Thesis for the Degree of Master of Science

Genome and Transcriptome Analyses of *Vibrio vulnificus* FORC_037 Isolated from Raw Seafood

우럭조개에서 분리한 패혈증 비브리오균 FORC_037에 대한 유전체 및 전사체 분석

August, 2017

Eun Jung Na

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

석사학위논문

Genome and Transcriptome Analyses of *Vibrio vulnificus* FORC_037 Isolated from Raw Seafood

지도교수 최 상 호

이 논문을 석사학위논문으로 제출함

2017년 6월

서울대학교 대학원

농생명공학부

나 은 정

나은정의 석사학위논문을 인준함

2017년 6월

위원장 서 진 호 (인)

부위원장 최 상 호 (인)

위원 유 상 렬 (인)

Abstract

Vibrio vulnificus naturally inhabits the coastal marine environments worldwide and is an opportunistic pathogen for humans as it may cause severe wound infections, gastroenteritis or life-threatening sepsis in susceptible individuals. To study this important pathogen at the genomic level, *V. vulnifcus* FORC_037 was isolated from soft-shell clam (*Mya arenaria oonogai)* and its whole genome was sequenced using Illumina MiSeq and PacBio platforms. The strain's genome, which is composed of two chromosomes and a plasmid, altogether contains 4,506 open reading frames, 118 tRNA, and 34 rRNA genes. Genes encoding several hemolysins, and iron uptake-related proteins were found by BLAST. Average nucleotide identity (ANI) analysis of FORC_037's genome with nine other completely sequenced *V. vulnificus* genomes showed that the genome, while being an outlier, is most closely related to those of FORC_017 and CMCP6. Comparative genome analysis of FORC_037 and CMCP6, a clinical isolate, revealed that FORC_037 has additional virulence factors such as accessory cholera enterotoxin (FORC37_3618) and zonula occludens toxin (FORC37_3619). This may explain why, despite being an environmental isolate, FORC_037 exhibits a high level of cytotoxicity toward INT-407 human epithelial cells as evidenced by lactose dehydrogenase (LDH) release assay. To further probe the genetic program of the strain upon contact with small octopus (*Octopus minor*), which is often consumed raw in Korea, transcriptome sequencing was used. Transcriptome analysis hinted that *V. vulnificus* uses the seafood as a reservoir as genes related to adhesion, galactose utilization, oxidative

stress resistance and iron-uptake were upregulated and genes related to motility were downregulated. Interestingly, a number of putative virulence factors were upregulated, including genes involved in vulnibactin utilization, and type II secretion system which is associated with pilus assembly. This may explain illness following consumption of the seafood. This report will help prevent *V. vulnificus* outbreaks in the future by providing genomic and transcriptomic insights on the species.

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

Student Number: 2015-23131

Contents

List of Figures

List of Tables

I. INTRODUCTION

The importance of studying *V. vulnificus* stems from the yearly recurrence of infections and the dangerous consequences of the infections. From 2011 to 2015 in South Korea, 269 people were infected and out of the 269 patients, 147 died, making the mortality rate of the period, 54.65% (Korea Centers for Disease Control and Prevention, http://is.cdc.go.kr/dstat/jsp/stat/stat0001.jsp). Especially, there have been cases of *V. vulnificus* infections following consumption of raw small octopus (*Octopus minor*) from as early as in 1990 and up to 2015.

The regular occurrence of *V. vulnificus* combined with the popularity of raw small octopus is proving to be a threat to food safety and public health. There is room for improvement in terms of how the raw small octopus has been prepared and served to customers. Needless to say, the process should be evidence-based and high throughput sequencing technologies can be a useful addition to the toolbox.

There are many benefits of the omics techniques in their applications to food safety. Genomics can be used to characterize pathogens, and transcriptomics can be used to characterize the pathogen's response to a stimuli such as stress and antimicrobial treatments (Bergholz *et al.*, 2014). At the time of writing this manuscript, there are 9 whole genome sequences of *V. vulnificus* on National Center for Biotechnology Information (NCBI), compared to the 322 whole genomes of *Escherichia coli.* More genome data of the species needs to be accumulated and analyzed in order to help prevent future outbreaks. Along with whole genome

sequencing, RNA sequencing will be a useful method to characterize the transcripts of bacteria when in contact with model food.

V. vulnificus FORC_037 was isolated from a soft-shell clam (*Mya arenaria oonogai*) by Gyeonggi Health and Environmental Institute and Ministry of Drug and Food Safety in South Korea. In this study, the bacterial genome was completely sequenced and analyzed, and its transcriptome was sequenced to uncover the genetic program that is set in motion upon contact with small octopus. It is hoped that this work will further our understanding on how *V. vulnificus* causes diseases and help prevent *V. vulnificus* outbreaks in the future.

.

II. MATERIALS AND METHODS

Strains and growth condition

Table 1 shows the *V. vulnificus* strains that were used in this study. The strains were aerobically incubated at 30°C for 12 h in Luria-Bertani (LB) medium supplemented with 2.0% (*w/v*) NaCl.

Genomic DNA extraction and identification

DNeasy Blood & Tissue Kit (QIAGEN, Valencia, California, USA) was used according to the manufacturers' protocol to extract genomic DNA. The 16S rRNA gene was amplified from the extracted genomic DNA and sequenced by an automated ABI3730XL capillary DNA sequencer (Applied Biosystems, Foster City, CA, USA) for taxonomic identification (Ku *et al*., 2014).

Virulence gene-specific PCR

Table 2 shows the primers of the genes (*vvhA, vvpE, vcgC, rtxA, nanA* and *gyrB*) that were used for virulence gene-specific PCR. After resolving the amplified DNA fragments in 1.5% agarose gel, the electrophoresis results were visualized by using Geldoc™ EZ Image (Bio-Rad, Richmond, California, USA).

Cytotoxicity test

Lactate dehydrogenase (LDH) release assay was employed to estimate the

cytotoxicity levels of the *V. vulnificus* strains. For the assay, INT-407 human epithelial cells (ATCC® CCL-6™) were grown and infected with bacterial culture in 96-well tissue culture plates (Nunc, Roskilde, Denmark). INT-407 cells in the positive control well were completely lysed via treatment of 4% Triton X-100 and the cells in the negative control well were conserved as only assay media was added. Biological duplicates were used and for each biological sample, triplicate wells were used. Using the cytotoxicity Detection Kit (Roche, Mannheim, Germany), LDH activity from the lysed INT-407 cells within the supernatant was measured and was finally calculated using the following formula:

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

Transmission electron microscope

FORC_037 was negatively stained with Phosphotungstic acid (PTA) for a few seconds and then washed two to three times. The organism was observed using LIBRA 120 (Carl Zeiss, Oberkochen, Germany) transmission electron microscope (TEM) at 120 kV.

Genome sequencing and annotation

Whole genome sequencing and assembly were performed at ChunLab Incorporation (Seoul, South Korea). Dual platforms of Illumina MiSeq (Illumina, San Diego, California, USA), and PacBio (Pacific Biosciences, Menlo Park,

California, USA) were used. The raw sequence reads from Illumina MiSeq were assembled with CLC Genomics workbench 7.5.1 (CLC Bio, Aarhus, Denmark), and those from PacBio were assembled with PacBio SMRT Analysis 2.3.0 (Pacific Biosciences). ORFs were predicted and annotated using GeneMarkS (Besemer, Lomsadze, and Borodovsky, 2001) and rapid annotations using subsytems technology (RAST) (Aziz *et al*., 2008). The ribosomal binding sites (RBSs) were predicted using RBS finder (J. Craig Venter Institute, Rockville, Maryland, USA). The existing annotations were modified after cross-checking with results from InterProScan 5 (Jones *et al*., 2014) and GAMOLA (Altermann and Klaenhammer, 2003). The end product was submitted to GenBank. Using virulence factor database (VFDB; http://www.mgc.ac/VFs/) as reference, putative virulence factors of FORC_037 were found by basic local alignment search tool (BLAST). The circular genome maps of the chromosomes and plasmid were visualized by using cluster of orthologous group (COG)-based WebMGA and GenVision (DNASTAR, Madison, Wisconsin, USA).

Phylogenetic tree analysis and comparative genome analysis

Phylogenetic tree analysis was conducted using 16S rRNA sequences of FORC_037 and other species within the *Vibrio* genus. MEGA7 was used (Kumar, Strecher, and Tamura, 2016), while neighbor joining method was employed and the number of bootstrap replicates used was 1000. To characterize the genetic relatedness amongst the currently sequenced *V. vulnificus* strains, average

nucleotide identity (ANI) analysis was conducted. The completed genome sequences of 10 strains (MO6-24/O, CMCP6, 93U204, YJ016, ATL 6-1306, FORC_009, FORC_016, FORC_017, FORC_036, and FORC_037) were used for the analysis. The ANI values were calculated using Jspecies (Richter and Móra, 2009). The result was visualized with R. As CMCP6 was determined to be second most closely related to FORC_037, comparative genome analysis between FORC_037 and CMCP6 was carried out. Artemis comparison tool (ACT) was used (Carver *et al*., 2005).

RNA extraction

FORC_037 was first prepared so that it can be added to the control and the experimental conditions: FORC_037 was grown to mid-log phase (OD_{600} of 0.8), and centrifuged at 1,500 x g for 10 min. After removing the supernatant, the pellet was washed with 1X phosphate buffered saline (PBS), and the resulting solution was centrifuged at 5,000 x g for 10 min. The wash and centrifugation step was repeated two or three times. Lastly, the pellet was re-suspended with pre-warmed *Vibrio fisheri* minimal medium containing glycerol (VFMG) medium.

