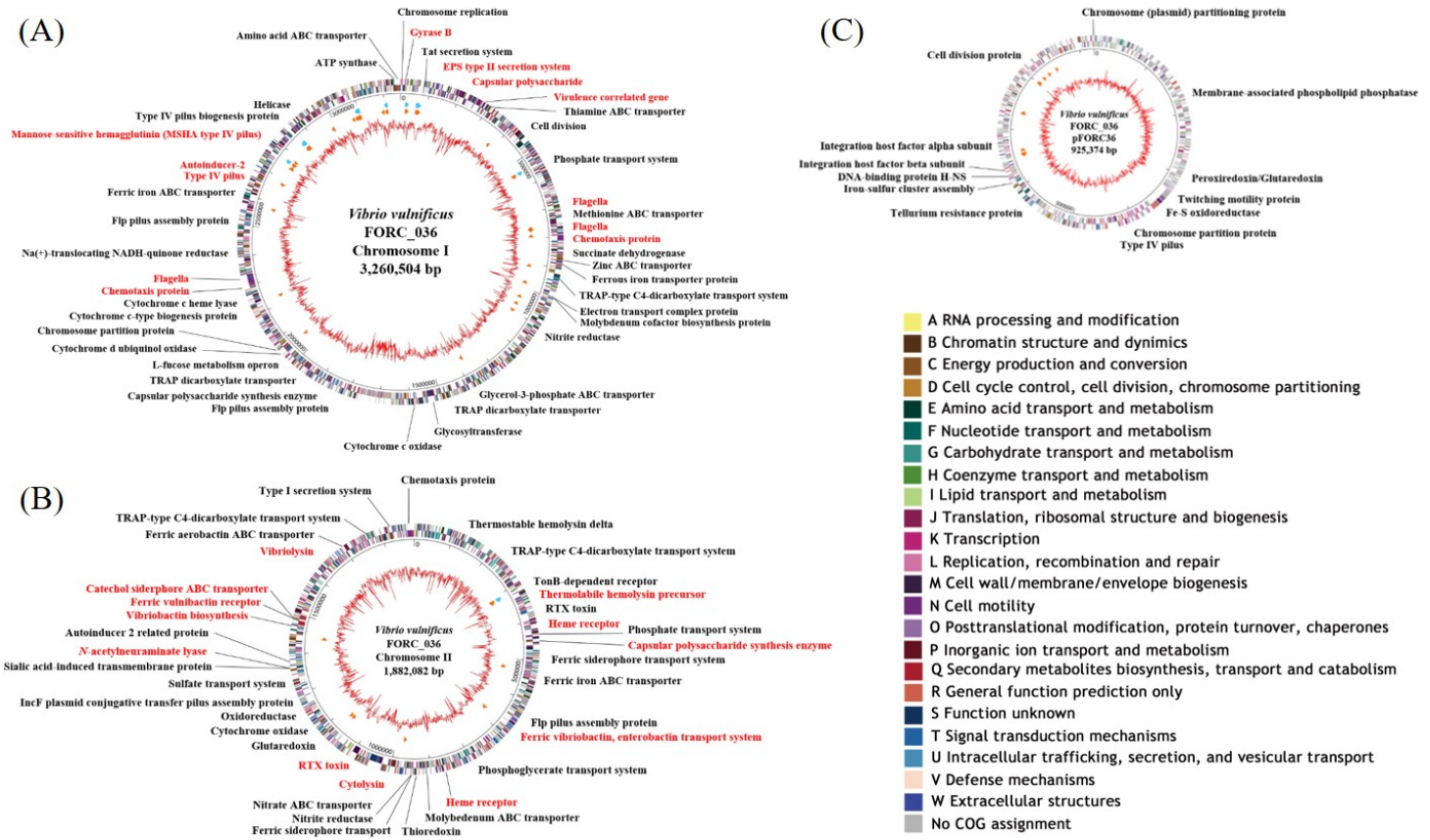
(A)
Chromosome replication
Amino acid ABC transporter
Gyrase B
ATP synthase
Tat secretion system
EPS type II secretion system
Capsular polysaccharide
Virulence correlated gene
Helicase
Thiamine ABC transporter
Type IV pilus biogenesis protein
Cell division
Mannose-sensitive hemagglutinin (MSHA type IV pilus)
Phosphate transport system
Autoinducer-2
Type IV pilus
Ferric iron ABC transporter
Flagella
Methionine ABC transporter
Flp pilus assembly protein
Flagella
Chemotaxis protein
Na(+)-translocating NADH-quinone reductase
Succinate dehydrogenase
Zinc ABC transporter
Flagella
Ferrous iron transporter protein
Chemotaxis protein
TRAP-type C4-dicarboxylate transport system
Cytochrome c heme lyase
Electron transport complex protein
Cytochrome c-type biogenesis protein
Molybdenum cofactor biosynthesis protein
Chromosome partition protein
Nitrite reductase
Cytochrome d ubiquinol oxidase
L-fucose metabolism operon
Glycerol-3-phosphate ABC transporter
TRAP dicarboxylate transporter
Capsular polysaccharide synthesis enzyme
TRAP dicarboxylate transporter
Flp pilus assembly protein
Glycosyltransferase
Cytochrome c oxidase
Vibrio vulnificus
FORC_036
Chromosome I
3,260,504 bp
(B)
Chemotaxis protein
Type I secretion system
TRAP-type C4-dicarboxylate transport system
Thermostable hemolysin delta
Ferric aerobactin ABC transporter
TRAP-type C4-dicarboxylate transport system
Vibriolysin
TonB-dependent receptor
Catechol siderphore ABC transporter
Thermolabile hemolysin precursor
Ferric vulnibactin receptor
RTX toxin
Vibriobactin biosynthesis
Heme receptor
Phosphate transport system
Autoinducer 2 related protein
Capsular polysaccharide synthesis enzyme
N-acetylneuraminate lyase
Ferric siderophore transport system
Sialic acid-induced transmembrane protein
Ferric iron ABC transporter
Sulfate transport system
IncF plasmid conjugative transfer pilus assembly protein
Flp pilus assembly protein
Oxidoreductase
Ferric vibriobactin, enterobactin transport system
Cytochrome oxidase
Glutaredoxin
RTX toxin
Phosphoglycerate transport system
Cytolysin
Nitrate ABC transporter
Heme receptor
Nitrite reductase
Molybdenum ABC transporter
Ferric siderophore transport
Thioredoxin
Vibrio vulnificus
FORC_036
Chromosome II
1,882,082 bp
(C)
Chromosome (plasmid) partitioning protein
Cell division protein
Membrane-associated phospholipid phosphatase
Vibrio vulnificus
FORC_036
pFORC36
925,374 bp
Integration host factor alpha subunit
Integration host factor beta subunit
DNA-binding protein H-NS
Peroxiredoxin/Glutaredoxin
Iron-sulfur cluster assembly
Twitching motility protein
Tellurium resistance protein
Fe-S oxidoreductase
Chromosome partition protein
Type IV pilus
A RNA processing and modification
B Chromatin structure and dynimics
C Energy production and conversion
D Cell cycle control, cell division, chromosome partitioning
E Amino acid transport and metabolism
F Nucleotide transport and metabolism
G Carbohydrate transport and metabolism
H Coenzyme transport and metabolism
I Lipid transport and metabolism
J Translation, ribosomal structure and biogenesis
K Transcription
L Replication, recombination and repair
M Cell wall/membrane/envelope biogenesis
N Cell motility
O Posttranslational modification, protein turnover, chaperones
P Inorganic ion transport and metabolism
Q Secondary metabolites biosynthesis, transport and catabolism
R General function prediction only
S Function unknown
T Signal transduction mechanisms
U Intracellular trafficking, secretion, and vesicular transport
V Defense mechanisms
W Extracellular structures
No COG assignment

Fig 4. Genome map of *V. vulnificus* FORC_036. (A) Chromosome I; (B) Chromosome II; (C) pFORC36. The outer circle indicates the locations of all annotated ORFs, and the inner circle with red peaks indicates the GC contents. Between two circles, the orange arrows are the tRNAs and the sky blue arrows are the rRNA. All annotated ORFs were colored according to the COG assignments. The information of each color was shown at the right side of this figure. The virulence-related genes were labeled with red and other functional genes were labeled with black.

