Identification of Genetic Bases of *Vibrio fluvialis* Species-Specific Biochemical Pathways and Potential Virulence Factors by Comparative Genomic Analysis

Xin Lu,a,b Weili Liang,a,b Yunduan Wang,a,c Jialiang Xu,a,d Jun Zhu,c Biao Kan b

State Key Laboratory for Infectious Disease Prevention and Control, Department of Diarrheal Diseases, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China; Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou, Zhejiang, China; Department of Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA; School of Food and Chemical Engineering, Beijing Technology and Business University, Beijing, China

**Vibrio fluvialis** is an important food-borne pathogen that causes diarrheal illness and sometimes extraintestinal infections in humans. In this study, we sequenced the genome of a clinical *V. fluvialis* strain and determined its phylogenetic relationships with other *Vibrio* species by comparative genomic analysis. We found that the closest relationship was between *V. fluvialis* and *V. furnissii*, followed by those with *V. cholerae* and *V. mimicus*. Moreover, based on genome comparisons and gene complementation experiments, we revealed genetic mechanisms of the biochemical tests that differentiate *V. fluvialis* from closely related species. Importantly, we identified a variety of genes encoding potential virulence factors, including multiple hemolysins, transcriptional regulators, and environmental survival and adaptation apparatuses, and the type VI secretion system, which is indicative of complex regulatory pathways modulating pathogenesis in this organism. The availability of *V. fluvialis* genome sequences may promote our understanding of pathogenic mechanisms for this emerging pathogen.

*Vibrio fluvialis* is a Gram-negative, oxidase-producing, halophilic bacterium that is normally found in coastal water and seafood. Clinically, *V. fluvialis* has been implicated as a cause of gastroenteritis with diarrhea (1) and even wound infection with primary septicaemia in immunocompromised individuals (2). In an epidemic situation, *V. fluvialis* behaved more aggressively than *Vibrio cholerae* O1, having a higher attack rate and a different clinical picture in the population (3). The gastrointestinal illness caused by this pathogen is usually associated with the consumption of raw or improperly cooked seafood. Additionally, this bacterium causes significant economic loss, like in the lobster industries of the eastern coasts of the United States and Canada (4).

Therefore, *V. fluvialis* has gained importance as an epidemic-causing *Vibrio*, especially in coastal areas.

*V. fluvialis* was broadly defined as a nonagglutinating *Vibrio* species (5). It was first isolated in 1975 from patients with severe diarrhea (6) and was originally called “group F Vibrios” by Furniss et al. and “EF-6 Vibrios” by the Centers for Disease Control (6, 7). Both phenotypic tests and DNA relatedness indicated that these organisms were much closer to the genus *Vibrio* than to the genus *Aeromonas* (8). In some taxonomic studies, group F was separated into two subgroups based on gas production during glucose fermentation (9, 10). The aerogenic group F strains were in a different DNA relatedness group from the anaerogenic strains, and the two biogroups were separated into two species within the genus *Vibrio* (8). The name *V. fluvialis* was proposed for both aerogenic and anaerogenic strains of group F and the synonymous group EF-6. However, only anaerogenic strains have been isolated from human samples, even though both aerogenic and anaerogenic strains of *V. fluvialis* have been found in the environment (11).

Subsequently, the aerogenic strains of group F were confirmed to be a separate species from *V. fluvialis* and were named a new species, *Vibrio furnissii* (12). *V. furnissii* produces gas from β-glucose, while *V. fluvialis* does not.

The clinical symptoms of gastroenteritis caused by *V. fluvialis* are quite similar to those caused by *V. cholerae*, except for the frequent occurrence of blood in stools. In spite of several virulence factors, including reports of enterotoxin-like substances, lipase, proteases, cytotoxins and hemolysin, their definitive contributions to the clinical manifestations remain to be characterized (2).

We recently have identified a quorum-sensing regulatory cascade that regulates several potential virulence genes (13); however, questions regarding the microbiological characteristics, mechanism of pathogenicity, and ecology of the organism remain mostly unanswered.

Whole-genome sequencing has served as a routine tool in bacteriological studies by revealing the basic genetic information of microorganisms, which provides the foundation for further functional studies. To date, seven whole-genome sequences of *Vibrio* spp. (including *V. cholerae*, *V. parahaemolyticus*, *V. splendidus*, *V. vulnificus*, *V. anguillarum*, *V. harveyi*, and *V. furnissii*) have been released on GenBank, in addition to several draft genome sequences of *Vibrio* spp. *V. cholerae*, *V. parahaemolyticus*, and *V. vulnificus* are among the intensively studied and high-risk human-pathogenic species. The whole-genome sequence of *V. furnissii*...
(strain NCTC 11218), which is considered one of the “emerging Vibrio species” potentially implicated in human infection (14), was recently reported as well (15). During our preparation of this article, two draft genome sequences of V. fluvialis were reported (16). In this study, we sequenced and annotated the genome of a clinical strain of V. fluvialis, which was isolated from a patient with diarrhea. The genome sequences outlined herein provide a global view of the genome of V. fluvialis and a clear map of the phylogenetic relationship with other Vibrio spp. The molecular mechanism that underlies the biochemical differences that discriminate V. fluvialis, V. furnissii, and V. cholerae is also demonstrated.