Prepared FORC_037 was incubated at 30°C for 4 h in VFMG medium either with or without small octopus. After incubation, the culture was first filtrated with syringe, which was filled with sterilized gauze, vacuum-filtered with Whatman no. 1 filter paper (Whatman International Ltd, Maidstone, England), and filtrated with another identically prepared syringe for the last time. The filtrated culture was

centrifuged at 5,000 x g at 4°C for 10 min. The pellet was re-suspended 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 the RBR-treated samples using miRNeasy Mini Kit (QIAGEN) according to the manufacturers' protocol. To prevent DNA contamination, TURBO DNase (AMbion, Austin, Texas, USA) was used. RNAs were cleaned up using RNeasy MinElute Cleanup Kit (QIAGEN). The quality of RNAs were confirmed by ChunLab Incorporation, with Agilent 2100 Bioanalyzer and Agilent RNA 6000 Nano reagents (Agilent Technologies, Waldbronn, Germany).

Strand-specific cDNA library construction and RNA sequencing

Strand-specific complementary DNA (cDNA) library construction and RNA sequencing were performed by ChunLab Incorporation. rRNAs were depleted by using Ribo-Zero™ rRNA Removal Kit (Epicentre, Madison, Wisconsin, USA). Enriched mRNA and TruSeq Stranded mRNA Sample Preparation kit (Illumina) was used to construct cDNA library. The quality of cDNA libraries was checked by ChunLab Incorporation, with Agilent 2100 Bioanalyzer and Agilent DNA 1000 reagents (Agilent Technologies). Strand-specific paired-ended 100 nucleotide reads from each cDNA library were obtained using Illumina HiSeq2500. For biological replication, two cDNA libraries were constructed and sequenced from RNAs isolated from two independent filtered culture supernatants of FORC_037.

Transcriptomic data analysis

The reads obtained from RNA sequencing were mapped to the reference genome of FORC_37 by using CLC Genomics Workbench 5.5.1 (CLC Bio). The GenBank (http://www.ncbi.nlm.nih.gov) accession numbers of the reference genome are CP016321 (Chromosome I), CP016322 (Chromosome II), and CP016323 (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). The genes were considered to be differentially expressed when the fold change was equal to 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).

RNA purification and transcript analysis

Total RNA from FORC_037 was isolated using RNeasy Mini Kit (QIAGEN). The culture condition was the same as above (See the 'RNA extraction' part). 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 genes and their specific primers used 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 experiment was conducted with biological duplicates and for each biological sample, technical triplicate wells were used.

Growth kinetics of FORC_037

The growth condition and bacteria preparation step was the same as described above (See the 'RNA extraction' part), except that the incubation times were 0 h, 2 h, 4 h, and 6 h. After the filtration step, the filtrated culture was serially diluted by 1X PBS and plated on LB agar supplemented with 2.0% (*w/v*) NaCl.

The plates were incubated in 30°C for 12 h and the colonies were counted. The experiments were run with biological duplicates and for each biological sample, technical triplicates were used.

Table 1. Bacterial strains used in this study

Bacterial strain	Relevant characteristics	Reference or Source
V. vulnificus		
MO6-24/O	Clinical isolate	Laboratory collection
FORC_037	Environmental isolate	Gyeonggi Health and Environmental Institute

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		
rtxA6_F	GCCAATATTCACGTGGGCGA	PCR analysis of *rtxA*	734

Oligonucleotide	Sequence (5` → 3`)	Use	Predicted size (bp)
rtxA6_R	CAAGCAGAGGCATCACCAGA		
nanA_F	TTATCGCCGCTCCCCATACA	PCR analysis of *nanA*	745
nanA_R	GCAACGCCACCGTATTCAAC		
gyrB_F	GTCCGCAGTGGAATCCTTCA	PCR analysis of *gyrB*	285
gyrB_R	TGGTTCTTACGGTTACGGCC	PCR analysis of *gyrB*	285
For qRT-PCR [b]			
FORC37_0694_qRT_F	GCAACTCTACTGCAACACGG	qRT-PCR of FOR37_0694	130
FORC37_0694_qRT_R	GCTTGCTGGAAACGCATCAT		
FORC37_0605_qRT_F	GTTCAAACAGATCGCTGGCTC	qRT-PCR of FOR37_0605	133
FORC37_0605_qRT_R	AATCGGCGTGGGTTCAGGA		
FORC37_1682_qRT_F	GGCTTGTCATTGAATCGGGC	qRT-PCR of FOR37_1682	128
FORC37_1682_qRT_R	AACGCTTGAGTCCGGTCATT		
FORC37_2291_qRT_F	GGTAACGACCGTGGTCAAGT	qRT-PCR of FOR37_2291	108
FORC37_2291_qRT_R	CGCTTTCTGCTCTTGGTCAT		
FORC37_3673_qRT_F	TTCATGGTGGAATGCGCTCT	qRT-PCR of FOR37_3673	184
FORC37_3673_qRT_R	TGAAAGGTGATCGTGTCGCA		

Oligonucleotide	Sequence (5` → 3`)	Use	Predicted size (bp)
FORC37_4282_qRT_F	TCGTCTTGTTTAACGTGGTGCATT	qRT-PCR of FOR37_4282	131
FORC37_4282_qRT_R	CACCATGGAAGCAGCGTGA		
FORC37_2197_qRT_F	AAGGAGAAACGCGCCAACTA	qRT-PCR of FOR37_2197	122
FORC37_2197_qRT_R	GCCATCCTCGATCGTCACTT		

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

[b] The oligonucleotides were designed using *V. vulnificus* FORC_37 genome sequence (GenBank™ accession numbers CP016321, CP016322, and CP016323).

III. RESULTS

Virulence gene-specific PCR screening

To check the virulent nature of FORC_037, which is an environmental strain, various virulence genes of *V. vulnificus* were amplified using PCR. *V. vulnificus* MO6-24/O and FORC_036 were used as positive controls. The target virulence genes and their expected sizes are listed in Table 2. All 6 genes were detected in all strains (Fig. 1).

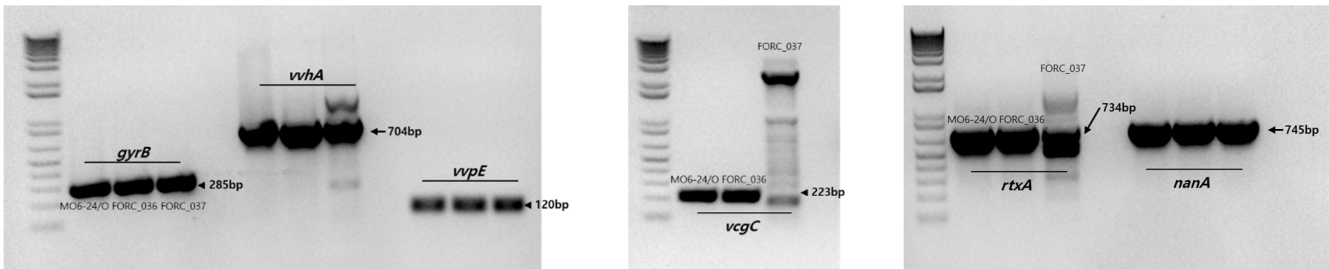

Fig 1. Virulence gene-specific PCR screening. 6 virulence genes of *V. vulnificus* were amplified using PCR. The size of each band is written as part of the figure. *vvhA*, hemolysin encoding gene; *vvpE*, elastase encoding gene; *vcgC*, virulence correlated gene; *rtxA*, repeated toxin protein A encoding gene; *nanA*, *N*-acetylneuraminate lyase encoding gene; *gyrB*, gyrase B encoding gene.

Cytotoxicity analysis of *V. vulnificus*

As FORC_037 is an environmental strain, it is necessary to estimate its virulence. Therefore, LDH release assay was performed using INT-407 cell lines. The host cells were infected with FORC_037 and, the clinical strain, *V. vulnificus* MO6-24/O was used as the positive control at various multiplicity of infections (MOIs) for 2 h. The result showed that FORC_037 may be as virulent as MO6-24/O (Fig. 2).

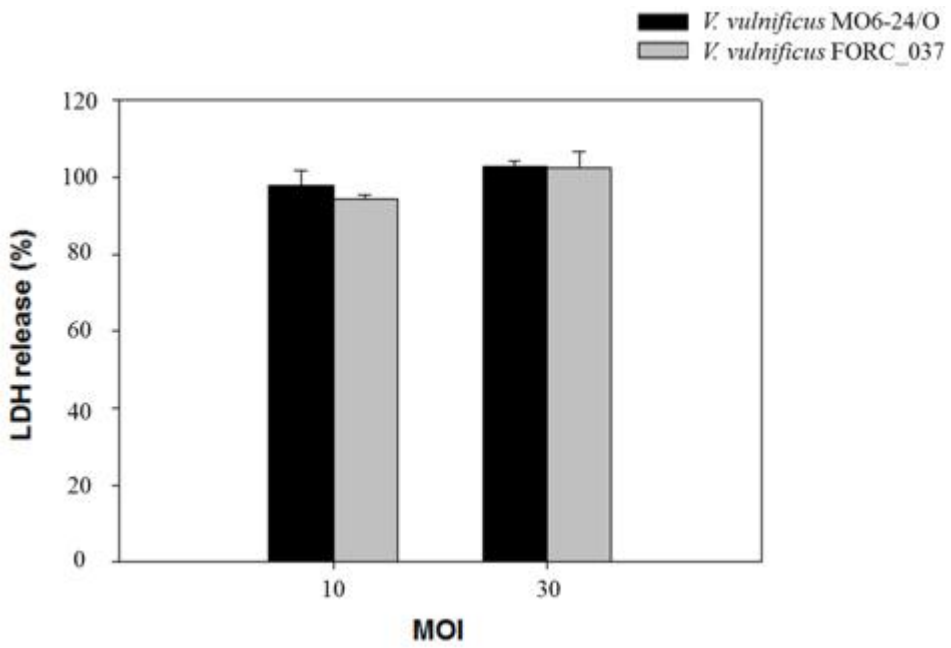

Fig 2. Cytotoxicity analysis of *V. vulnificus* strains. INT-407 cells were infected with *V. vulnificus* stains at various MOIs for 2 h. The cytotoxicity was determined by the LDH release activity. The total LDH (100%) is the amount of LDH released from the cells which were completely lysed by 2% Triton 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.

