International Journal of Food Microbiology 244 (2017) 90–95

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Vibrio parahaemolyticus O4:K8 forms a potential predominant clone in southern China as detected by whole-genome sequence analysis

Baisheng Li, Xingfen Yang, Hailing Tan, Bixia Ke, Dongmei He, Changwen Ke, Yonghui Zhang *

Center for Diseases Control and Prevention of Guangdong Province, China

ARTICLE INFO

Article history:
Received 14 July 2016
Received in revised form 3 January 2017
Accepted 3 January 2017
Available online 04 January 2017

Keywords:
Vibrio parahaemolyticus
Whole genome sequence
Phylogenetic analysis
SNPs

ABSTRACT

Vibrio parahaemolyticus has been the most common food-borne pathogen in southern China, especially the O3:K6 pandemic clone and its serovariants. Recently, the serotype O4:K8 became more and more prevalent in southern China, which was different from the O3:K6 pandemic clone. Thus, the aim of the present work was to elucidate the molecular characteristics of the O4:K8. Some O3:K6 pandemic clone and its serovariants isolated in the same period were selected for comparative analysis, which were still dominant clone locally. The whole genome sequencing (WGS) was applied to characterize 20 strains of *V. parahaemolyticus* isolated from food-borne diarrheal cases and belonging to the serotype O4:K8, O3:K6 and O1:KUT (untypable), prevalent serotypes in recent southern China. The results showed that all these isolates were positive for the thermostable direct hemolysin gene (*tdh*), while negative for the TDH-related hemolysin gene (*trh*). We compared the *V. parahaemolyticus* strains to those of 31 strains isolated overseas and were available from NCBI genome database. A WGS-SNPs phylogenetic analysis of all the genomes revealed that the strains formed an important genetic lineage, which was genetically distinct from the O3:K6, O1:KUT and other internationals strains. Comparative genome analysis also revealed that all the O4:K8 strains carried the entire T3SS-1 and VpaI-7 (T3SS-2) regions, the most important virulent elements of the O3:K6 pandemic clone. However, all the O4:K8 strains lacked the entire VpaI-1 and VpaI-4 regions and carried only few ORFs of the VpaI-5 and VpaI-6, which were considered to be unique among post-1995 strains belonging to the O3:K6 pandemic clone. Our data showed that the O4:K8 strains possessed the virulence factors similar to the O3:K6 pandemic clone, which may have enabled them to become prevalent in southern China. Our study also revealed that WGS-bases analysis may help improve understanding epidemiology of this bacterium in food-borne disease surveillance.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Vibrio parahaemolyticus is a curved, rod-shaped, Gram-negative bacterium that inhabits brackish saltwater. Most *V. parahaemolyticus*-borne diarrhea cases were mediated by the consumption of raw or undercooked seafood, usually oysters (Vongxay et al., 2008). Isolates carrying one or both of the virulence factors such as thermostable direct hemolysin (TDH) and TDH-related hemolysin (TRH) are considered pathogenic (Shirai et al., 1990). Previous studies revealed that diseases caused by *V. parahaemolyticus* were sporadic before 1996 and usually attributed to multiple serotypes (Chen et al., 2011). However, a variant of *V. parahaemolyticus* serotype (O3:K6) reported first in February 1996 in India, which was implicated in large localized outbreaks (Chiou et al., 2000; Okuda et al., 1997). This O3:K6 serotype and its serovariant strains formed the pandemic clonal complex, known as CC3 (González-Escalona et al., 2008). During the last two decades, *V. parahaemolyticus* infections and outbreaks caused by this CC3 have increased throughout the world (Nair et al., 2007; González-Escalona et al., 2008).

V. parahaemolyticus is becoming the leading cause of food-borne disease nowadays in southern China. About dozens *V. parahaemolyticus* outbreaks occurred in this area from 2008 to 2014, causing hundreds of cases every year (Ma et al., 2014; Li et al., 2015b). Most of these cases were due to *V. parahaemolyticus* strains having O3:K6 and O1:KUT serotype, both belonging to CC3. However, recently another serotype O4:K8-associated infections increased in southern China (Xiao et al., 2014; Li et al., 2015a, 2015b). Previous study revealed that these O4:K8 isolates affecting southern China formed another important clonal complex called CC345 (Li et al., 2015a, 2015b). However, the molecular characteristics which contributed to the survival and spread of this particular clone were still unclear (Li et al., 2015a, 2015b). Therefore, whole genome sequence-based analysis of these isolates is of utmost importance to elucidate their genetic characteristics, pathogenicity and spread.