MATERIALS AND METHODS

Strains and genome sequencing. V. fluvialis strain 85003, which was isolated from an adult diarrheal patient in 1985 in Xinjiang Province, China, was used to construct shotgun libraries for genomic sequencing. The genomic DNA of strain 85003 was extracted from its overnight Luria-Bertani broth culture using the TIANamp bacterial DNA kit (Tiangen Biotech, Beijing, China). In addition, V. cholerae O1 El Tor strain N16961 was used in the biochemical tests and the genetic complement experiments.

An initial V. fluvialis shotgun genome sequence was obtained from a mixture of Solexa sequences from the pair-end Solexa sequencing of a 500-bp insert-size library to 160× coverage. The low-quality regions of raw reads were trimmed using the fastq quality trimmer of FASTX-Toolkit (version 0.0.13.2) (http://hannonlab.cshl.edu/fastx_toolkit/index.html), with 20 as the quality threshold. After trimming, reads shorter than 50 bp were removed as low-quality reads. High-quality reads were assembled to contigs with Velvet (version 1.2.07) (17). Contigs were then scaffolded by SPSPACE (18).

Annotation. Protein coding genes were predicted by Glimmer (version 3.02) (19) with the default parameters. Repeats were identified using RepeatMasker (version open-3.2.7) (http://repeatmasker.org). tRNAs were predicted by tRNAscan (version 1.23) (20). rRNAs were predicted by RNamer (version 1.2) (21).

The genes were annotated using the Swiss-Prot, Clusters of Orthologous Groups of proteins (COG), and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. BLAST (22) was used to compare sequences to those in the Swiss-Prot database (downloaded from the European Bioinformatics Institute on 8 August 2012) with E values of ≤1e−10 (23). They were assigned functional annotations via a sequence similarity comparison against the COG database with BLAST with E values of ≤1e−10. A Perl script was written to assign the functional classification to genes (24). Genes were first compared with the KEGG database (release 58) and compared each other subsequently. ClustalW was used to independently align the sequence members of each ortholog cluster to minimize gene rearrangement within the multiple-sequence alignment (27). A maximum likelihood method was used to generate a phylogenetic species tree with 100 replicates for bootstrapping (28) using phyML 3.0 and MEGAS (29). All three produced trees were of highly similar topologies.

Comparative genome analysis. The assembled genome was compared to that of V. furnissii NCTC 11218 (15) using CGView (Circular Genome Viewer) (30) at both the nucleotide and protein levels.

Biochemical pathway complement tests. In the biochemical tests, D-glucose fermentation and three amino acids utilizations (arginine, lysine, and ornithine) were used to discriminate V. fluvialis from V. furnissii and V. fluvialis from V. cholerae, respectively. The genes associated with these carbohydrate fermentation pathways were screened and listed according to the KEGG database (http://www.genome.jp/kegg/). All of these genes were correspondingly aligned to those of V. furnissii and V. cholerae with BLAST with an overlap of less than 80%. The absent genes were then identified in the species with negative phenotypes.

To investigate whether the lack of these homologs in the pathways is responsible for the negative fermentations, the genes were amplified from the genomes of the species with positive phenotypes, cloned into the plasmid vectors, confirmed by sequencing, and transformed into the negative strains by electroporation (31). The 500-bp upstream fragments of the cloned genes (operons) were included. The primers, restriction sites for cloning, and recipient strains are listed in Table S2 in the supplemental material.

In the biochemical tests, the wild-type strains and the complemented strains were tested using a microbial biochemical identification tube (073260; HuanKai, China) for gas from glucose, API20E (bioMérieux, France) for the lysine decarboxylase test and ornithine decarboxylase test, and microbial biochemical identification tubes (073130 and 073140; HuanKai, China) for the arginine dihydrolase test. All procedures were performed following the manufacturer’s instructions.

Nucleotide sequence accession number. The genome sequences of V. fluvialis reported in the present study have been deposited in the Sequence Read Archive under accession no. SRX397301.