Transmission electron microscopy image

The morphology of FORC_037 was observed through transmission electron microscope (TEM). One can see that FORC_037 is curved and rod-shaped with a single polar flagellum. (Fig. 3).

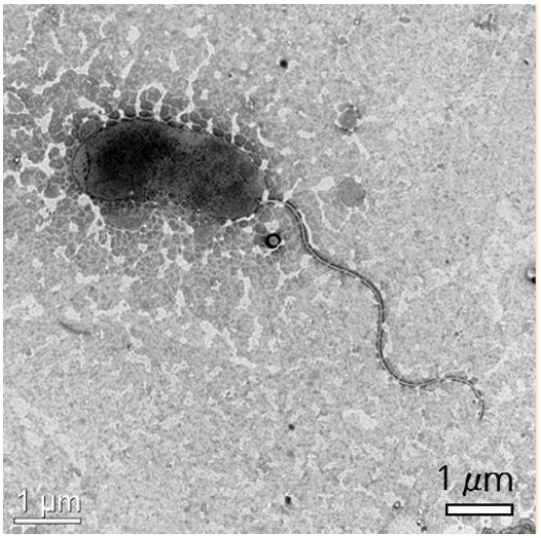

Fig 3. Transmission electron microscopy image of *V. vulnificus* FORC_037. The cells were negatively stained with PTA. Bars, 1 μm.

Genome properties of *V. vulnificus* FORC_037

The results from the whole genome sequencing was summarized in Table 3. The genome of *V. vulnificus* FORC_037 had two circular chromosomes and one plasmid. Chromosome I consists of 3,224,939 bp with a GC content of 46.73% containing 2,860 predicted ORFs, 105 tRNA genes and 31 rRNA genes. Among the predicted ORFs, 2,294 ORFs (80.21%) were predicted to encode the functional protein and 566 ORFs (19.79%) were expected to encode hypothetical proteins. Chromosome II consists of 1,852,101 bp with a GC content of 47.73% containing 1,571 predicted ORFs, 13 tRNA genes, and 3 rRNA genes. Among the 1,571, 1,237 ORFs (78.74%) were predicted to be functional and 334 ORFs (21.26%) were expected to encode hypothetical proteins. The plasmid, called pFORC37, consists of 40,850 bp with a GC content of 41.18% containing 75 predicted ORFs. Out of the 75, 17 ORFs (22.67%) were predicted to encode functional proteins and remaining 58 ORFs (77.33%) were expected to encode hypothetical proteins. All the results are summarized in Table 4.

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

Table 3. Summary of *V. vulnificus* FORC_037 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	463.28X
Contig length (bp)	3,224,939 (Chromosome I)
	1,852,101 (Chromosome II)
	40,850 (Plasmid)
Contig No.	3
Scaffold No.	3
N50	3,224,939
Locus tag	FORC37
GenBank accession No.	CP016321, CP016322, CP016323
GenBank release date	2018-06-28
BioProject No.	PRJNA323980
Source material identifier	FORC_037
Project relevance	Agricultural

Table 4. Chromosomal properties of *V. vulnificus* FORC_037

Property	Term		
Sample	*V. vulnificus* FORC_037		
Classification	Chromosome I	Chromosome II	Plasmid
GenBank accession No.	CP016321	CP016322	CP016323
Genome size (bp)	3,224,939	1,852,101	40,850
Protein coding genes	2,860	1,571	75
Annotated genes	2,294	1,237	58
Hypothetical genes	566	334	17
tRNA	105	13	0
rRNA	31	3	0

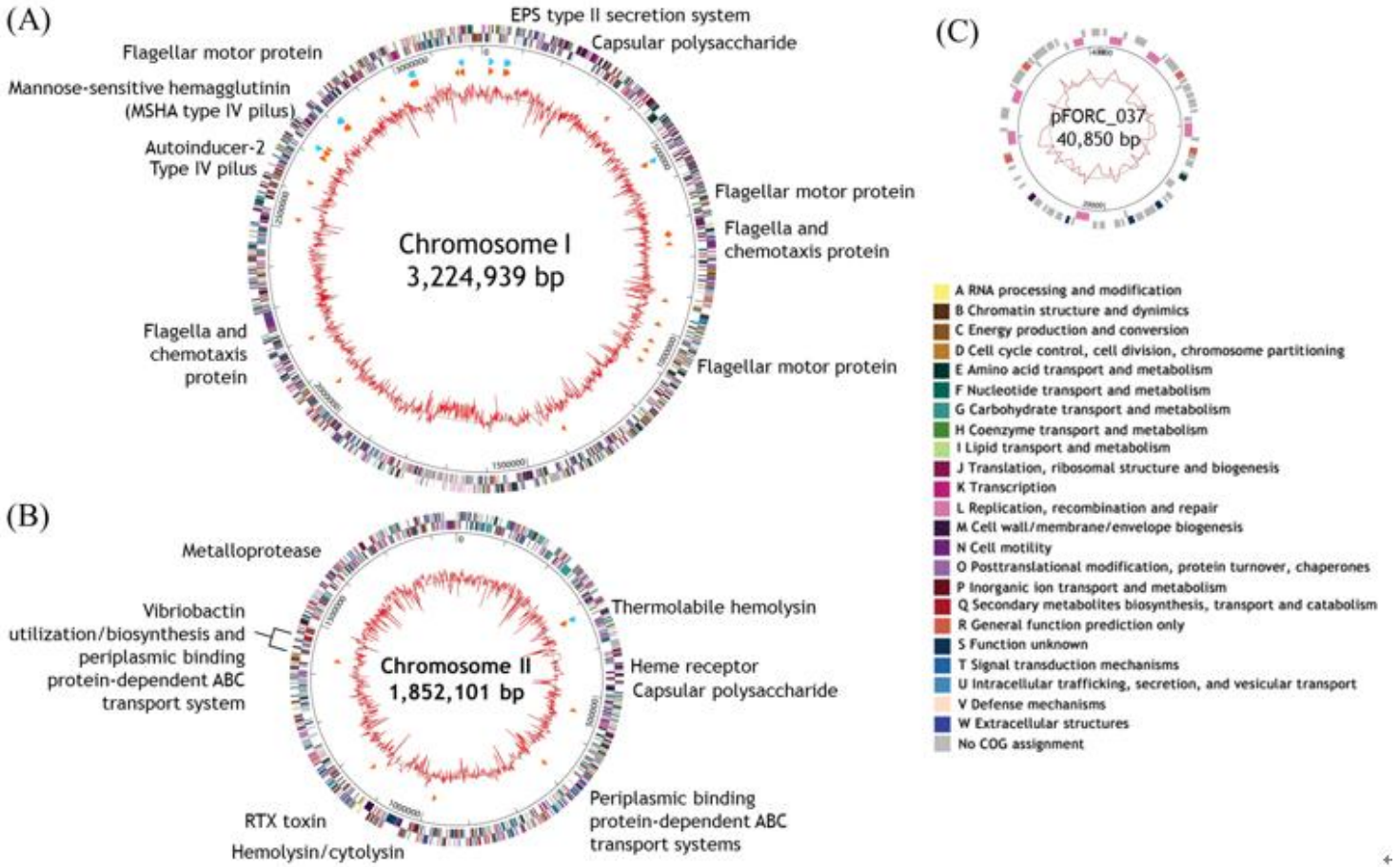
(A)
EPS type II secretion system
Capsular polysaccharide
Flagellar motor protein
Mannose-sensitive hemagglutinin (MSHA type IV pilus)
Autoinducer-2
Type IV pilus
Flagellar motor protein
Flagella and chemotaxis protein
Chromosome I
3,224,939 bp
Flagella and chemotaxis protein
Flagellar motor protein
(B)
Metalloprotease
Vibriobactin utilization/biosynthesis and periplasmic binding protein-dependent ABC transport system
Thermolabile hemolysin
Heme receptor
Capsular polysaccharide
Chromosome II
1,852,101 bp
RTX toxin
Hemolysin/cytolysin
Periplasmic binding protein-dependent ABC transport systems
(C)
pFORC_037
40,850 bp
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_037. (A) Chromosome I; (B) Chromosome II; (C) pFORC_037. The outer circle indicates the locations of all annotated ORFs, and the inner circle with red peaks indicates the GC contents. The orange arrows between the two circles are tRNAs and the sky blue arrows, rRNAs. All annotated ORFs are colored according to COG assignments and for more information on each category, refer to the legend on the right.