Pulsed Field Gel Electrophoresis (PFGE) and Multilocus Sequence Typing (MLST) have been widely used for genotyping *V. parahaemolyticus* isolates and outbreak investigations. However, PFGE

* Corresponding author.
E-mail address: zyh@cdcp.org.cn (Y. Zhang).

http://dx.doi.org/10.1016/j.ijfoodmicro.2017.01.001
0168-1605/© 2017 Elsevier B.V. All rights reserved.

is useful for short-term epidemiology since it does not provide details of the genetic relationships among the strains isolated for a long period, while MLST cannot distinguish the sequence types (STs) of isolates recovered from one area or one period of time and the ones recovered from another area or period of time (Angelo et al., 2015; González-Escalona et al., 2008). Whole genome sequencing (WGS) and WGS-SNPs data analysis possess higher discriminatory ability and sensitivity, to better understand the heredity and variation characteristics of these pathogens, providing new insights into outbreak investigations (Bryant et al., 2012). Thus, in order to make a deep insight into the *V. parahaemolyticus* serotype O4:K8 isolated from southern China, we used WGS to compare the genomes of this serotype involved in the outbreaks in the years 2008–2015. We sequenced 14 strains of serotype O4:K8, then 6 additional *V. parahaemolyticus* CC3 strains (3 O1:KUT and 3 O3:K6) used for phylogenetic comparative analysis. The WGS of additional *V. parahaemolyticus* strains available from international depository organizations were also used.

2. Materials and methods

2.1. Bacterial strains

The 20 *V. parahaemolyticus* strains belonging to O4:K8, O1:KUT or O3:K6 isolated from food-borne diarrheal cases between 2008 and 2015 were used in this study for WGS determination (Table 1). These strains were identified after inoculation onto thiosulfate citrate bile-salts sucrose (TCBS) agar and nutrient agar medium containing 3.5% NaCl at 37 °C for 18 h and serotyped by slide agglutination with a *V. parahaemolyticus* antiserum (Denka Seiken, Tokyo, Japan). The WGS of 31 strains of *V. parahaemolyticus* isolated in China and other countries and belonging to the three serotypes and related serotypes, or with untypeable serovar were available from NCBI data base (Table 2). These WGS data were employed for comparison with those of the strains listed in Table 1.

2.2. Chromosomal DNA preparation

DNA was extracted using QIAamp DNA Mini kit according to the manufacturer's instructions (Qiagen, Inc., Shanghai, China). The DNA quality was checked using a NanoDrop 1000 (Thermo Scientific, Rockford, IL) and the concentration was determined using a Qubit double-stranded DNA-HS assay kit and a Qubit2.0 fluorimeter (Life Technologies, Grand Island, NY), according to the respective manufacturer's instructions.

2.3. Whole genome sequencing, contig assembly and annotation

The genomes of the isolates were sequenced using 125 bp reads using an Illumina HiSeq 2500 sequencer, according to the manufacturer's instructions, at approximately 200× coverage. Genomic sequence contigs were de novo assembled and annotated using the MicrobeTrakr Plus (Zeta Biosciences, Shanghai, China). According to manufacturer, Quake (Kelley et al., 2010) and BWA (Li, 2013) were used in pre- and post-assembly sequences correction, respectively. A MGAP-like (Huntemann et al., 2015) structure annotation method incorporated in MicrobeTrakr was used to predict gene models. Genome sequencing coverage was estimated by bases of raw illumina reads divide by the length of assembly. All the genome assemblies used in this study were generated by high coverage random WGS reads (>200×), it can be inferred that there are no physical gaps within and between scaffolds.

2.4. In silico MLST phylogenetic analysis

The initial analysis and identification of the isolates were performed using an in silico *V. parahaemolyticus* MLST approach, based on the information available at the *V. parahaemolyticus* MLST website (http://pubmlst.org/vparahaemolyticus). Seven *V. parahaemolyticus* loci (*dnaE*, *gyrB*, *recA*, *dtdS*, *pntA*, *pyrC*, and *tnaA*) previously described (González-Escalona et al., 2008) were used for MLST analysis. The same *V. parahaemolyticus* MLST database was also used to assign numbers for alleles and STs.

2.5. Clonal complex assignment

eBURST algorithm was re-implemented in Perl module Graph according to literature (Feil et al., 2004) and integrated in MicrobeTrakr (Zeta Biosciences, Shanghai, China). The most restrictive group definition was used to define the clonal complexes, such as at least six of the seven alleles had to be identical to be included in the same group or clonal complex. The statistical confidences for the ancestral types were assessed using 1000 bootstrap resembling. Two different STs

Table 1
Summary of information on the clinical strains of *V. parahaemolyticus* O4:K8, O3:K6, and O1:KUT isolated in southern China for which the whole genome sequence was determined in this study.