Pathogenesis and virulence factor

V. vulnificus is one of the major food-borne pathogens and causes the primary septicemia. Therefore, it has various virulence factors in its genome.

V. vulnificus FORC_036 had major virulence related genes such as *vvhA* gene (FORC36_3773 in chromosome II) and *rtx* genes (FORC36_3840, 3841, 3843, 3844 in chromosome II). And, this microorganism had the other hemolysin encoding gene (FORC36_3212 in chromosome II). Also, it had type II secretion system (T2SS). T2SS related genes were found in the specific region of chromosome I in this strain (from FORC36_0092 to FORC36_0103). Additionally, the genome of the strain FORC_036 included various iron uptake related genes; heme receptor encoding gene (FORC36_3267 in chromosome II), iron ABC transporter related genes (FORC36_3503, 3504, 3505, and 4193 in chromosome II), vibriobactin receptor related genes (FORC36_4181, 4183, 4185, 4186, 4187, 4189, and 4195 in chromosome II), and vulnibactin receptor related genes (FORC36_4188 and 4194 in chromosome II).

Furthermore, the strain FORC_036 had many virulence genes which did not describe here. These genes which localized in the strain FORC_036 genome are summarized in Table 5.

Table 5. Virulence factors of *V. vulnificus* FORC_036

Virulence factor	Gene product	Chromosome	Location (locus tag)	Function
Adherence				
mshH, mshI, mshJ, mshK, mshL, mshM, mshN, mshE, mshG, mshF, mshB, mshA, mshC, mshD	Mannose-sensitive hemagglutinin (MSHA type IV pilus)	Chromosome I	2816698~2830388 (FORC36_2538~2551)	Adherence
pilA, pilB, pilC, pilD	Type IV pilus	Chromosome I	2649553~2653946 (FORC36_2388~2391)	Adherence
Antiphagocytosis				
wza, hp1, wzb, wzc	Capsular polysaccharide	Chromosome I	229183~233842 (FORC36_0205~0208)	Antiphagocytosis
wbfV/wcvB	Capsular polysaccharide	Chromosome I	253968~255134 (FORC36_0228)	Antiphagocytosis
cpsA, cpsB, cpsC, cpsD, cpsF, cpsH, cpsI, cpsJ	Capsular polysaccharide synthesis enzyme	Chromosome II	460435~470233 (FORC36_3292~3299)	Antiphagocytosis
Chemotaxis and motility				
flaC, flaA, flgL, flgK, flgJ, flgI, flgH, flgG, flgF, flgE, flgD, flgC, flgB, flgA, flgM, flgN	Flagella	Chromosome I	791742~808099 (FORC36_0717~0734)	Motility

Virulence factor	Gene product	Chromosome	Location (locus tag)	Function
cheR, cheV	Chemotaxis protein	Chromosome I	793483~795250 (FORC36_0720~0721)	Chemotaxis
flhB, fliR, fliQ, fliP, fliO, fliN, fliM, fliL, fliK, fliJ, fliI, fliH, fliG, fliF, fliE, flrC, flrB, flrA, fliS, flaI, fliD, flaG, flaB, flaD, flaE	Flagella	Chromosome I	2310737~2335973 (FORC36_2085~2109)	Motility
cheW, cheB, cheA, cheZ, cheY	Chemotaxis protein	Chromosome I	2298257~229875 (FORC36_2074) 2300632~2305258 (FORC36_2077~2080)	Chemotaxis
fliA, flhG, flhF, flhA	Flagella	Chromosome I	2305030~2310576 (FORC36_2081~2084)	Motility
motA, motB, motY, motX	Flagella	Chromosome I	709061~710791 (FORC36_0642~0643) 1052282~1053163 (FORC36_0943) 2963591~2964226 (FORC36_2661)	Motility
Exoenzyme				
hap/vvp	Vibriolysin, extracellular zinc protease	Chromosome II	1656911~1658740 (FORC36_4339)	Metalloprotease

Virulence factor	Gene product	Chromosome	Location (locus tag)	Function
Iron uptake				
hutR	TonB-dependent heme receptor HutR	Chromosome II	428135~430270 (FORC36_3267)	Heme receptor
vctP, vctD, vctG, vctC	Periplasmic binding protein-dependent ABC transport systems	Chromosome II	687671~690300 (FORC36_3503~3505)	Iron uptake
vibB, vibE, vibC vibA, vibH, vibD, vibF	Vibriobactin biosynthesis	Chromosome II	1494198~1495106 (FORC36_4193) 1475589~1476284 (FORC36_4181) 1479412~1483974 (FORC36_4183) 1485512~1489503 (FORC36_4185~4187) 1490498~1491385 (FORC36_4189) 1497596~1498867 (FORC36_4195)	Iron uptake
viuB, viuA	Ferric vulnibactin receptor	Chromosome II	1489620~1490435 (FORC36_4188) 1495192~1497255 (FORC36_4194)	Iron uptake
Quorum sensing				
luxS	Autoinducer-2	Chromosome I	2661885~2662403	Quorum sensing

Virulence factor	Gene product	Chromosome	Location (locus tag)	Function
			(FORC36_2402)	
Secretion system				
epsN, epsM, epsL, epsK, epsJ, epsI, epsH, epsG, epsF, epsE, gspD, epsC	EPS type II secretion system	Chromosome I	116560~127879 (FORC36_0092~0103)	General secretion pathway protein
Toxin				
vvhA	Hemolysin/cytolysin	Chromosome II	1009573~1010988 (FORC36_3773)	Cytolysin precursor
rtxA, rtxC, rtxB, rtxD	RTX toxin	Chromosome II	1089984~1104513 (FORC36_3840~3841) 1105404~1108867 (FORC36_3843~3844)	Toxin
tlh	Thermolabile hemolysin precursor	Chromosome II	349784~351037 (FORC36_3212)	Toxin

Phylogenetic analysis

To check the phylogenetic status of *V. vulnificus* FORC_036, the 16S rRNA sequence of the strain FORC_036 was compared with that of other *Vibrio* species (*V. vulnificus*, *V. parahaemolyticus*, and *V. cholerae*). The result demonstrated that the strain FORC_036 belongs to *V. vulnificus* (Fig. 5).

Additionally, ANI analysis using complete genome sequences of various *V. vulnificus* strains which were available (MO6-24/O, CMCP6, 93U204, YJ016, and FORC_009) was conducted to figure out the evolutionary relationship in the same species (*V. vulnificus*). The result showed that the strain FORC_036 had the highest ANI value (98.44) with the strain CMCP6, a clinical isolate from Chonnam National University Hospital (Kim *et al*., 2003) (Fig. 6).