Pathogenesis and virulence factor

V. vulnificus is known to have various virulence factors in its genome. According to VFDB, virulence factors of *V. vulnificus* can be categorized as those related to adhesion, antiphagocytosis, chemotaxis and motility, exoenzymes, iron uptake, quorum sensing, secretion system, and toxin. Table 5 summarizes the virulence factors of FORC_037.

Table 5. Virulence factors of *V. vulnificus* FORC_037

Virulence factor	Gene product	Chromosome	Location (locus tag)
Adhesion			
- *mshD, mshC, mshA, mshB, mshF, mshG, mshE, mshN, mshM, mshL, mshK, mshJ, mshI, mshH*	Mannose-sensitive hemagglutinin (MSHA type IV pilus)	Chromosome I	-2772698~2786475 (FORC37_2469~2482)
pilA, pilB, pilC, pilD	Type IV pilus	Chromosome I	-2603296~2607201 (FORC37_2321~2324)
Antiphagocytosis			
- *wza* - *wzb, wzc* - *wcaJ* - *wbfY, wbfV/wcvB*	Capsular polysaccharide	Chromosome I	-219401~220537 (FORC37_0187) -221251~224097 (FORC37_0189~0190) -235235~235774 (FORC37_0201) -238201~241491 (FORC37_0205~ 0206)
- *cpsJ, cpI, cpsH, cpsF, cpsD, cpsC, cpsB, cpsA*	Capsular polysaccharide	Chromosome II	-473873~483671 (FORC37_3239~ 3246)
Chemotaxis			
- *cheV, cheR*	Chemotaxis protein	Chromosome I	-764506~766273 (FORC37_0683~ 0684)
- *cheW* - *cheB, cheA, cheZ, cheY*	Chemotaxis protein	Chromosome I	-2254991~2255485 (FORC37_2015) -2257366~2262004 (FORC37_2018~ 2021)
Motility			

- *flgN, flgM, flgA* - *flgB, flgC, flgD, flgE, flgF, flgG, flgH, flgI, flgJ, flgK, flgL, flaA, flaC*	Flagella	Chromosome I	-762765~764430 (FORC37_0680~ 0682) -766637~780841 (FORC37_0685~ 0697)
- *fliA, fleN, flhF, flhA, flhB, fliR, fliQ, fliP, fliO, fliN, filM, fliL, fliK, fliJ, fliI, fliH, fliG, fliF, fliE, fleQ, fleS, flrA, fliS, flaI, fliD, flaG, flaA, flaD* - *flaF*	Flagella	Chromosome I	-2262037~2291262 (FORC37_2022~ 2049) -2291568~2292701 (FORC37_2051)
- *motA, motB* - *motY* - *motX*	Flagellar motor protein	Chromosome I	-678988~680718 (FORC37_0604~ 0605) -1052602~1053483 (FORC37_0956) -2919397~2920032 (FORC37_2592)
Exoenzyme			
- *hapA/vvp*	Metalloproteases	Chromosome II	-1620654~1622483 (FORC37_4236)
Iron uptake			
- *hutR*	Heme receptors	Chromosome II	-441445~443580 (FORC37_3214)
- *vctG, vctC* - *vctD* - *vctP*	Periplasmic binding protein-dependent ABC transport systems	Chromosome II	-700500~702213 (FORC37_3454~3455) -699650~700507 (FORC37_3543) -1482219~1483127 (FORC37_4110)

- *vibD* - *vibF* - *vibA, vibC, vibE* - *vibB* - *vibH*	Vibriobactin biosynthesis	Chromosome II	-1463641~1464336 (FORC37_4098) -1467468~1472030 (FORC37_4100) -1473546~1477534 (FORC37_4102~4104) -1478519~1479406 (FORC37_4106) -1485617~1486933 (FORC37_4112)
- *viuB* - *viuA*	Vibriobactin utilization	Chromosome II	-1477640~1478455 (FORC37_4105) -366826~368889 (FORC37_4111)
Quorum sensing			
luxS	Autoinducer-2	Chromosome I	-2615124~2615642 (FORC37_2335)
Secretion system			
- *epsC, gspD, epsE, epsF, epsG, epsH, epsI, epsJ, epsK, epsL, epsM, epsN*	EPS type II secretion system	Chromosome I	-117451~128770 (FORC37_0090~0101)
Toxin			
- *vvhA*	Hemolysin/cytolysin	Chromosome II	-1019639~1021054 (FORC37_3714)
- *rtxA, rtxC* - *rtxB, rtxD*	RTX toxin	Chromosome II	-1105142~1111496 (FORC37_3780~3781) -1112243~1115784 (FORC37_3783~3784)
- *tlh*	Thermolabile hemolysin	Chromosome II	-361457~362710 (FORC37_3156)

Phylogenetic analysis and ANI analysis

To check the phylogenetic status of *V. vulnificus* FORC_037, the 16S rRNA sequence of FORC_037 was compared with that of other *Vibrio* species (*V. vulnificus*, *V. parahaemolyticus*, and *V. cholerae*). The result demonstrated that FORC_037 indeed belongs to *V. vulnificus* (Fig. 5).

Additionally, ANI analysis was performed using the chromosomal sequences of total ten *V. vulnificus* strains: MO6-24/O, CMCP6, 93U204, YJ016, ATL 6-1306, FORC_009, FORC_016, FORC_017, FORC_036, and FORC_037. The result showed that FORC_037 had the second highest ANI value (96.565) with CMCP6, a clinical isolate from Chonnam National University Hospital (Kim *et al*., 2003) (Fig. 6). Although FORC_037 had the highest ANI value (96.72) with FORC_017, a clinical isolate from a 63-year-old woman with symptoms of hemorrhagic rash after consuming a raw dotted gizzard shad on Jeju Island (Chung *et al.,* 2016), CMCP6 was selected for further comparative genome analysis as CMCP6 is regarded as one of the type strains.

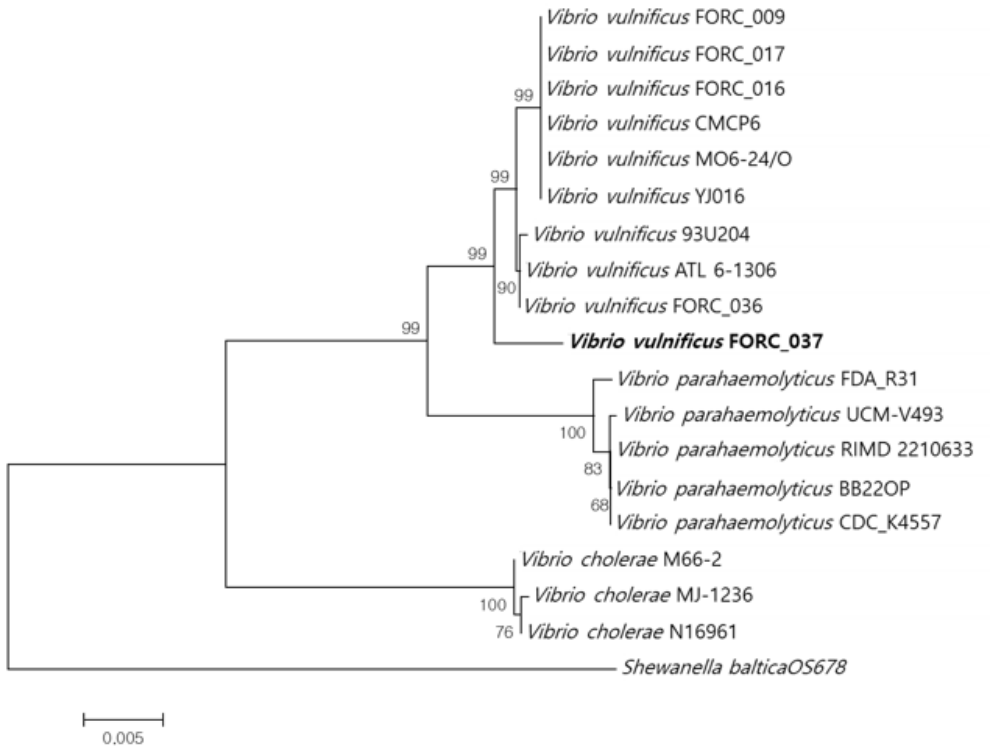

Fig 5. Phylogenetic tree of FORC_037 with *Vibrio* species. The 16S rRNA sequences from various *Vibrio* species (*V. vulnificus, V. parahaemolyticus,* and *V. cholerae*) were aligned by using ClustalW. The phylogenetic tree was constructed using the neighbor joining method with 1000 bootstrap replicates. *S. balitca* OS678 was used as the outgroup. The scale bar is below the tree and the bootstrap values are on the branch.