Isolate name	Accession number	MLST								Isolation information				
		Clonal complex	STs	*pyrC*	*recA*	*pntA*	*gyrB*	*dtdS*	*dnaE*	*tnaA*	Serotype	County	Province	Collection year	Source
100135	MORW00000000	CC345	189	48	3	26	48	48	11	26	O4:K8	Guangzhou	Guangdong	2010	Stool
100136	MORX00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Zhaoqing	Guangdong	2008	Stool
100138	MORY00000000	CC3	3	4	19	29	4	4	3	22	O1:KUT	Zhaoqing	Guangdong	2008	Stool
100139	MORM00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Guangzhou	Guangdong	2014	Stool
100142	MORZ00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Dongguan	Guangdong	2010	Stool
100143	MOSA00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Guangzhou	Guangdong	2013	Stool
100144	MOSB00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Guangzhou	Guangdong	2014	Stool
100145	MOSC00000000	CC3	3	4	19	29	4	4	3	22	O3:K6	Guangzhou	Guangdong	2013	Stool
100146	MOSD00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Chaozhou	Guangdong	2009	Stool
100147	MOSE00000000	CC345	189	48	3	26	48	48	11	26	O4:K8	Zhangjiang	Guangdong	2009	Stool
100148	MOSF00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Chaozhou	Guangdong	2009	Stool
100149	MORO00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Guangzhou	Guangdong	2015	Stool
100150	MORN00000000	CC3	3	4	19	29	4	4	3	22	O1:KUT	Zhaoqing	Guangdong	2008	Stool
100151	MORP00000000	CC3	3	4	19	29	4	4	3	22	O1:KUT	Foshan	Guangdong	2010	Stool
100152	MORQ00000000	CC3	3	4	19	29	4	4	3	22	O3:K6	Foshan	Guangdong	2010	Stool
100153	MORR00000000	CC3	3	4	19	29	4	4	3	22	O3:K6	Dongguan	Guangdong	2009	Stool
100154	MORSS00000000	CC345	189	48	3	26	48	48	11	26	O4:K8	Guangzhou	Guangdong	2012	Stool
100156	MORU00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Guangzhou	Guangdong	2013	Stool
100157	MORV00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Dongguan	Guangdong	2009	Stool
100158	MORT00000000	CC345	265	48	107	26	48	48	11	26	O4:K8	Zhuhai	Guangdong	2015	Stool

Table 2
National and international *V. parahaemolyticus* isolates genome used for phylogenomic analyses available at the National Center for Biotechnology Information.

GenBank assembly accession	BioSample	Isolation information			
		Collection Year	Geographic location	Isolation_source	Serovar
GCA_001270885.1	SAMN02781333	1998	Chile: Antofagasta	Feces	O3:K6
GCA_000525005.1	SAMN02641511	2008	Bangladesh	Water	NA
GCA_000786835.1	SAMN03195665	2007	Canada: BC	Feces	NA
GCA_000960685.1	SAMN03349606	2009	Canada: British Columbia	Feces	O3:K6
GCA_000972035.1	SAMN03452294	2009	Canada: British Columbia	Feces	O1:KUT
GCA_000972045.1	SAMN03452290	2007	Canada: British Columbia	Feces	O3:K6
GCA_000972125.1	SAMN03452292	2009	Canada: Alberta	Feces	O1:KUT
GCA_000736325.1	SAMN02910585	2010	China	Shrimp pond	NA
GCA_001541615.1	SAMN04349747	2007	China: Guangxi	Stool	NA
GCA_001541625.1	SAMN04349756	2009	China: Guangxi	Stool	NA
GCA_001013435.1	SAMN03646983	2007	Greece	Environment	NA
GCA_000773455.1	SAMN02866458	2013	India: Andhra Pradesh	*Litopenaeus vannamei*	NA
GCA_001011015.1	SAMN03458146	1951	Japan	Missing	O1:KUT
GCA_001270945.1	SAMN02781340	1996	Japan: Kansai International Airport	Feces	O3:K6
GCA_001558495.1	SAMN03996250	1951	Japan	Shirasu	O1:KUT
GCA_000732995.1	SAMN02912085	2014	Mexico: Hermosillo	Sediment	NA
GCA_000523375.1	SAMN02584906	2013	Mexico: Sinaloa	Stomach	NA
GCA_001270815.1	SAMN02781338	2005	Chile: Puerto Montt	Feces	O3:K6
GCA_001270825.1	SAMN02781339	2007	Chile: Puerto Montt	Feces	O3:K6
GCA_000701045.1	SAMN02252525	1996	Peru	Clinical	O4:K8
GCA_001244315.1	SAMN03159469	2014	South Korea: Gyeongnam	Finespotted flounder	NA
GCA_000736335.1	SAMN02910583	1999	Thailand	Hepatopancreas	NA
GCA_000736345.1	SAMN02910586	2008	Thailand	Shrimp pond sediment	NA
GCA_001433415.1	SAMN03140317	2014	South Korea: Busan	Aquarium water	NA
GCA_000707625.1	SAMN02741350	1997	USA: WA	Clinical	O4:K12
GCA_000707865.1	SAMN02741362	2013	USA: MD	Stool	NA
GCA_000707885.1	SAMN02741351	2004	USA: AK	Missing	O4:K12
GCA_000707605.1	SAMN01816336	1998	USA: LA	Gulf Coast isolate	O4:K9
GCA_000430405.1	SAMN02179882	2007	USA: LA	NA	O1:Kut
GCA_000707585.1	SAMN01816333	2004	USA: AK	Missing	O4:K12
GCA_000737635.1	SAMN02866418	2013	Viet Nam	Stomach	NA

NA, not applicable.

were considered single-locus variant (SLV) when they differed from each other at a single locus. Double-locus variants (DLVs) were referred to any two different STs differing in two loci.