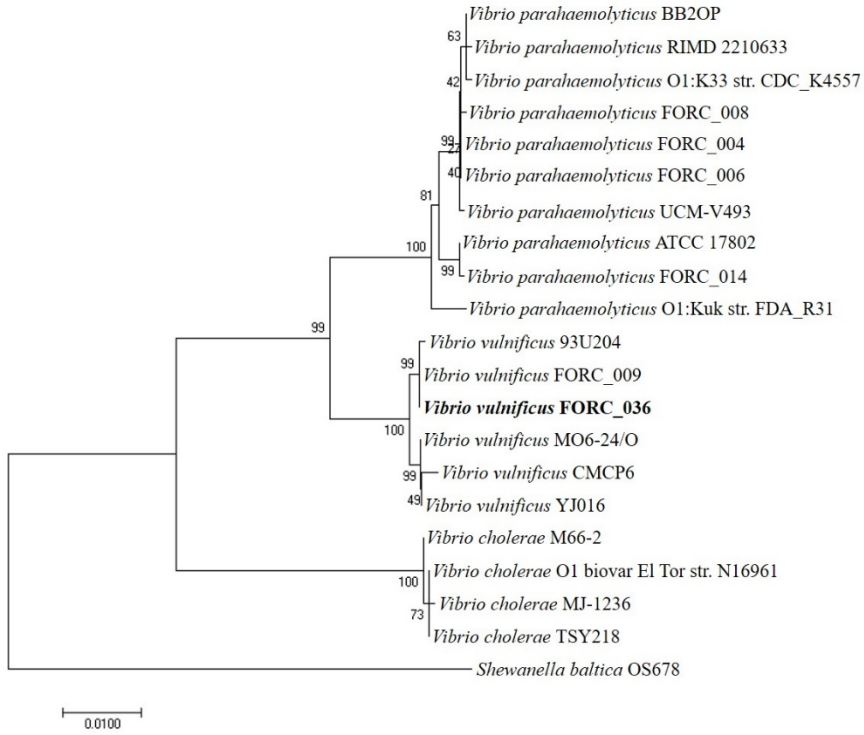

Fig 5. Phylogenetic tree of the strain FORC_036 with *Vibrio* species. The 16S rRNA sequences from various *Vibrio* species (*V. vulnificus*, *V. parahaemolyticus*, and *V. cholerae*) were aligned by using ClustalW. The phylogenetic tree was constructed using the neighbor joining method with 1000 bootstrap replicates. *S. balitca* OS678 was used as the out-group. The scale bar was below the tree and the bootstrap values were on the branch.

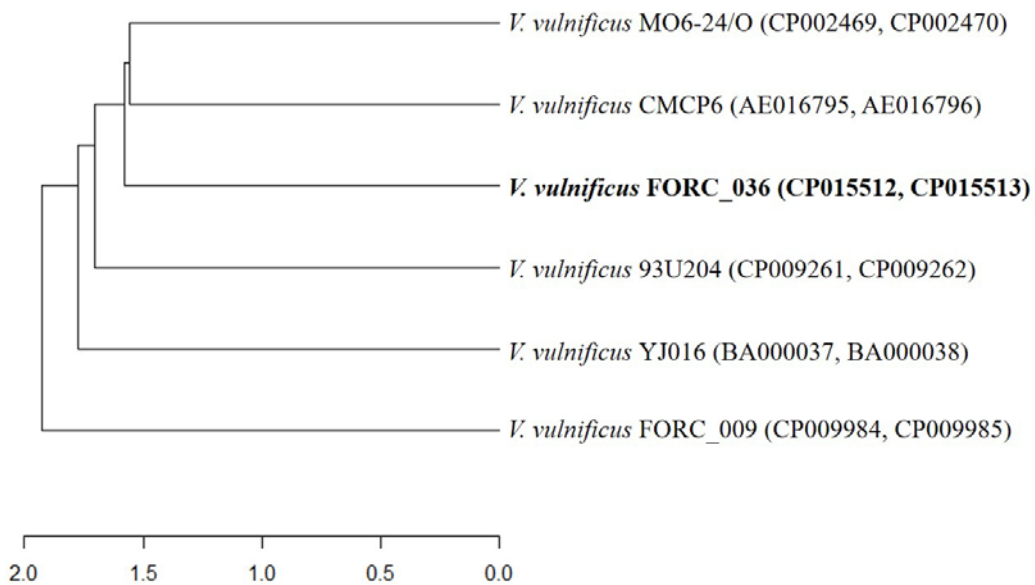

Fig 6. ANI analysis of *V. vulnificus* strains. This phylogenetic tree was created using the ANI values. The complete genome sequences of various *V. vulnificus* strains were used when the ANI values were calculated. The shorter the bar between two strains was, the closer the relativeness between two strains was. The scale bar was below the tree.

Comparative genome analysis between *V. vulnificus* FORC_036 and CMCP6

The comparative genome analysis between *V. vulnificus* FORC_036 and the closest strain CMCP6 indicated that the strain FORC_036 has additional gene cluster or gene (Fig. 7). The first region is gene cluster encoding L-fucose utilization proteins (from FORC36_1828 to 1834 in chromosome I) (Fig. 7A). The second additional gene is the thermostable hemolysin delta-VPH encoding gene (FORC36_3031 in chromosome II) (Fig. 7B). The information of these regions was listed in Table 6.

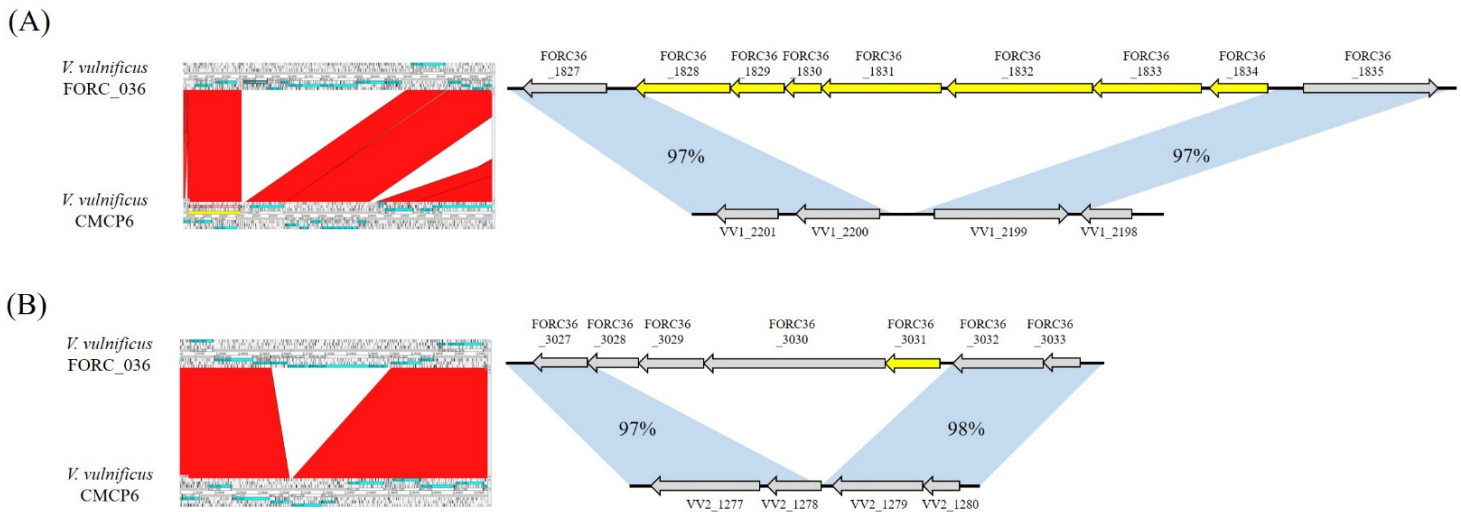

Fig 7. Comparative genome analysis between *V. vulnificus* FORC_036 and CMCP6. The region included (A) the L-fucose utilization genes, and (B) the thermostable hemolysin delta-VPH encoding gene was expressed, respectively. (Left part of each figure) The red regions were the homologue regions and the white regions were the additional regions in the strain FORC_036. (Right part of each figure) The direction of arrows was that of corresponding genes. The yellow arrows represented the important functional genes in the strain FORC_036. The percentage in the figure was indicated the nucleotide sequence similarity of the homologue regions.