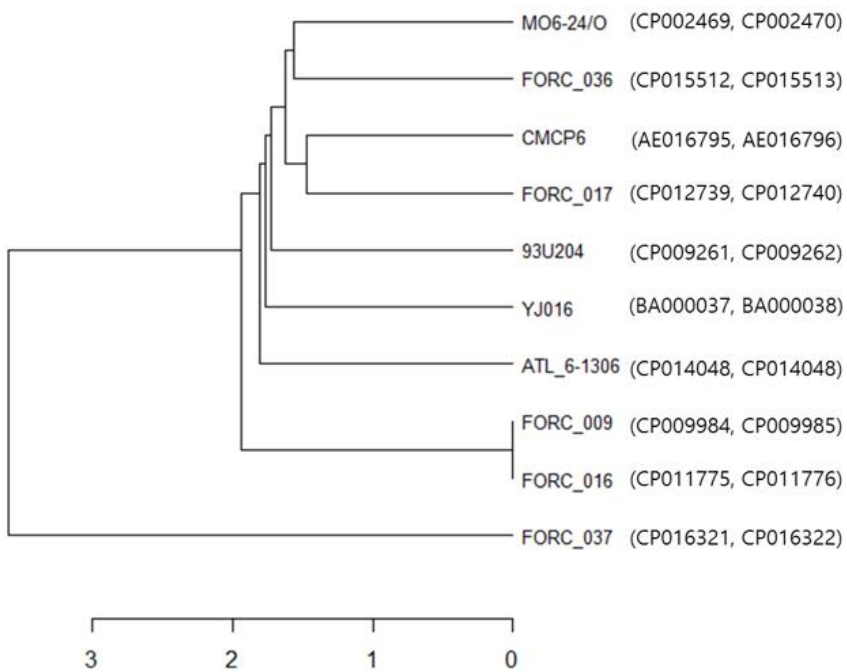

Fig 6. ANI analysis of *V. vulnificus* strains. The figure was created with R using ANI values, which were computed based on chromosomal sequences of the selected *V. vulnificus* strains (i.e. plasmids were excluded).

Comparative genome analysis between FORC_037 and *V. vulnificus* CMCP6

The comparative genome analysis between FORC_037 and CMCP6 indicated that FORC_037 has a non-homologous region which may be relevant to virulence (Fig. 7). FORC37_3608 to FORC37_3622 is the region and among the genes, FORC37_3618 and FORC37_3619 encode accessory cholera enterotoxin and zona occludens toxin, respectively (Fig. 7). Table 6 shows the products encoded by the genes on the non-homologous region.

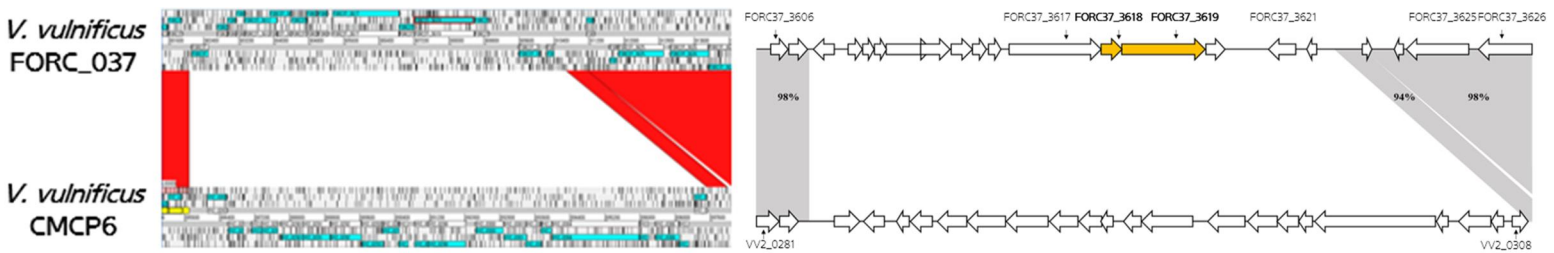

Fig 7. Comparative genome analysis between FORC_037 and *V. vulnificus* CMCP6. The region included phage protein, accessory cholera enterotoxin and zona occludens toxin (left). The red regions are the homologue regions and the white regions are the non-homologous regions (right). The direction of each arrow corresponds to the directionality of each corresponding gene. The yellow arrows represent the functionally important genes in FORC_037. The gray regions, which correspond to the red regions on the left, respectively were 98%, 94%, and 98% homologous.

Table 6. Non-homologous region of FORC_037 compared to CMCP6

Locus tag	Gene product
FORC37_3617	Phage protein
FORC37_3618	Accessory cholera enterotoxin
FORC37_3619	Zona occludens toxin

Identification of differentially expressed genes of FORC_037 upon exposure to small octopus

Transcriptome analysis was performed to characterize the transcriptional profile of FORC_037 upon exposure to small octopus. RNA samples extracted from FORC_037 incubated in VFMG with or without small octopus were sequenced. Average RPKM values from the biological duplicate samples were used to represent the expression level of each gene. Figure 8 is a volcano plot which depicts the overall transcriptome profile. Genes that are differently expressed significantly with statistical significance (fold change $\geq \pm 2$, P-value < 0.05) were identified. A total of 1,301 genes were differently expressed. Among them, 752 genes were up-regulated and 549 genes were down-regulated.

The motility related genes were down-regulated and genes related to adherence, galactose utilization, and iron uptake were up-regulated (Table 7). Additionally, heat map was used to visualize the RNA sequencing result (Fig. 9).

To confirm the result of RNA sequencing, qRT-PCR was conducted (Fig. 10).

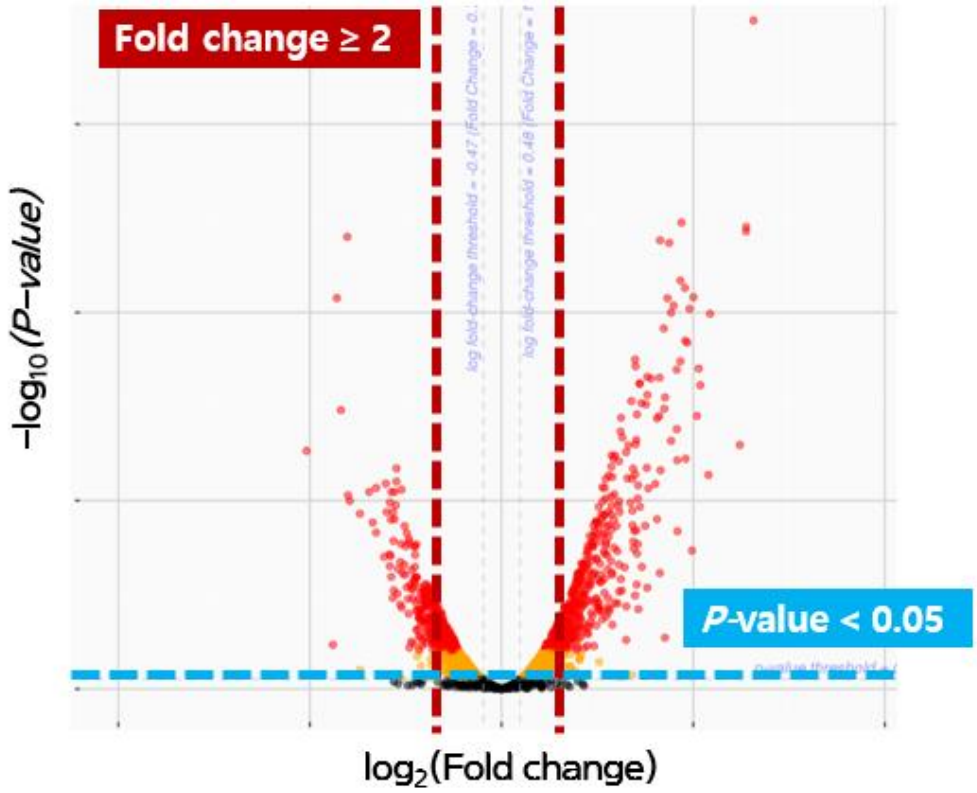

Fig 8. Transcriptome comparison of the RNA sequencing samples. The volcano plot was created to express differently expressed genes when FORC_037 was exposed to small octopus.

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

Locus tag	Gene product	Fold change [a]	*P*-value [b]
Motility			
FORC37_0604	Flagellar motor rotation protein MotA	-1.83	0
FORC37_0605	Flagellar motor rotation protein MotB	-1.98	0
FORC37_2031	Flagellar motor switch protein FliN	-1.06	0.000105184
FORC37_2032	Flagellar motor switch protein FliM	-1.32	0
FORC37_3406	Ribose ABC transport system, periplasmic ribose-binding protein RbsB	3.13	0
FORC37_4426	Signal transduction histidine kinase CheA	-1.10	0.000172732
FORC37_4427	Fis family transcriptional regulator	-1.58	0
FORC37_0693	Flagellar protein FlgJ / peptidoglycan hydrolase	-1.99	0
FORC37_0696	Flagellin protein FlaC	-2.74	0
FORC37_0697	Flagellin protein FlaD	-2.93	0
FORC37_0956	Sodium-type flagellar protein MotY	-1.96	0
FORC37_2045	Flagellar rod protein flaI	-1.84	0

Locus tag	Gene product	Fold change [a]	*P*-value [b]
FORC37_2049	Flagellin protein FlaD	-2.53	0
FORC37_2592	Sodium-type flagellar protein MotX	-1.65	0
FORC37_0680	Flagellar biosynthesis protein FlgN	-1.75	0
FORC37_0681	Negative regulator of flagellin synthesis FlgM	-1.80	0
FORC37_0685	Flagellar basal-body rod protein FlgB	-1.88	0
FORC37_0686	Flagellar basal-body rod protein FlgC	-1.87	0
FORC37_0687	Flagellar basal-body rod modification protein FlgD	-1.79	0
FORC37_0688	Flagellar hook protein FlgE	-1.54	0
FORC37_0689	Flagellar basal-body rod protein FlgF	-2.62	0
FORC37_0690	Flagellar basal-body rod protein FlgG	-2.91	0
FORC37_0691	Flagellar L-ring protein FlgH	-2.88	0
FORC37_0692	Flagellar P-ring protein FlgI	-2.26	0
FORC37_0694	Flagellar hook-associated protein FlgK	-2.82	0
FORC37_0695	Flagellar hook-associated protein FlgL	-2.50	0
FORC37_2033	Flagellar biosynthesis protein FliL	-1.51	0
FORC37_2034	Flagellar hook-length control protein FliK	-2.43	0