2.6. Phylogenetic analysis

Whole-genome phylogeny of the 20 *V. parahaemolyticus* assemblies was determined by the software package Harvest Suite (version 1.2) using *V. parahaemolyticus* strain BB22OP as reference and enabling the recombination filter (Treangen et al., 2014). The Harvest was run in default configuration, because it was optimized for the data used in this study (closely related bacterial genomes).

2.7. Virulence related gene detection and visualization

Selected virulence factor-coding sequences were extracted from BB22OP genome assembly using the program extractfeat from the European Molecular Biology Open Software Suite (EMBOSS) (Rice et al., 2000). To detect the presence or absence of virulent genes in *V. parahaemolyticus* assemblies, GMAP (Wu and Watanabe, 2005) was used to align all the CDS to the genomes we assembled. Outputs were parsed using an in-house Perl script and visualized using Circos package (Krzywinski et al., 2009) version 0.69.

2.8. Nucleotide sequence accession number

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession MORM00000000-MOSF00000000 (Table 1).

3. Results and discussion

3.1. In silico V. parahaemolyticus strains MLST

In this study, we performed the whole genome sequencing of 14 *V. parahaemolyticus* O4:K8 strains isolated from food-borne diarrheal cases between 2008 and 2015, as well as three O3:K6 and three O1:KUT strains for comparative analysis. Three STs were identified by in silico MLST: ST-3, ST-189 and ST-265. All the O3:K6 and O1:KUT strains belonged to the ST-3, while three O4:K8 strains belonged to ST-189 and another eleven O4:K8 strains belonged to the ST-265. Most of the identified STs varied by location, population and isolated years in southern China (Table 1). Our eBURST analysis (goeBURST) showed that all the O4:K8 isolates formed clonal complexes CC345, while all the O3:K6 and O1:KUT isolates formed clonal complexes CC3. This result indicated that the O4:K8 strains forming another important clonal complex CC345, might be responsible for the increasing O4:K8 associated infections in southern China during the selected period.

3.2. Phylogenomic analysis

The phylogenetic analysis of the O4:K8, O3:K6 and O1:KUT strains in southern China was evaluated using a WGS-SNP analysis, along with another 31 strains of *V. parahaemolyticus* available from NCBI genome database (*www.ncbi.nlm.nih.gov/genome/?term=vibrio parahaemolyticus*) (Table 2). The WGS-SNP analyses provide clearer insights than earlier methods: many of these strains were unable to be distinguished from one another by MLST and PFGE analysis (Haendiges et al., 2015). Our results obtained a high-resolution tree using the WGS-SNP defining each branch. Two clusters of related strains were identified, cluster I (ST189 and ST265) and cluster II (ST3) (Fig. 1). Previous studies revealed the O4:K8 with ST-189 emerged in 1996 and quickly spread throughout