Table 6. Additional genes of *V. vulnificus* FORC_036 compared to *V. vulnificus* CMCP6

Locus tag	Gene product
L-fucose utilization	
FORC36_1828	Lactaldehyde dehydrogenase involved in fucose or rhamnose utilization
FORC36_1829	L-fuculose phosphate aldolase
FORC36_1830	L-fucose mutarotase
FORC36_1831	L-fuculokinase
FORC36_1832	L-fucose isomerase
FORC36_1833	L-fucose:H^+ symporter permease
FORC36_1834	L-fucose operon activator
Thermostable hemolysin delta-VPH	
FORC36_3028	Hypothetical protein
FORC36_3029	Short-chain dehydrogenase/reductase SDR
FORC36_3030	Long-chain-fatty-acid-CoA ligase / long-chain acyl-CoA synthetase
FORC36_3031	Thermostable hemolysin delta-VPH

Identification of differentially expressed genes of *V. vulnificus* FORC_036 under exposure to the small octopus

The transcriptome analysis was used to compare the transcriptional profile when the strain FORC_036 was exposed to the small octopus. As a control, the same strain which was not exposed to the model food was used. Using the RNA samples which were extracted from the strain FORC_036 incubated in VFMG with presence or absence of the small octopus, RNA sequencing was conducted. Average RPKM values from the biological duplicate samples were used to represent the expression level of each gene. As the Figure 8 showed, the genes differently expressed with the significance (2 fold change, *P*-value < 0.05) were identified. A total of 1,542 genes were differently expressed when *V. vulnificus* FORC_036 was exposed to the small octopus. Among them, 912 genes were up-regulated and 630 genes were down-regulated. The differently expressed genes were clustered into the functional groups using WebMGA server (http://weizhong-lab.ucsd.edu/metagenomic-analysis/) (Fig. 9). This graph showed that the genes in various functional groups were differently expressed.

The motility related genes were down-regulated and genes related to adherence, galactose utilization, and iron uptake were up-regulated (Table 7). Additionally, the heat map was constructed to visualize the RNA sequencing result (2 fold change, *P*-value < 0.05) (Fig. 10).

To confirm the result of RNA sequencing, qRT-PCR was conducted. The results showed that the inferred contents from RNA sequencing were reliable (Fig. 11).

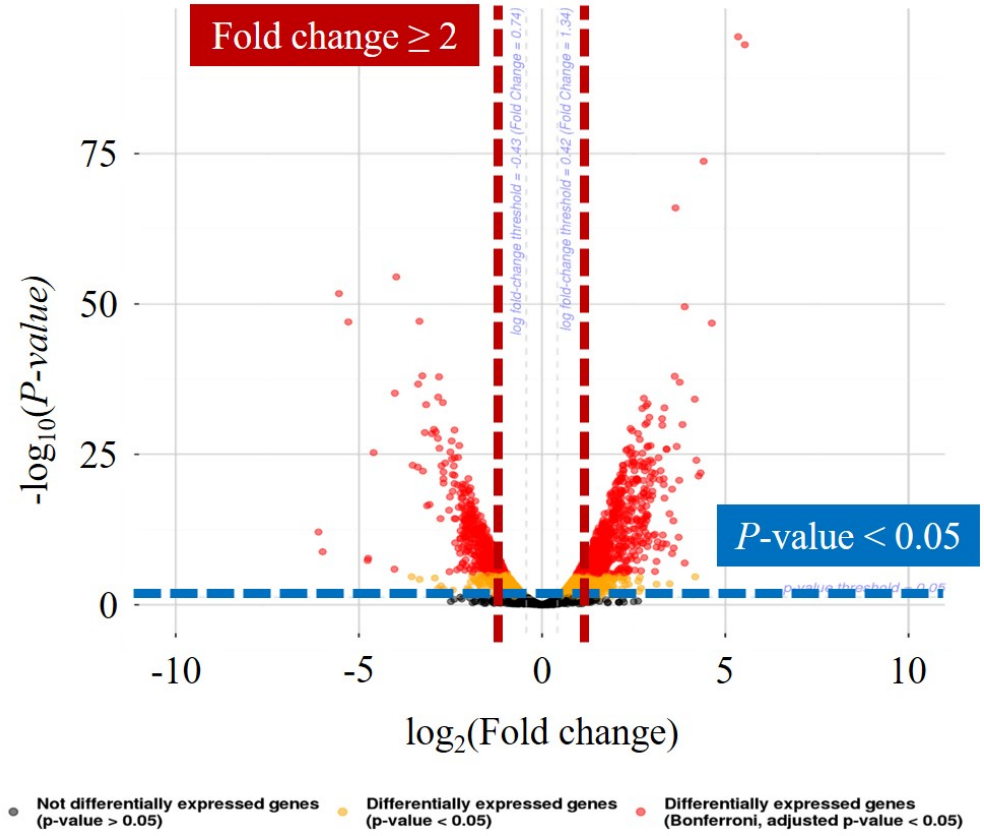

Fig 8. Transcriptome comparison of the RNA sequencing samples. The volcano plot was created to express differently expressed genes when the strain FORC_036 was exposed to the small octopus. Numbers on the X- and Y-axis represented the fold change (log_2) and the *P*-value ($-log_{10}$), respectively.

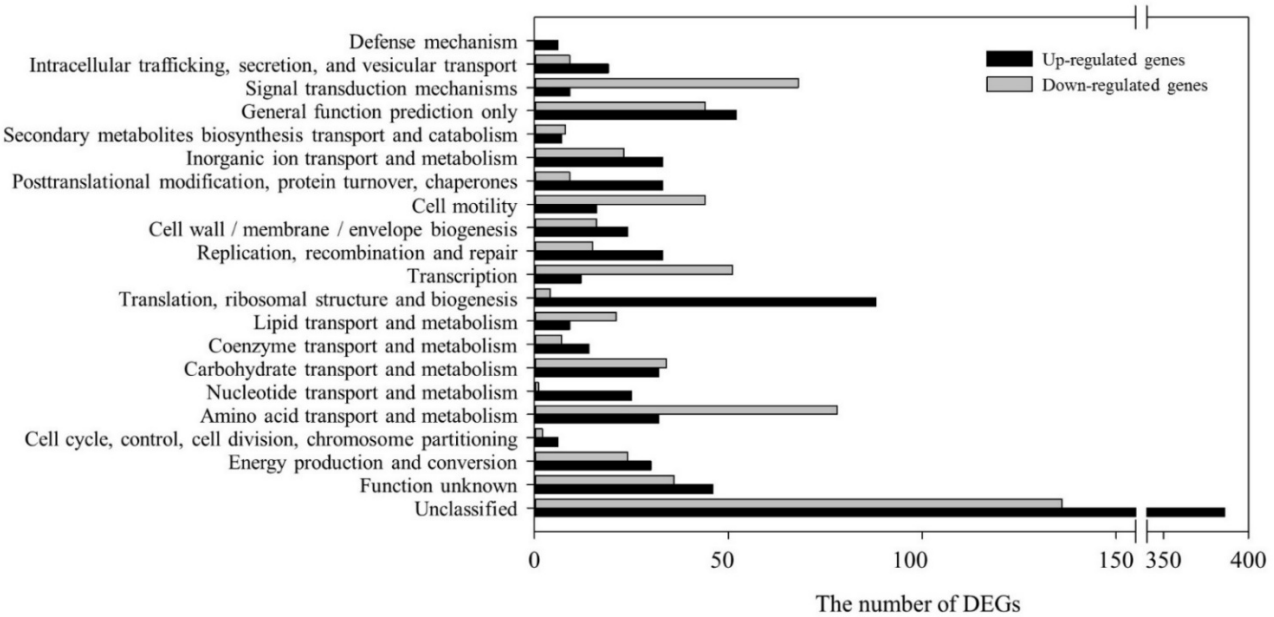

Fig 9. Functional categorization of genes differentially expressed under exposure to the small octopus. The differently expressed genes when the strain FORC_036 was exposed to the small octopus were categorized by their function. The functional categories are based on the *V. vulnificus* FORC_036 genome. The black bars represented the up-regulated genes and the gray bars indicated the down-regulated genes. DEGs, differently expressed genes.