Locus tag	Gene product	Fold change [a]	*P*-value [b]
FORC37_2044	Flagellar biosynthesis protein FliS	-1.99	0
FORC37_2046	Flagellar hook-associated protein FliD	-1.32	0
FORC37_2051	Flagellin protein FlaF	-2.74	0
Adherence			
FORC37_1681	Flp pilus assembly protein TadD, contains TPR repeat	2.95	0
FORC37_1682	Type II/IV secretion system protein TadC, associated with Flp pilus assembly	2.88	0
FORC37_1683	Flp pilus assembly protein TadB	2.25	0
FORC37_1684	Flp pilus assembly protein TadA	2.76	0
FORC37_1685	Septum site-determining protein MinD	2.79	0
FORC37_1687	Flp pilus assembly protein, secretin CpaC	2.83	0
FORC37_2267	Outer membrane protein OmpU	1.50	0
Iron uptake			
FORC37_2289	Ferric iron ABC transporter, ATP-binding protein	2.89	0
FORC37_2290	Ferric iron ABC transporter, permease protein	2.67	0
FORC37_2291	Ferric iron ABC transporter, iron-binding protein	5.10	0

Locus tag	Gene product	Fold change [a]	*P*-value [b]
FORC37_3214	TonB-dependent heme receptor HutR	5.00	0
FORC37_3273	Ferric siderophore transport system, biopolymer transport protein ExbB	4.01	0
FORC37_3274	Ferric siderophore transport system, periplasmic binding protein TonB	4.05	0
FORC37_3671	Ferric siderophore transport system, periplasmic binding protein TonB	4.42	0
FORC37_3673	Ferric siderophore transport system, biopolymer transport protein ExbB	4.23	0
FORC37_3453	Ferric vibriobactin, enterobactin transport system, permease protein VctD	4.05	0
FORC37_3454	Ferric vibriobactin, enterobactin transport system, permease protein VctG	3.41	0
FORC37_3455	Ferric vibriobactin, enterobactin transport system, ATP-binding protein	2.59	0
FORC37_4110	Catechol siderophore ABC transporter, substrate-binding protein	4.80	0
FORC37_4098	Phosphopantetheinyl transferase component of siderophore synthetase	3.28	0
FORC37_4100	hypothetical protein	4.42	0
FORC37_4102	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase	3.66	0
FORC37_4103	Isochorismate synthase of siderophore biosynthesis	3.25	0
FORC37_4104	2,3-dihydroxybenzoate-AMP ligase	2.47	0
FORC37_4106	Isochorismatase of siderophore biosynthesis	5.44	0
FORC37_4105	Vulnibactin utilization protein VuuB	4.84	0

Locus tag	Gene product	Fold change [a]	*P*-value [b]
FORC37_4111	Ferric vulnibactin receptor VuuA	3.44	0
FORC37_4272	Hypothetical protein in aerobactin uptake cluster	3.36	0
FORC37_4273	Ferric aerobactin ABC transporter, ATPase component	4.90	0
FORC37_4274	Ferric aerobactin ABC transporter, periplasmic substrate binding protein	3.80	0
FORC37_4275	Ferric aerobactin ABC transporter, permease component	3.80	0
Galactose metabolism			
FORC37_0625	PTS system, trehalose-specific IIB component / PTS system, trehalose-specific IIC component	-3.70	0.003081
FORC37_3535	Fructose-specific phosphocarrier protein HPr / PTS system, fructose-specific IIA component	-1.72	3.72E-05
FORC37_2197	Evolved beta-D-galactosidase, beta subunit	3.52	0
FORC37_4280	N-acetylgalactosamine-6-phosphate deacetylase	1.79	0.000803
FORC37_4281	PTS system, N-acetylgalactosamine- and galactosamine-specific IIA component	1.92	0.003939
FORC37_4282	PTS system, N-acetylgalactosamine-specific IID component	2.02	0.000161
FORC37_4283	PTS system, N-acetylgalactosamine-specific IIC component	1.14	0.014326

[a] The mRNA expression level in *V. vulnificus* FORC_037 after four hours' exposure to small octopus relative to the mRNA expression level without exposure to small octopus.

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

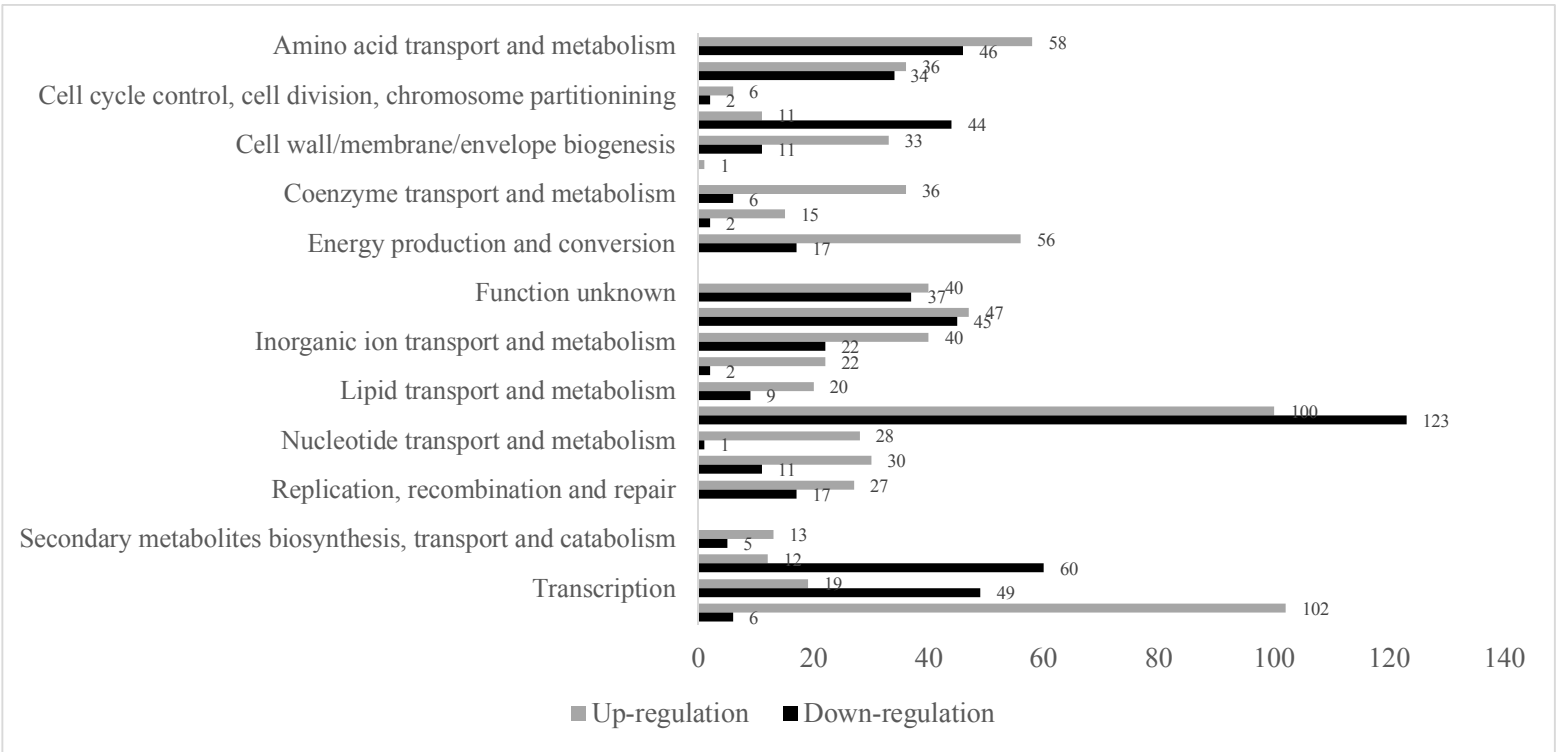

Fig 9. Functional categorization of genes differentially expressed upon exposure to small octopus. Genes were considered to be differentially expressed when p-value < 0.05 and either fold change ≥ 2 or fold change ≤ -2.