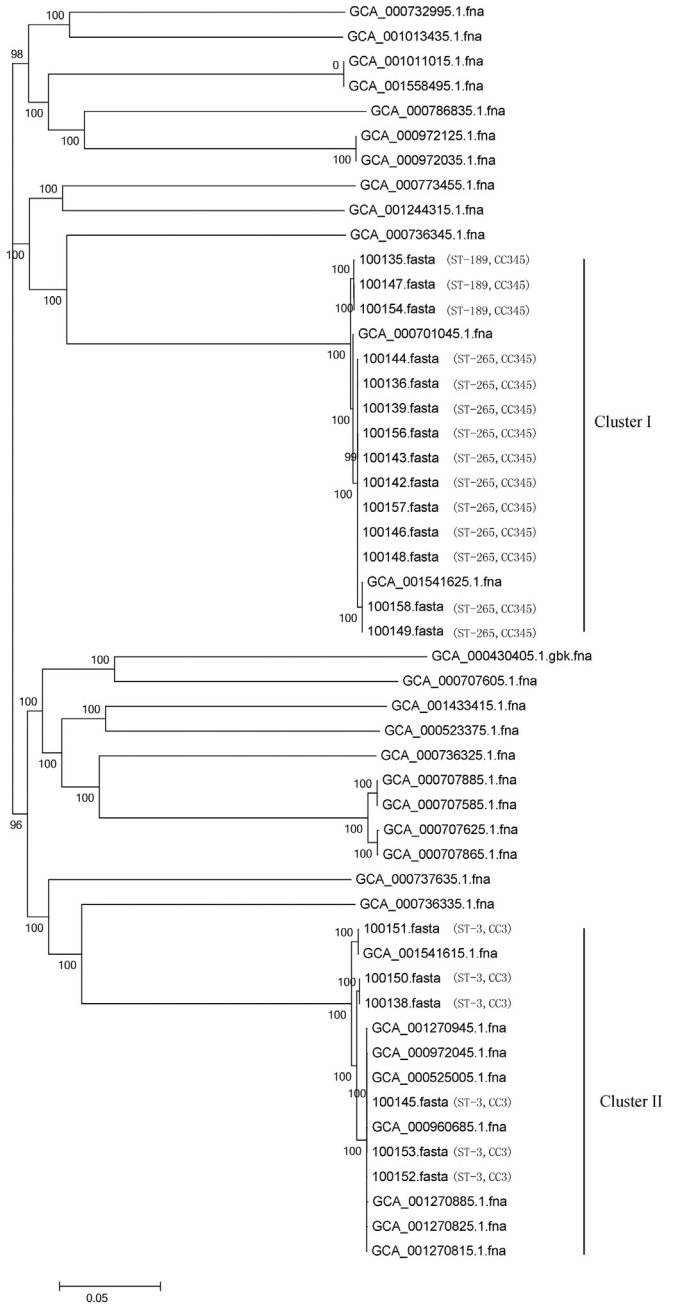

Fig. 1. Phylogenies of lineages of *V. parahaemolyticus* O4:K8, O3:K6 and O1:KUT strains by WGS-SNPs analysis.

the Peru, replacing the O4:K8 with ST-88 variant originated in Asia and becoming the predominant clone (Martinez-Urtaza et al., 2008; González-Escalona et al., 2015). However, the ST-189 as the dominant ST was replaced by the pandemic clone ST-3 in 1997 in Peru, which also originated in Asia (Gavilan et al., 2013). In this study, all the O3:K6 and O1:KUT strains from different countries and regions formed the cluster II. However, all the O4:K8 strains associated with rinsing cases in southern China and this novel genetic variant from Peru reported in 1996 were placed in cluster I, suggesting a possibility of past international transfer of the O4:K8 clone.

3.3. Identification of tdh, trh genes, T3SS and other genomic regions

The thermostable direct hemolysin (TDH) and the TDH-related hemolysin (TRH), encoded by *tdh* and *trh* gene, respectively, are considered as the most important virulent elements of *V. parahaemolyticus* (Shirai et al., 1990). Previous studies had revealed five different subtypes of *tdh* genes (*tdh*1 through *tdh*5) and two recognized subtypes of *trh* genes (*trh*1 and *trh*2) that differ by sequence (Nishibuchi and Kaper, 1990; Kishishita et al., 1992). All of the *V. parahaemolyticus* analyzed in this study carried only the *tdh* genes, but not the *trh* genes (Table 1). The comparison of the TDH amino acid sequences gathered from nucleotide sequences indicated that all of the *tdh* positive strains in this study carried the *tdh* 2 gene.

We performed an in silico analysis of each strain's genome to examine 24 known regions >10 kb described for pandemic strains RIMD2210633, such as the class 1 integron, phages-like, a lipopolysaccharide (LPS)/capsule polysaccharide (CPS) region, two type III secretion systems (T3SS), two osmotic stress response clusters, a type VI secretion system (T6SS) and 7 *Vibrio parahaemolyticus* genomic islands (VPaI-1 to VPaI-7) (Fig. 2). All the O3:K6 and O1:KUT strains carried nearly all the 24 regions except for one O3:K6 which did not carry ORFs of Biofilm region and three O1:KUT strains which did not carry ORFs of the LPS, phage f237 and Biofilm regions. However, all the O4:K8 strains carried the 8 complete regions of NK, T3SS-1, Osmotolerance (chromosome I), Osmotolerance (chromosome II), CPS, Type I secretion, Type I pilus and Ferric uptake, while did not carry ORFs of the LPS, T6SS, phage f237, Integron class-1, Degradative, phage f237-like, Biofilm, Gametolysin and Multidrug efflux regions, indicating that the O4:K8 strains lost a number of genomic fragments which existed completely in the pandemic clone, such as O3:K6 and O1:KUT isolates.

T3SS-1 and T3SS-2 were identified as a type III secretion system (T3SS) on each chromosome of the *V. parahaemolyticus* RIMD2210633, a clinical isolate found in Japan in 1996 (Makino et al., 2003). In addition, this study revealed that T3SS-1 was present in both clinical and environmental isolates with a G + C content similar to the rest of the genome, indicating that this region is ancestral in the species (Makino et al., 2003). T3SS-2 is predominantly present in the *V. parahaemolyticus* O3:K6 strains collected after 1995, indicating that this region is not essential for virulence, although may enhance virulence when present (Makino et al., 2003). All the O4:K8 strains in this study carried the entire T3SS-1 and VpaI-7 (T3SS-2) regions, indicating that the O4:K8 strains have a similar virulence as the O3:K6 pandemic clone.