Table 7. Transcriptome of various genes under exposure to the small octopus

Locus tag	Gene product	Fold change [a]	*P*-value [b]
Motility			
FORC36_0642	Flagellar motor rotation protein MotA	0.45	0.00006
FORC36_0643	Flagellar motor rotation protein MotB	0.42	0.00002
FORC36_2090	Flagellar motor switch protein FliN	0.49	0.00002
FORC36_2091	Flagellar motor switch protein FliM	0.44	0
FORC36_3456	Ribose ABC transport system, periplasmic ribose-binding protein RbsB	3.50	0
FORC36_4526	Signal transduction histidine kinase CheA	0.31	0
FORC36_4527	Fis family transcriptional regulator	0.25	0
FORC36_4883	Aerotaxis sensor receptor protein	2.84	0
FORC36_5304	Diguanylate cyclase/phosphodiesterase with PAS/PAC sensor(s)	5.66	0
FORC36_0730	Flagellar protein FlgJ	0.35	0
FORC36_0734	Flagellin protein FlaC	0.23	0
FORC36_0737	Flagellin protein FlaD	0.31	0
FORC36_0943	Sodium-type flagellar protein MotY	0.32	0

Locus tag	Gene product	Fold change [a]	*P*-value [b]
FORC36_2104	Flagellar rod protein flaI	0.45	0
FORC36_2108	Flagellin protein FlaD	0.33	0
FORC36_2661	Sodium-type flagellar protein MotX	0.45	0
FORC36_0717	Flagellar biosynthesis protein FlgN	0.42	0
FORC36_0718	Negative regulator of flagellin synthesis FlgM	0.43	0
FORC36_0722	Flagellar basal-body rod protein FlgB	0.35	0
FORC36_0723	Flagellar basal-body rod protein FlgC	0.35	0
FORC36_0724	Flagellar basal-body rod modification protein FlgD	0.35	0
FORC36_0725	Flagellar hook protein FlgE	0.38	0
FORC36_0726	Flagellar basal-body rod protein FlgF	0.27	0
FORC36_0727	Flagellar basal-body rod protein FlgG	0.23	0
FORC36_0728	Flagellar L-ring protein FlgH	0.22	0
FORC36_0729	Flagellar P-ring protein FlgI	0.28	0
FORC36_0731	Flagellar hook-associated protein FlgK	0.18	0
FORC36_0732	Flagellar hook-associated protein FlgL	0.23	0
FORC36_2092	Flagellar biosynthesis protein FliL	0.44	0

Locus tag	Gene product	Fold change [a]	*P*-value [b]
FORC36_2093	Flagellar hook-length control protein FliK	0.26	0
FORC36_2103	Flagellar biosynthesis protein FliS	0.46	0
FORC36_2105	Flagellar hook-associated protein FliD	0.37	0
FORC36_2109	Flagellin protein FlaF	0.26	0
Adherence			
FORC36_1688	TPR repeat containing Flp pilus assembly protein TadD	4.97	0
FORC36_1689	Type II/IV secretion system protein TadC	3.44	0
FORC36_1690	Flp pilus assembly protein TadB	5.52	0
FORC36_1691	Flp pilus assembly protein TadA	13.39	0
FORC36_1692	Type II/IV secretion system ATPase TadZ/CpaE,	6.32	0
FORC36_1694	Flp pilus assembly protein, secretin CpaC	4.44	0
FORC36_2276	Type II/IV secretion system secretin RcpA/CpaC	0.39	0.00005
FORC36_2277	Flp pilus assembly protein RcpC/CpaB	0.28	0.00009
FORC36_2334	Outer membrane protein OmpU	3.13	0
FORC36_5121	Type IV pilus biogenesis protein PilQ	7.03	0
FORC36_5130	Putative cell-surface adhesion	4.06	0

Locus tag	Gene product	Fold change [a]	*P*-value [b]
Galactose metabolism			
FORC36_0663	PTS system, trehalose-specific IIB or IIC component	0.30	0
FORC36_3587	PTS system, fructose-specific IIA component	0.49	0.00461
FORC36_2253	Aldose 1-epimerase	2.27	0.00001
FORC36_2259	Evolved beta-D-galactosidase, alpha subunit	6.66	0
FORC36_2260	Evolved beta-D-galactosidase, beta subunit	18.52	0
FORC36_4384	N-acetylgalactosamine-6-phosphate deacetylase	2.69	0.00049
FORC36_4385	PTS system, N-acetylgalactosamine- and galactosamine-specific IIA component	2.89	0.02578
FORC36_4386	PTS system, N-acetylgalactosamine-specific IID component	5.16	0
FORC36_4387	PTS system, N-acetylgalactosamine-specific IIC component	2.24	0.00394
Iron uptake			
FORC36_2356	Ferric iron ABC transporter, ATP-binding protein	2.73	0
FORC36_2357	Ferric iron ABC transporter, permease protein	2.19	0
FORC36_2358	Ferric iron ABC transporter, iron-binding protein	4.24	0
FORC36_3267	TonB-dependent heme receptor HutR	24.79	0

Locus tag	Gene product	Fold change [a]	*P*-value [b]
FORC36_3654	TonB-dependent heme and hemoglobin receptor HutA	2.66	0
FORC36_3327	Ferric siderophore transport system, biopolymer transport protein ExbB	6.99	0
FORC36_3328	Ferric siderophore transport system, periplasmic binding protein TonB	12.08	0
FORC36_3731	Ferric siderophore transport system, periplasmic binding protein TonB	6.01	0
FORC36_3733	Ferric siderophore transport system, biopolymer transport protein ExbB	6.04	0
FORC36_3503	Ferric vibriobactin, enterobactin transport system, permease protein VctD	3.45	0
FORC36_3504	Ferric vibriobactin, enterobactin transport system, permease protein VctG	2.75	0
FORC36_3505	Ferric vibriobactin, enterobactin transport system, ATP-binding protein	2.33	0
FORC36_4193	Catechol siderophore ABC transporter, substrate-binding protein	6.98	0
FORC36_4181	Phosphopantetheinyl transferase component of siderophore synthetase	3.46	0
FORC36_4183	Non-ribosomal peptide synthetase modules	6.94	0
FORC36_4185	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase	3.53	0.00003
FORC36_4186	Isochorismate synthase of siderophore biosynthesis	4.73	0
FORC36_4187	2,3-dihydroxybenzoate-AMP ligase	2.72	0
FORC36_4189	Isochorismatase of siderophore biosynthesis	9.42	0
FORC36_4188	Vulnibactin utilization protein VuuB	9.23	0

Locus tag	Gene product	Fold change [a]	*P*-value [b]
FORC36_4194	Ferric vulnibactin receptor VuuA	2.59	0
FORC36_4376	Hypothetical protein in aerobactin uptake cluster	4.48	0
FORC36_4377	Ferric aerobactin ABC transporter, ATPase component	6.79	0
FORC36_4378	Ferric aerobactin ABC transporter, periplasmic substrate binding protein	7.06	0
FORC36_4379	Ferric aerobactin ABC transporter, permease component	8.91	0
FORC36_4380	Ferric aerobactin ABC transporter, permease component	11.15	0
Toxin			
FORC36_3212	Thermolabile hemolysin precursor	1.73	0.00086
FORC36_3773	Cytolysin precursor	1.42	0.03075
FORC36_3841	RTX toxin activating lysine-acyltransferase	2.00	0.00043
FORC36_3843	RTX toxin transporter, ATP-binding protein	2.38	0
FORC36_3844	RTX toxin transporter, determinant D	2.38	0.00009

[a] The mRNA expression level in *V. vulnificus* FORC_036 when this strain was exposed to the small octopus relative to that in *V. vulnificus* FORC_036 when this strain was not exposed to the small octopus.

[b] The *P*-value less than six decimal places were denoted as zero.