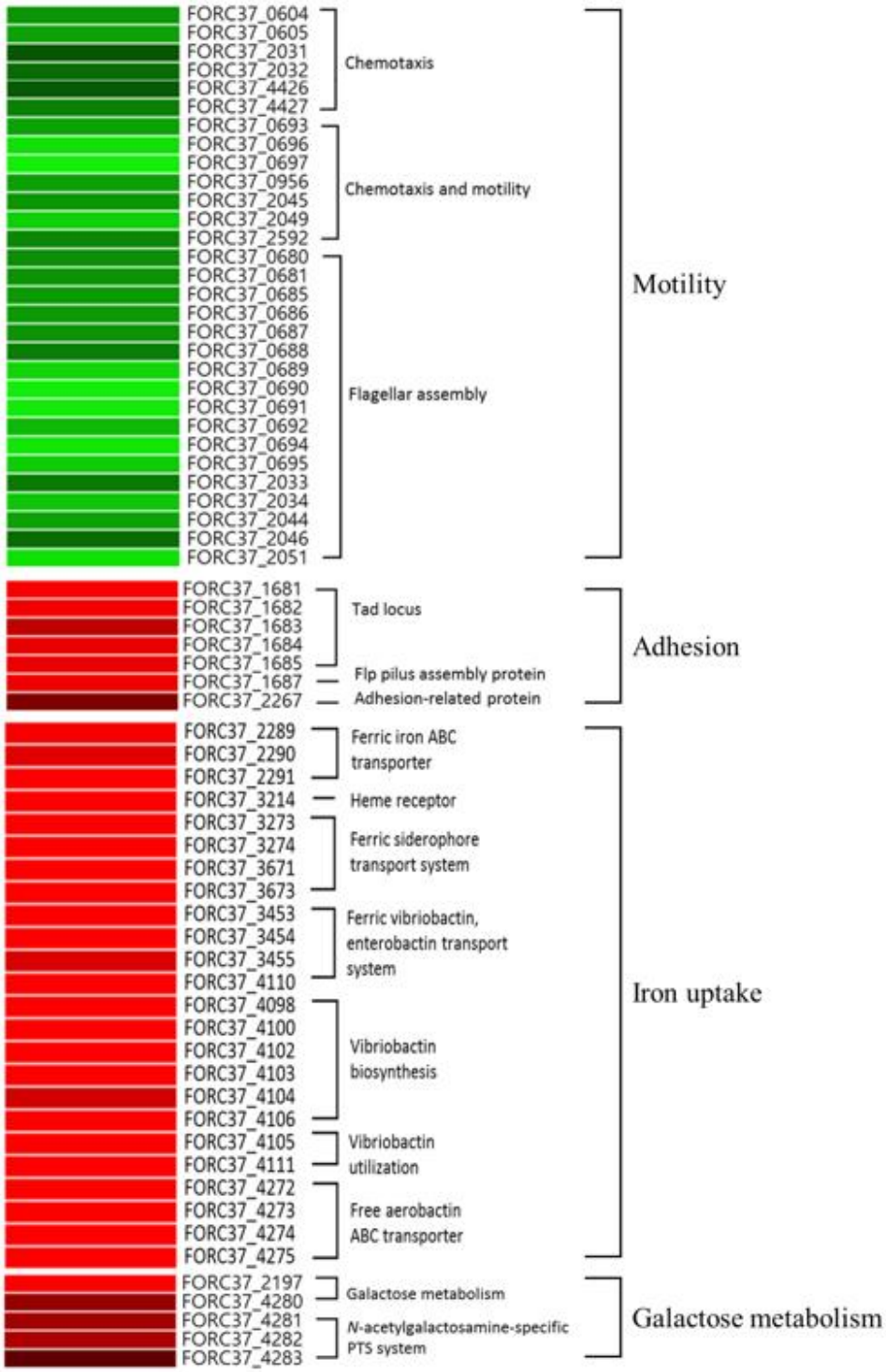
FORC37_0604
FORC37_0605
FORC37_2031
FORC37_2032
FORC37_4426
FORC37_4427
Chemotaxis
FORC37_0693
FORC37_0696
FORC37_0697
FORC37_0956
FORC37_2045
FORC37_2049
FORC37_2592
Chemotaxis and motility
FORC37_0680
FORC37_0681
FORC37_0685
FORC37_0686
FORC37_0687
FORC37_0688
FORC37_0689
FORC37_0690
FORC37_0691
FORC37_0692
FORC37_0694
FORC37_0695
FORC37_2033
FORC37_2034
FORC37_2044
FORC37_2046
FORC37_2051
Flagellar assembly
Motility
FORC37_1681
FORC37_1682
FORC37_1683
FORC37_1684
FORC37_1685
Tad locus
FORC37_1687
Flp pilus assembly protein
FORC37_2267
Adhesion-related protein
Adhesion
FORC37_2289
FORC37_2290
FORC37_2291
Ferric iron ABC transporter
FORC37_3214
Heme receptor
FORC37_3273
FORC37_3274
FORC37_3671
FORC37_3673
Ferric siderophore transport system
FORC37_3453
FORC37_3454
FORC37_3455
FORC37_4110
Ferric vibriobactin, enterobactin transport system
FORC37_4098
FORC37_4100
FORC37_4102
FORC37_4103
FORC37_4104
FORC37_4106
Vibriobactin biosynthesis
FORC37_4105
FORC37_4111
Vibriobactin utilization
FORC37_4272
FORC37_4273
FORC37_4274
FORC37_4275
Free aerobactin ABC transporter
Iron uptake
FORC37_2197
FORC37_4280
Galactose metabolism
FORC37_4281
FORC37_4282
FORC37_4283
N-acetylgalactosamine-specific PTS system
Galactose metabolism

Fig 10. Heat map of selected genes' transcriptome after 4 h exposure to small octopus. The red bars represented the up-regulated genes and the green bars indicated that the down-regulated genes. The scale bar is above the heat map.

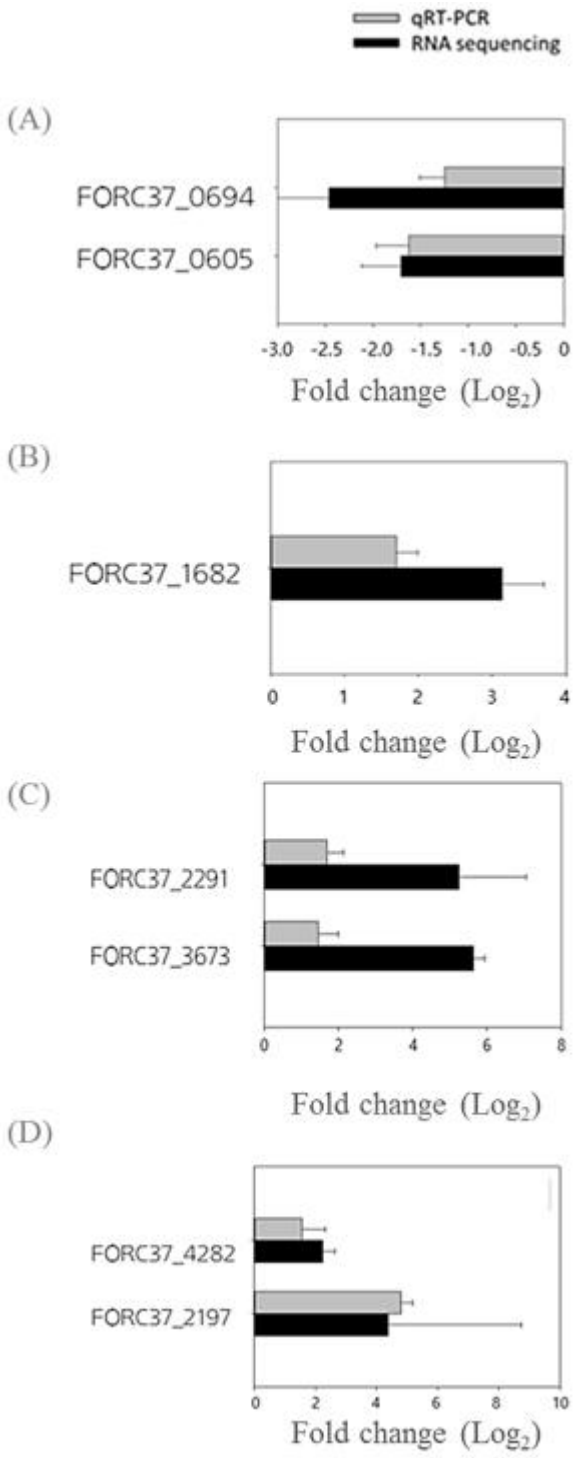
qRT-PCR
RNA sequencing
(A)
FORC37_0694
FORC37_0605
-3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0
Fold change (Log2)
(B)
FORC37_1682
0 1 2 3 4
Fold change (Log2)
(C)
FORC37_2291
FORC37_3673
0 2 4 6 8
Fold change (Log2)
(D)
FORC37_4282
FORC37_2197
0 2 4 6 8 10
Fold change (Log2)

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, (B) adhesion related gene, (C) iron uptake related gene, and (D) galactose metabolism related gene were confirmed, respectively. Table 7 shows the products of the represented genes. The error bars represent the standard deviations.

Growth kinetics of FORC_037 exposed to small octopus

Because the *N*-acetylgalactosamine is one of the major components of small octopus skin (Florkin and Scheer, 1970) and *V. vulnificus* can transport this molecule, FORC_037 may use this nutrient for growth when exposed to small octopus. The growth kinetics were compared between two samples which were incubated in VFMG either with or without small octopus, respectively. FORC_037 exposed to the small octopus grew faster than the control and the colony forming unit (CFU) was 10 fold higher (Fig. 11).