As regard the T6SS regions (ORFs VP1386 to VP1420), this region encodes T6SS (ORFs VP1401 to VP1409) and a range of proteins that could be translocated by the T6SS, such as the hemo utilization/adhesion proteins, Outer membrane protein A (OmpA), as well as a number of hypothetical proteins (Hurley et al., 2006; Boyd et al., 2008). In this study, all the O4:K8 strains were missing the ORFs VP1389, VP1390 and VP1416 to VP1419. This result suggested that the O4:K8 strains had the complete T6SS, although lacking some hypothetical proteins encoded by this region, which may be the cause of the different virulence mechanism compared to the O3:K6 pandemic clone.

Regarding the pathogenicity islands described for the pandemic strain RIMD2210633 (VPaI 1–7), all the O4:K8 strains in this study carried the entire VPaI-7 region, while they were missing the entire VpaI-1 and VpaI-4 regions (see the supplemental material). The VPaI-1, VpaI-4, VpaI-5 and VpaI-6 regions were considered to be unique among O3:K6 pandemic clone post-1995, indicating that these regions may play a role in the emergence of O3:K6 pandemic clone (Hurley et al., 2006; Boyd et al., 2008). Our analysis revealed that the O4:K8 rising in southern China lacked the VpaI-1 and VpaI-4 regions, while carried partial ORFs of the VpaI-5 and VpaI-6. This molecular characteristic might enable this O4:K8 become a potential dominant clone in southern China.

4. Conclusions

The WGS to provide a high resolution investigation for deep insight into the genetic diversity and molecular characteristics of *V. parahaemolyticus* O4:K8 rising in recent southern China. Comparative

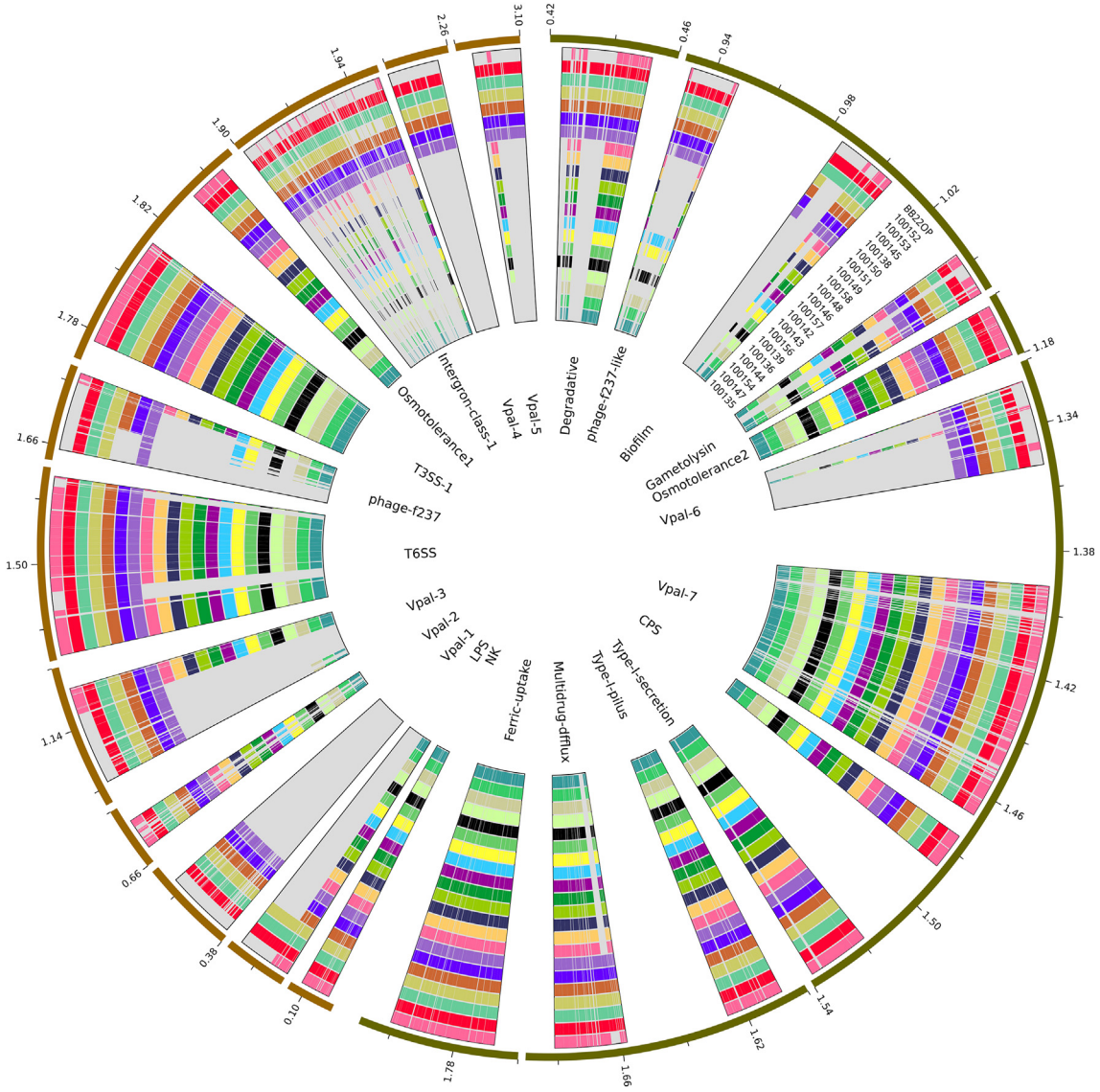

Fig. 2. Genome BLAST Atlas of *V. parahaemolyticus* RIMD2210633 as reference strain (outermost circle) versus *V. parahaemolyticus* BB22OP (O4:K8) and isolates isolated in recent southern China in this study for chromosome 1 and chromosome II. The gaps or holes in the inner circles represent regions present in *V. parahaemolyticus* strain RIMD2210633 that are absent from the other isolates.