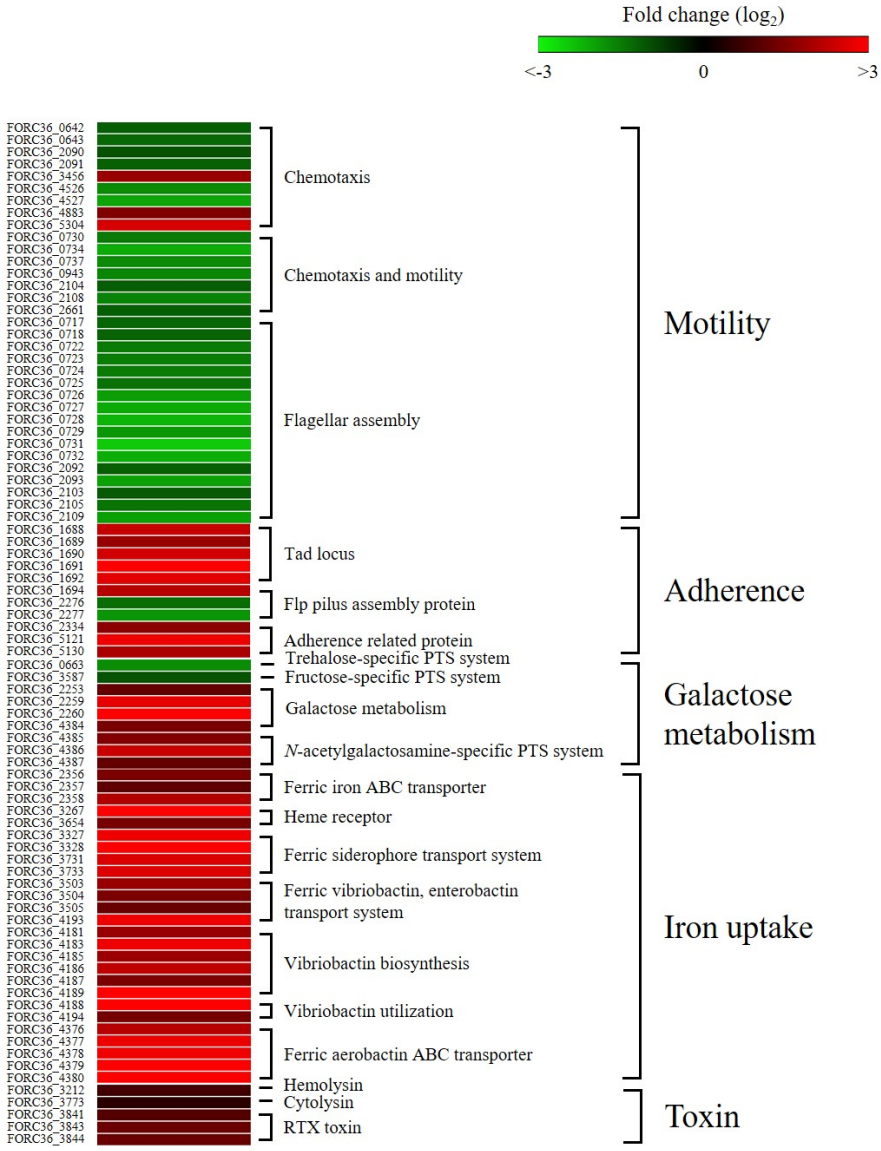

Fig 10. Transcriptome of various genes under exposure to the small octopus. The RNA sequencing results were visualized using the heat map. The red bars represented the up-regulated genes and the green bars indicated that the down-regulated genes. The scale bar was above the heat map.

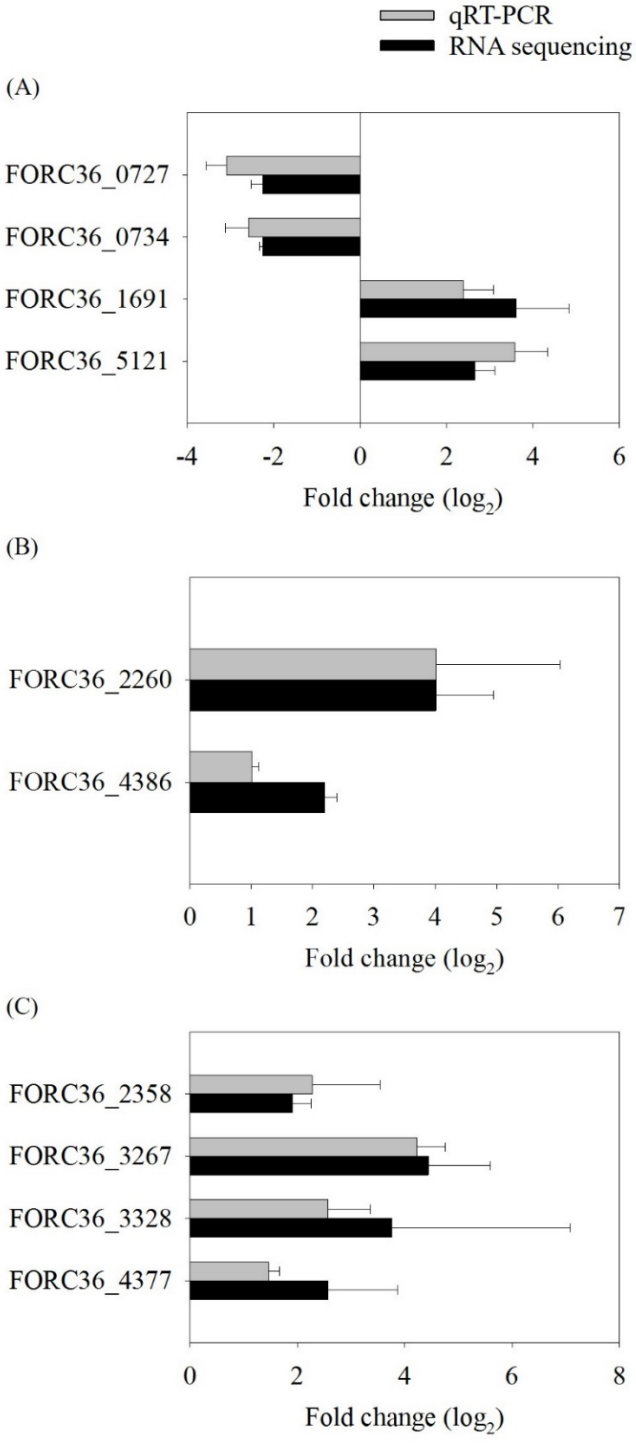
qRT-PCR
RNA sequencing
(A)
FORC36_0727
FORC36_0734
FORC36_1691
FORC36_5121
-4
-2
0
2
4
6
Fold change (log2)
(B)
FORC36_2260
FORC36_4386
0
1
2
3
4
5
6
7
Fold change (log2)
(C)
FORC36_2358
FORC36_3267
FORC36_3328
FORC36_4377
0
2
4
6
8
Fold change (log2)

Fig 11. Expression comparison between RNA sequencing and qRT-PCR. The results of RNA sequencing were confirmed using qRT-PCR targeting the represented genes. (A) Motility and adherence related gene, (B) galactose metabolism related gene, and (C) iron uptake related gene were confirmed, respectively. The black bars represented the RNA sequencing results and the gray bars indicated the qRT-PCR results. Information of each gene was shown at Table 7. The error bars represented the standard deviations (SD).

Growth kinetics of *V. vulnificus* FORC_036 exposed to small octopus

Because the *N*-acetylgalactosamine is one of the major component of the small octopus skin (Florkin and Scheer, 1970) and *V. vulnificus* can transport this molecule, the strain FORC_036 may use this nutrient for growth when exposed to the small octopus. The growth kinetics were compared between two samples which were incubated in VFMG with the presence or absence of the model food (small octopus), respectively. *V. vulnificus* FORC_036 grew faster and the colony forming unit (CFU) was 10 fold higher when the strain FORC_036 was exposed to small octopus (Fig. 12).