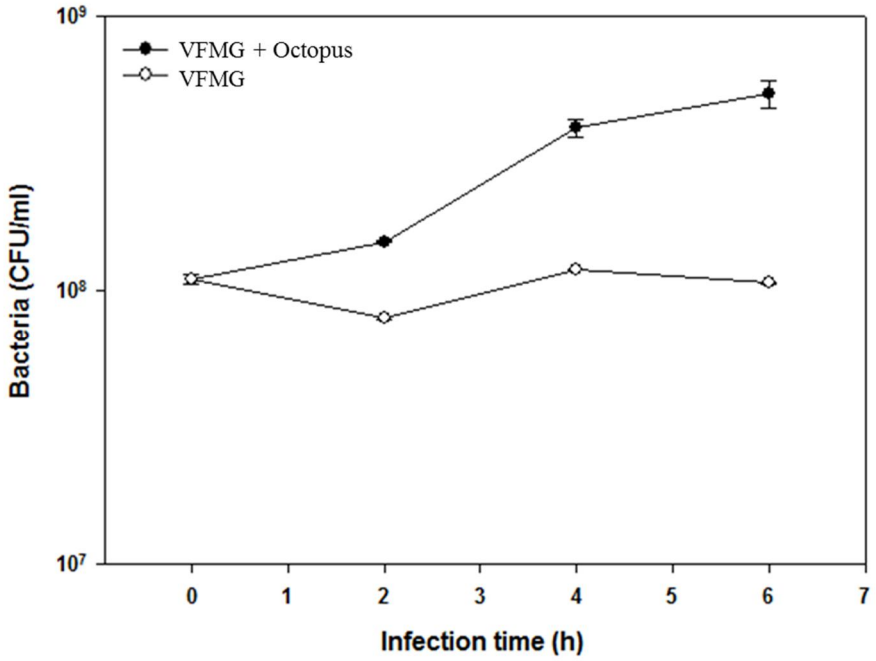

Fig 12. Growth kinetics of FORC_037 incubated in VFMG either in the presence or absence of small octopus. The CFUs were counted at the specific time points (0 h, 2 h, 4 h, and 6 h). The experiments were repeated with biologically duplicated samples and technically triplicate were run for each sample. Open circle (○), the small octopus-exposed sample; closed circle (●), control sample. The error bars represented the standard deviations.

IV. DISCUSSION

V. vulnificus is an opportunistic pathogen, which can cause gastroenteritis and deadly septicemia in susceptible individuals. Every summer in South Korea, *V. vulnificus* is isolated from seawater and food sources, and fatalities are reported. Preventing future outbreaks can be aided by studying the genome and by analyzing the change in transcriptome of the pathogen, upon contact with food. In this study, FORC_037, a *V. vulnificus* strain isolated from a soft-shell clam, was analyzed.

In order to characterize the potential virulence of FORC_037, virulence gene-specific PCR screening and LDH release assay were performed. FORC_037 had various virulence genes and showed a high level of cytotoxicity toward INT-407 cell lines (Figures 1 and 2). The result suggested that this strain may be pathogenic toward human cells.

FORC_037's genomic DNA was extracted and sequenced, and its complete genome sequence was analyzed. The phylogenetic tree analysis using 16S rRNA sequence of FORC_037 with those of other *V. vulnificus*, *V. cholerae* and *V. parahaemolyticus* strains, showed that FORC_037 is indeed *V. vulnificus*. But ANI analysis suggested that FORC_037 is an outlier as this strain seemed to show a lower ANI value compared to other *V. vulnificus* strains (Table 4).

In order to take a closer look at the strain's genome, BLAST searches were performed. The goal was to predict whether the genome had the virulence factors that are present in genomes of virulent strains of *V. vulnificus* (Table 5). The major virulence factors of *V. vulnificus* such as a hemolysin/cytolysin (VvhA) and repeats-

in-toxin (Rtx) (Jones and Oliver, 2009) were located in the genome of this strain. Also, FORC_037's genome harbors genes that encode Type II secretion system (T2SS), which is responsible for secreting folded proteins from periplasm to extracellular environment (Costa *et al*., 2015). T2SS is required to export at least two virulence factors, VvhA and VvpE in *V. vulnificus* (Hwang *et al*., 2011). FORC_037's genome also harbored multiple iron uptake systems. Iron acquisition from the host cell is important for *V. vulnificus* to survive and grow (Wright, Simpson and Oliver, 1981). In short, although FORC_037 is not a clinical isolate, its genome harbors the various virulence factors, such as hemolysin/cytolysin, RTX toxins, secretion systems for toxin transport, and iron uptake systems. Also, the level of cytotoxicity of FORC_037 was similar to that of a clinical isolate. This is an example of an environmental isolate which shows a potential to be as virulent as a clinical isolate. Furthermore, comparative genome analysis was carried out between FORC_037 and CMCP6, its second most closely related strain according to ANI analysis. FORC_037 has a non-homologous region, which carries accessory cholera enterotoxin and zonula occludens toxin, which have been reported to play a role in increasing the permeability of the small intestinal mucosa by affecting the structure of the tight junction. This suggests that FORC_037 may have an advantage in surviving in the human intestine.

To examine the behavior of *V. vulnificus* FORC_037 when this strain is exposed to small octopus, RNA sequencing was performed using two RNA samples which were extracted from the bacterium incubated in VFMG either with or without

small octopus. Genes related to motility were down-regulated but the genes related to adherence were up-regulated. So, FORC_037 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). Expression of genes related to the galactose metabolism and the *N*-acetylgalactosamine transporter was up-regulated. As *N*-acetylgalactosamine is a major component of the small octopus's skin (Florkin and Scheer, 1970), this expression change indicates that *V. vulnificus* FORC_037 may be utilizing this nutrient for survival and growth when in contact with a small octopus. Lastly, various iron uptake systems were up-regulated.

These combined results suggest that FORC_037, when exposed to small octopus, adapts in four different ways: (1) reduction in motility, (2) attachment to the small octopus, (3) utilization of *N*-acetylgalacotsamine from small octopus skin, and (4) increased iron uptake. This suggests that FORC_037 perceives the small octopus as a reservoir rather than a host. This study contributes the accumulation of database about the domestic pathogens and helps to deal with the coming outbreak involved in *V. vulnificus*. Thus studying the genetic program set in motion upon contact with small octopus suggests that better practices in processing small octopuses should be recommended, given that the bacterium seems to exhibit the ability adhere to the surface and use the small octopus as a reservoir.

V. REFERENCES

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.

Chung, H. Y., Kim, Y., Kim, S., Na, E. J., Ku, H., Lee, K. H., Heo, S. T., Ryu, S., Kim, H., Choi, S., H., & Lee, J. (2016). Complete genome sequence of *Vibrio vulnificus* FORC_017 isolated from a patient. *Gut Pathogens*, 8(1), 22.

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.

Bergholz, T. M., Switt, A. I. M., & Wiedmann, M. (2014). Omics approaches in food safety: fulfilling the promise? *Trends in Microbiology*, 22(5), 275-281.

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.

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.

Ku, H., & Lee, J. (2014). Development of a Novel Long-Range 16S rRNA Universal Primer Set for Metagenomic Analysis of Gastrointestinal Microbiota in Newborn Infants. *Journal of Microbiology and Biotechnology,* 24, 812-822.

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.

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.

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., Galen, J. E., Michalski, J., Fasano, A., Kaper, J. B. (1993). Accessory cholera enterotoxin (Ace), the third toxin of a Vibrio cholerae virulence cassette. *Proceedings of National Academy of Science*, 90(11), 5267-5271.

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 해안가에 자연적으로 존재하며 인간에 기회감염을 일으켜 상처를 덧나게 하거나, 위장질환, 혹은 패혈증까지 일으킬 수 있는 식중독균이다. 매년 발생하는 패혈증 비브리오균으로 인해 발생하는 피해를 예방하기 위해 유전체 및 전사체 수준의 연구가 필요하다. 우럭조개로부터 분리된 패혈증 비브리오균 'FORC_037'의 독성을 PCR과 세포주 실험을 통해 확인했고 이 균주의 특성을 더 알아보기 위해서 유전체와 전사체 분석을 진행하였다.

FORC_037은 두 개의 염색체와 하나의 플라스미드를 가지고 있고, 4,506개의 전사해석틀, 118개의 운반 RNA와 34개의 리보솜 RNA를 지니고 있다고 예측된다. FORC_037은 주요 독성인자 및 병원성에 기여하는 다양한 유전자들을 가지고 있었다. 비브리오 속의 16S 리보솜 RNA를 이용해 계통 분석을 진행했고, 따라서 FORC_037이 패혈증 비브리오균에 속한다는 것을 확실히 하였다. 또한, 유전체가 완전히 밝혀져 NCBI에 등재된 총 9개의 타 패혈증 비브리오균과의 종 간 유사도를 확인한 결과, FORC_037이 전반적으로 타 패혈증 비브리오균들과 비교할 때 유사도가 낮다는 것과, 환자분리 균주인 CMCP6와 계통학적으로 거의 가장 유사함을 알 수 있었다.

이에 CMCP6와의 유전체 비교 분석을 통해 FORC_037이 accessory cholera enterotoxin과 zona occludens toxin 유전자를 추가적으로 가지고 있음을 확인하였다.

낙지는 한국에서 날 것으로 섭취되는 식품 중 하나이고, 낙지 섭취 후 패혈증 환자가 발생한 국내 사례가 있기에 낙지를 모델 식품으로 선정했다. FORC_037이 낙지에 노출되었을 때 유전자 발현양상 변화를 RNA sequencing을 이용해 확인했을 때, 운동성과 관련된 유전자는 발현이 감소했지만, 부착, 철 흡수, 유당 대사와 관련된 유전자의 발현은 증가했다.

우럭조개로부터 분리된 FORC_037은 여러 독성인자를 가져 병원성을 나타낼 가능성이 높고, 낙지에 노출되었을 때 낙지를 저장소로 인지해 성장할 것으로 예측된다. 본 연구를 통해 국내 식중독균에 대한 데이터베이스를 축적함으로써 이후 패혈증 비브리오균에 의한 질병 발생을 예방하는 것에 도움이 될 것이라 기대된다.

주요어: 패혈증 비브리오균, 식중독균, 유전체, 전사체, 전장유전체 시퀀싱, 리보 핵산 시퀀싱, 유전체 비교분석

학 번: 2015-23131