genomic analysis revealed the O4:K8 strains carried the entire T3SS-1 and VpaI-7 (T3SS-2) regions and missed the entire or part of the VpaI-1, 2, 3, 4, 5 and 6, which were described as pathogenicity islands and considered to be unique among post-1995 strains belonging to the serotype O3:K6 pandemic clone. Our data showed that the O4:K8 strains possessed the virulence factors similar to the O3:K6 pandemic clone, which may have recently enabling them to become prevalent in southern China. Our study also revealed that WGS-bases analysis may help improve understanding the epidemiology of this bacterium in foodborne disease surveillance.

Disclosure statement

The authors declare no competing financial interests.

Acknowledgments

This work was funded by the Guangdong Province Science and Technology Plan Projects (No. 2014A020219004).

References

Angelo, K.M., Chu, A., Anand, M., Nguyen, T.A., Bottichio, L., Wise, M., Williams, I., Seelman, S., Bell, R., Fatica, M., Lance, S., Baldwin, D., Shannon, K., Lee, H., Trees, E., Strain, E., Gieraltowski, L., Centers for Disease Control and Prevention (CDC), 2015. Outbreak of *Salmonella Newport* infections linked to cucumbers–United States, 2014. MMWR Morb. Mortal. Wkly Rep. 64 (6), 144-7 (20).

Boyd, E.F., Cohen, A.L., Naughton, L.M., Ussery, D.W., Binnewies, T.T., Stine, O.C., Parent, M.A., 2008. Molecular analysis of the emergence of pandemic *Vibrio parahaemolyticus*. BMC Microbiol. 8, 110.

Bryant, J., Chewapreecha, C., Bentley, S.D., 2012. Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. Future Microbiol 7 (11), 1283–1296.

Chen, Y., Stine, O.C., Badger, J.H., Gil, A.I., Nair, G.B., Nishibuchi, M., Fouts, D.E., 2011. Comparative genomic analysis of *Vibrio parahaemolyticus*: serotype conversion and virulence. BMC Genomics 12, 294.

Chiou, C.S., Hsu, S.Y., Chiu, S.I., Wang, T.K., Chao, C.S., 2000. *Vibrio parahaemolyticus* serovar O3:K6 as cause of unusually high incidence of food-borne disease outbreaks in Taiwan from 1996 to 1999. J. Clin. Microbiol. 38 (12), 4621-5.

Feil, E.J., Li, B.C., Aanensen, D.M., Hanage, W.P., Spratt, B.G., 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J. Bacteriol. 186 (5), 1518–1530.

Gavilan, R.G., Zamudio, M.L., Martinez-Urtaza, J., 2013. Molecular epidemiology and genetic variation of pathogenic *Vibrio parahaemolyticus* in Peru. PLoS Negl. Trop. Dis. 7 (5), 16, e2210.

González-Escalona, N., Martinez-Urtaza, J., Romero, J., Espejo, R.T., Jaykus, L.A., DePaola, A., 2008. Determination of molecular phylogenetics of *Vibrio parahaemolyticus* strains by multilocus sequence typing. J. Bacteriol. 190, 2831–2840.

González-Escalona, N., Gavilan, R.G., Brown, E.W., Martinez-Urtaza, J., 2015. Transoceanic spreading of pathogenic strains of *Vibrio parahaemolyticus* with distinctive genetic signatures in the *recA* gene. PLoS One 10 (2), e0117485.

Haendiges, J., Timme, R., Allard, M.W., Myers, R.A., Brown, E.W., González-Escalona, N., 2015. Characterization of *Vibrio parahaemolyticus* clinical strains from Maryland (2012−2013) and comparisons to a locally and globally diverse *V. parahaemolyticus* strains by whole-genome sequence analysis. Front. Microbiol. 6, 125.

Huntemann, M., Ivanova, N.N., Mavromatis, K., Tripp, H.J., Paez-Espino, D., Palaniappan, K., Szeto, E., Pillay, M., Chen, I.M., Pati, A., Nielsen, T., Markowitz, V.M., Kyrpides, N.C., 2015. The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4). Stand. Genomic Sci. 10, 86.

Hurley, C.C., Quirke, A., Reen, F.J., Boyd, E.F., 2006. Four genomic islands that mark post-1995 pandemic *Vibrio parahaemolyticus* isolates. BMC Genomics 7, 104.