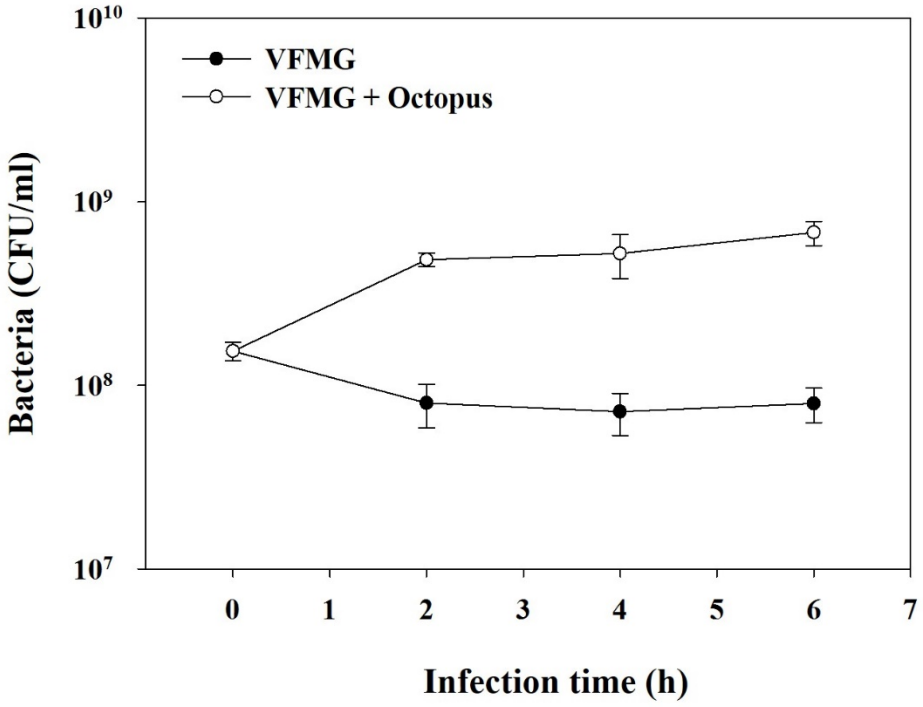

Fig 12. Growth kinetics of *V. vulnificus* FORC_036 incubated in VFMG with the presence or absence of the small octopus. The CFUs were counted at the specific time points (0 h, 2 h, 4 h, and 6 h). The experiments were repeated with biologically duplicated samples and technically triplicate were run for each sample. Open circle (○), the small octopus-exposed sample; closed circle (●), control sample. The error bars represented the standard deviations (SD).

IV. DISCUSSION

V. vulnificus is one of the opportunistic pathogen. In South Korea, some outbreaks involved in infection of this bacterium were occured. So, it is important to understand the genome and transcriptome of *V. vulnificus*. In this study, *V. vulnificus* isolated from a surf clam was analyzed and characterized.

To check the virulence possibility of this isolate, the virulence gene-specific PCR screening and LDH release assay were performed. *V. vulnificus* 1001957 had various virulence genes and showed a high level of cytotoxicity toward INT-407 cell lines (Fig 1 and 2). The result indicated that this strain was possibly pathogenic toward the human cells. After these experiments, *V. vulnificus* 1001957 was designated as 'FORC_036'.

Using its genomic DNA, the whole genome sequencing was conducted and the complete genome sequence was analyzed. Compared to other *V. vulnificus* strains, this strain had an additional large-sized contig (Table 4). Because of the size, it can be regarded as a novel chromosome. However, this contig did not satisfy the requirement of a chromosome (Trucksis *et al.*, 1998). And there was a report suggesting that *Pseudomonas syringe* had a 1 Mb-sized plasmid (Romanchuk *et al.*, 2014). Therefore, I concluded that the additional contig was a plasmid.

Because the strain FORC_036 was isolated from environment, it was needed to check whether its genome had the virulence factors which other virulent strains had. Through the BLAST search referring VFDB, the virulence factors of the strain

FORC_036 were predicted (Table 5).

First, the major virulence factors of *V. vulnificus* such as a hemolysin/cytolysin (VvhA) and a repeats-in-toxin (Rtx) (Jones and Oliver, 2009) were found in the genome of this strain. Also, additional hemolysin was detected. Many researches about these virulence factors said that they affected the pathogenesis of *V. vulnificus*. So, this result showed that *V. vulnificus* FORC_036 may be pathogenic.

Second, the strain FORC_036 had T2SS. T2SS secretes folded proteins from periplasm to extracellular environment (Costa *et al*., 2015). In *V. cholerae*, the cholera toxin is exported through this system (Costa *et al*., 2015). However, the strain FORC_036 does not have the cholera toxin. Instead, T2SS is required to export at least two virulence factors, VvhA and VvpE in *V. vulnificus* (Hwang *et al*., 2011). Addition to T1SS (Type I secretion system), T2SS is also responsible for the pathogenesis of *V. vulnificus*. As this secretion system influenced the virulence, the strain FORC_036 may cause the diseases. Although *V. vulnificus* had the secretion system which predicted to contribute the pathogenesis of this strain, more researches are required to confirm this prediction.

Third, *V. vulnificus* FORC_036 had multiple iron uptake systems. Iron acquisition from the host cell is important for *V. vulnificus* to survive and grow (Wright, Simpson and Oliver, 1981). Also, the loss of iron by the pathogen (in this case, *V. vulnificus*) gives negative influence to the host cells. As *V. vulnificus* FORC_036 can uptake free iron from the host cell using these proteins, it can survive and be pathogenic in host environment.

Though *V. vulnificus* FORC_036 was isolated from the environment (from a surf clam), it has lots of the virulence factors (hemolysin/cytolysin, RTX toxins, secretion systems for toxin transport, and iron uptake systems) and the cytotoxicity of *V. vulnificus* FORC_036 was similar to cytotoxicity of *V. vulnificus* MO6-24/O, a clinical isolate, toward the host cell. So, the strain FORC_036 is a possible pathogen which can cause various disease.

To characterize the strain FORC_036, the comparative genome analysis was carried out. Compared with the closest strain CMCP6 (a clinical isolate), the strain FORC_036 had additional region and gene. The first region is L-fucose utilization gene cluster. L-fucose accounts for 4~14% of the oligosacharide of the human intestinal mucin (Staib and Fuchs, 2015). So, some bacteria which can use this sugar have the advantage of surviving in the human intestine. The second one is thermostable hemolysin delta-VPH encoding gene. Because the hemolysin is major virulence factor, it is possible to predict the strain FORC_036 is virulent toward the host. Consequentially, the strain FORC_036 could be pathogenic and has benefit on survival in the human intestine although it was isolated from a surf clam, one of shellfish.

To elucidate the behavior of *V. vulnificus* FORC_036 when this strain was exposed to the small octopus, the RNA sequencing was performed using two RNA samples which were extracted from the bacterium incubated in VFMG with the presence or absence of the small octopus, respectively. As figure 10 showed, a large number of genes were differently expressed. The genes related to motility were down-regulated

but the genes related to adherence were up-regulated. So, the strain FORC_036 may attach to the small octopus through the *tad* (tight adherence) locus. The *tad* locus includes various genes necessary for bacterial adhesion to surface, biofilm formation, and pathogenesis. (Kram *et al*., 2008). And, expression of genes related to the galactose metabolism and the *N*-acetylgalactosamine transporter was up-regulated. Because the *N*-acetylgalactosamine is the major component of the small octopus's skin (Florkin and Scheer, 1970), this expression change indicated that *V. vulnificus* FORC_036 could uptake and use this nutrient for survival and growth when it contacted to the small octopus. Lastly, various iron uptake systems were up-regulated. As previously mentioned, iron is important to the survival and growth of *V. vulnificus*. So, this response gave advantage to *V. vulnificus* FORC_036. These results described above suggested the behavior of the strain FORC_036 when this strain was exposed to the small octopus; (1) motility reduction, (2) attachment to the small octopus through the *tad* locus, (3) utilization the *N*-acetylgalacotsamine of the small octopus skin, and (4) increase the iron uptake. However, the expression of toxin-related genes did not change. This result suggested that the strain FORC_036 perceived the small octopus as a reservoir rather than a host.

In conclusion, *V. vulnificus* FORC_036 had lots of virulence factors and showed a high level of cytotoxicity despite an environmental isolate. It had two chromosomes and one plasmid in its genome. This bacterium had additional genes compared with the closest strain, CMCP6, one of the clinical isolates. Also, it suggested that the strain FORC_036 had advantage to the growth and virulence in the intestine of human.