Kelley, D.R., Schatz, M.C., Salzberg, S.L., 2010. Quake: quality-aware detection and correction of sequencing errors. Genome Biol. 11 (11), R116.

Kishishita, M., Matsuoka, N., Kumagai, K., Yamasaki, S., Takeda, Y., Nishibuchi, M., 1992. Sequence variation in the thermo stable direct hemolysin-related hemolysin (*trh*) gene of *Vibrio parahaemolyticus*. Appl. Environ. Microbiol. 58, 2449–2457.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., Marra, M.A., 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19 (9), 1639–1645.

Li, H., 2013. Aligning Sequence Reads, Clone Sequences and Assembly Contigs With BWA-MEM (arXiv:1303.3997v2).

Li, B., Luo, J., Tan, H., Ke, B., He, D., Ke, C., Klena, J.D., Zhang, Y., 2015a. Phenotypic and phylogenetic analysis of *Vibrio parahaemolyticus* isolates recovered from diarrhea cases in Guangdong Province, China. Int. J. Food Microbiol. 200, 13-7.

Li, B., Tan, H., Ke, B., He, D., Ke, C., Zhang, Y., 2015b. Etiologic characteristics of food-borne *Vibrio parahaemolyticus* strains isolated in Guangdong, 2014. Zhonghua Liu Xing Bing Xue Za Zhi 36 (11), 1283-7 (Article in Chinese).

Ma, C., Deng, X., Ke, C., He, D., Liang, Z., Li, W., Ke, B., Li, B., Zhang, Y., Ng, L., Cui, Z., 2014. Epidemiology and etiology characteristics of foodborne outbreaks caused by *Vibrio parahaemolyticus* during 2008–2010 in Guangdong province, China. Foodborne Pathog. Dis. 11 (1), 21–29.

Makino, K., Oshima, K., Kurokawa, K., Yokoyama, K., Uda, T., Tagomori, K., Iijima, Y., Najima, M., Nakano, M., Yamashita, A., Kubota, Y., Kimura, S., Yasunaga, T., Honda, T., Shinagawa, H., Hattori, M., Iida, T., 2003. Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholerae*. Lancet 361 (9359), 743–749.

Martinez-Urtaza, J., Huapaya, B., Gavilan, R.G., Blanco-Abad, V., Ansede-Bermejo, J., Cadarso-Suarez, C., Figueiras, A., Trinanes, J., 2008. Emergence of Asiatic *Vibrio* diseases in South America in phase with El Niño. Epidemiology 19 (6), 829–837.

Nair, G.B., Ramamurthy, T., Bhattacharya, S.K., Dutta, B., Takeda, Y., Sack, D.A., 2007. Global dissemination of *Vibrio parahaemolyticus* serotype O3:K6 and its serovariants. Clin. Microbiol. Rev. 20, 39–48.

Nishibuchi, M., Kaper, J.B., 1990. Duplication and variation of the thermo stable direct haemolysin (*tdh*) gene in *Vibrio parahaemolyticus*. Mol. Microbiol. 4, 87–99.

Okuda, J., Ishibashi, M., Hayakawa, E., Nishino, T., Takeda, Y., Mukhopadhyay, A.K., Garg, S., Bhattacharya, S.K., Nair, G.B., Nishibuchi, M., 1997. Emergence of a unique O3:K6 clone of *Vibrio parahaemolyticus* in Calcutta, India, and isolation of strains from the same clonal group from Southeast Asian travelers arriving in Japan. J. Clin. Microbiol. 35 (12), 3150–3155.

Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. 16 (6), 276–277.

Shirai, H., Ito, H., Hirayama, T., Nakamoto, Y., Nakabayashi, N., Kumagai, K., Takeda, Y., Nishibuchi, M., 1990. Molecular epidemiologic evidence for association of thermostable direct hemolysin (TDH) and TDH-related hemolysin of *Vibrio parahaemolyticus* with gastroenteritis. Infect. Immun. 58, 3568–3573.

Treangen, T.J., Ondov, B.D., Koren, S., Phillippy, A.M., 2014. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 15 (11), 524.

Vongxay, K., Pan, Z., Zhang, X., Wang, S., Cheng, S., Mei, L., Xu, C., Fang, W., 2008. Occurrence of pandemic clones of *Vibrio parahaemolyticus* isolates from seafood and clinical samples in a Chinese coastal province. Foodborne Pathog. Dis. 5, 127–134.

Wu, T.D., Watanabe, C.K., 2005. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21 (9), 1859–1875.

Xiao, N., Li, B., Liu, X., Xiao, Y., Huang, J., Ke, B., Tan, H., Ke, C., Yu, S., 2014. Etiologic characteristics of *Vibrio parahaemolyticus* strains causing outbreaks and sporadic cases in Guangdong, 2013. Zhonghua Liu Xing Bing Xue Za Zhi 35 (12), 1379–1383 (Article in Chinese).