And transcriptional analysis of this strain when the strain FORC_036 was exposed to the model food, small octopus, indicated that it kept survive and grow in the small octopus.

This study contributes the accumulation of database about the domestic pathogens and helps to deal with the coming outbreak involved in *V. vulnificus*.

V. REFERENCE

Altermann, E., & Klaenhammer, T. R. (2003). GAMOLA: a new local solution for sequence annotation and analyzing draft and finished prokaryotic genomes. *Omics A Journal of Integrative Biology*, 7(2), 161-169.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., ... & Meyer, F. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC genomics*, 9(1), 1.

Besemer, J., Lomsadze, A., & Borodovsky, M. (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic acids research*, 29(12), 2607-2618.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B. G., & Parkhill, J. (2005). ACT: the Artemis comparison tool. *Bioinformatics*, 21(16), 3422-3423.

Chan, F. K. M., Moriwaki, K., & De Rosa, M. J. (2013). Detection of necrosis by release of lactate dehydrogenase activity. *Immune Homeostasis: Methods and Protocols*, 65-70.

Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology*, 13(6), 343-359.

Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E., & Crook, D. W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nature Reviews Genetics*, 13(9), 601-612.

Schoffeniels, E., Florkin, R. G. M., & Scheer, B. I. (1970). Chemical Zoology (Volume 5, Arthropoda. Part A), *Elsevier*

Gulig, P. A., Bourdage, K. L., & Starks, A. M. (2005). Molecular pathogenesis of *Vibrio vulnificus*. *The Journal of Microbiology*, 43(Suppl 1), 118-131.

Hwang, W., Lee, N. Y., Kim, J., Lee, M. A., Kim, K. S., Lee, K. H., & Park, S. J. (2011). Functional characterization of EpsC, a component of the type II secretion system, in the pathogenicity of *Vibrio vulnificus*. *Infection and immunity*, 79(10), 4068-4080.

Jeong, K. C., Jeong, H. S., Rhee, J. H., Lee, S. E., Chung, S. S., Starks, A. M., ... & Choi, S. H. (2000). Construction and phenotypic evaluation of a *Vibrio vulnificus vvpE* mutant for elastolytic protease. *Infection and immunity*, 68(9), 5096-5106.

Jones, M. K., & Oliver, J. D. (2009). *Vibrio vulnificus*: disease and pathogenesis. *Infection and immunity*, 77(5), 1723-1733.

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., ... & Pesseat, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30(9), 1236-1240.

Kim, Y. R., Lee, S. E., Kim, C. M., Kim, S. Y., Shin, E. K., Shin, D. H., ... & Handfield, M. (2003). Characterization and pathogenic significance of *Vibrio vulnificus* antigens preferentially expressed in septicemic patients. *Infection and immunity*, 71(10), 5461-5471.

Kram, K. E., Hovel-Miner, G. A., Tomich, M., & Figurski, D. H. (2008). Transcriptional regulation of the *tad* locus in *Aggregatibacter actinomycetemcomitans*: a termination cascade. *Journal of bacteriology*, 190(11), 3859-3868.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, msw054.

Lim, J. G., & Choi, S. H. (2014). IscR is a global regulator essential for pathogenesis of *Vibrio vulnificus* and induced by host cells. *Infection and immunity*, 82(2), 569-578.

Linkous, D. A., & Oliver, J. D. (1999). Pathogenesis of *Vibrio vulnificus*. *FEMS Microbiology Letters*, 174(2), 207-214.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, 5(7), 621-628.

Pinto, A. C., Melo-Barbosa, H. P., Miyoshi, A., Silva, A., & Azevedo, V. (2011). Application of RNA-seq to reveal the transcript profile in bacteria. *Genetic and Molecular Research*, 10(3), 1707-1718.

Perez-Llamas, C., & Lopez-Bigas, N. (2011). Gitools: analysis and visualisation of genomic data using interactive heat-maps. *PloS one*, 6(5), e19541.

Richter, M., & Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, 106(45), 19126-19131.

Romanchuk, A., Jones, C. D., Karkare, K., Moore, A., Smith, B. A., Jones, C., ... & Baltrus, D. A. (2014). Bigger is not always better: Transmission and fitness burden of ~ 1MB *Pseudomonas syringae* megaplasmid pMPPla107. *Plasmid*, 73, 16-25.

Staib, L., & Fuchs, T. M. (2015). Regulation of fucose and 1, 2-propanediol utilization by *Salmonella enterica* serovar Typhimurium. *Frontiers in microbiology*, 6.

Strom, M. S., & Paranjpye, R. N. (2000). Epidemiology and pathogenesis of *Vibrio vulnificus*. *Microbes and infection*, 2(2), 177-188.

Trucksis, M., Michalski, J., Deng, Y. K., & Kaper, J. B. (1998). The *Vibrio cholerae* genome contains two unique circular chromosomes. *Proceedings of the National Academy of Sciences*, 95(24), 14464-14469.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, 10(1), 57-63.

Wright, A. C., Simpson, L. M., & Oliver, J. D. (1981). Role of iron in the pathogenesis of *Vibrio vulnificus* infections. *Infection and immunity*, 34(2), 503-507.

VI. 국문초록

패혈증 비브리오균은 그람음성 간균으로 해양에 서식하는 식중독균이다. 이는 가벼운 위장질환에서부터 사망률이 50%에 달하는 패혈증을 야기한다. 해산물을 날 것으로 섭취하거나 오염된 해수에 상처부위가 노출되면 감염이 발생한다. 동죽에서 분리한 패혈증 비브리오균의 독성을 PCR과 세포주 실험을 통해 확인하고 'FORC_036'으로 명명하였다. FORC_036의 특성을 알아보기 위해서 유전체와 전사체 분석을 진행하였다.

FORC_036은 두 개의 염색체와 하나의 플라스미드를 가지고 있었으며, 5,642개의 전사해석틀, 203개의 운반 RNA와 34개의 리보솜 RNA가 지니고 있다고 예측된다. FORC_036은 주요 독성인자는 물론 병원성에 기여하는 다양한 유전자들을 가지고 있었다. 비브리오 속의 16S 리보솜 RNA를 이용한 계통 분석을 통해서 FORC_036이 패혈증 비브리오균에 속한다는 것을 다시 확인하였다. 더불어, 유전체 분석이 완료된 5개의 패혈증 비브리오균과의 종 간 유사도를 확인한 결과, FORC_036은 환자로부터 분리한 CMCP6와 계통학적으로 가장 유사하였다. FORC_036과 CMCP6의 유전체 비교 분석을 통해 FORC_036이 fucose 대사 유전자와 thermostable hemolysin delta-VPH

유전자를 추가적으로 가지고 있음을 확인하였다.

국내에서 날 것으로 섭취한 후, 패혈증을 일으킨 사례가 있는 낙지를 모델 식품으로 선정하여 FORC_036가 낙지에 노출되었을 때의 유전자 발현양상 변화를 RNA sequencing 기법을 이용해 확인하였다. 운동성과 관련된 유전자는 발현이 감소한 반면, 부착, 유당 대사, 철 흡수와 관련된 유전자의 발현은 증가하였다. 하지만 독성 관련 유전자의 발현은 큰 변화를 나타내지 않았다.

결론적으로, 동죽에서 분리한 패혈증 비브리오균 FORC_036은 여러 독성인자를 가지고 있어 병원성을 나타낼 가능성이 높고, 모델 식품인 낙지에 노출되면 낙지를 하나의 저장소로 인지하고 생존과 성장을 유지할 것으로 예측된다.

본 연구를 통해 국내 식중독균에 대한 데이터베이스를 축적하고 이 후에 발생하는 패혈증 비브리오균에 의한 질병 발생을 효율적으로 처리할 수 있을 것으로 기대된다.

주요어: 패혈증 비브리오균, 식중독균, 유전체 분석, 전사체 분석

학 번: 2015-21758