RESULTS AND DISCUSSION

General genome features of V. fluvialis strain 85003. A total of 5,135,920 reads (770,388,000 nucleotide bases) were assembled into 132 large contigs. The contig N50 was 193,015 bp in length, and the longest assembled contig was 532,968 bp. The draft genome of V. fluvialis 85003 consists of 4,689,093 bp with 50.0% GC content and contains 4,596 predicted coding sequences (CDSs) (Table 1), just the same as the genome characteristics of V. fluvialis strains PG41 and 121563 (16). The putative functions could be assigned to 60.7% of the CDSs, while 14.4% possessed similarity to genes of unknown function, and no function could be proposed for 25.0% of the CDSs. A total of 53 tRNA operons are predicted in the draft genome. The assembled contigs from the newly sequenced V. fluvialis 85003 strain are shown aligned to the V. furnissii NCTC 11218 genome in Fig. 1. The general features of the genome are summarized in Table 1 and Fig. 1.

Genome comparison revealed the phylogenetic position of V. fluvialis in Vibrio. We compared the genome sequences of 11 Vibrio spp. and two Photobacterium spp. from GenBank with the
identified 961 common single-copy ortholog clusters. The patterns of the phylogenetic tree showed that all Vibrio spp. clustered together and formed a phylogroup separate from Photobacterium spp. and indicated the taxonomic position of V. fluvialis among the Vibrionaceae (Fig. 2), which were similar to those observed in the 16S rRNA gene gene-based tree. The taxonomic position was also similar to those recorded in Bergey’s Manual of Systematic Bacteriology, 2nd ed. (32). As expected, V. fluvialis 58003 and V. furnissii NCTC 11218 clustered together, which suggested that both species may have evolved recently from a common ancestor. Both species formed a sister taxon with V. cholerae IEC224 and Vibrio mimicus MB451.

Within the phylogenetic tree, V. cholerae, V. mimicus, V. fluvialis, and V. furnissii were distributed into the Cholerae clade, which was recognized as the nearby branch of the Anguillarum clade (33). V. parahaemolyticus, V. campbellii, and V. harveyi were recognized as members of the Harveyi clade; they formed a distinct sublineage in our genomic analysis (Fig. 2). V. ordalii was previously recognized as biotype 2 of V. anguillarum; the genome-based phylogenic tree showed that they were in the same lineage (Fig. 2). The phylogenetic analysis showed the relationships among Vibrio species and confirmed the taxonomy of these species, which will be useful in an evolution context.

We further calculated the core genome of V. fluvialis and other Vibrio spp. The pairwise comparisons of core genes between V. fluvialis 85003 and the other 10 Vibrionaceae species used in the above phylogenetic analysis showed that V. fluvialis 85003 shared the most core genes (3,754 genes) with V. furnissii NCTC 11218 and less than 2,810 genes with the rest of the Vibrionaceae species. V. fluvialis 85003 shared the fewest core genes (2,388 genes) with V. ordalii ATCC 33509 (Fig. 3A) because V. ordalii ATCC 33509 has the smallest genome of these Vibrionaceae species. Figure 3B

### TABLE 1 Global genomic view of V. fluvialis 85003 and V. furnissii NCTC 11218

<table>
<thead>
<tr>
<th>Parameter</th>
<th>V. fluvialis</th>
<th>V. furnissii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of sequence, bp</td>
<td>4,689,093</td>
<td>4,916,408</td>
</tr>
<tr>
<td>GC content, %</td>
<td>50.03</td>
<td>50.67</td>
</tr>
<tr>
<td>Repeat content length, bp (% coverage)</td>
<td>5,608 (0.12)</td>
<td>19,104 (0.39)</td>
</tr>
<tr>
<td>Simple repeat length, bp (% coverage)</td>
<td>557 (0.01)</td>
<td>1,067 (0.02)</td>
</tr>
<tr>
<td>Low-complexity length, bp (% coverage)</td>
<td>1,438 (0.03)</td>
<td>1,881 (0.04)</td>
</tr>
<tr>
<td>Repeat element length, bp (% coverage)</td>
<td>504 (0.01)</td>
<td>1,131 (0.02)</td>
</tr>
<tr>
<td>Small RNA no. (bp length (% coverage))</td>
<td>27 (3.292 [0.07])</td>
<td>73 (15,537 [0.32])</td>
</tr>
</tbody>
</table>

SRR length in genome, bp^a^ 4,016,205 4,220,358

**Features**
- ORFs, no. 4,596 4,455
- Genome coverage, % 85.65 85.84
- Genome density, genes/kb 0.98 0.91
- Avg gene length, bp 874 947
- Maximum gene length, bp 4,596 9,450

**Functions**
- ORFs with assigned functions, no. b 2,788 2,928
- Function prediction only or unknown^a^ 660 633
- ORFs without COG match, no. 1,148 894

**RNA**
- rRNA^c^ 5S rRNA, no. (length, bp) 2 (115) 8 (115)
- 16S rRNA, no. (length, bp) 1 (1,541) 7 (1,541)
- 23S rRNA, no. (length, bp) 1 (2,906) 7 (2,889)
- tRNA, no. (total bp length)^d^ 53 (4,205) 95 (7,474)

---

^a^ Coding genes were predicted by Glimmer version 3.02.
^b^ BLAST against the COG database with an E value of < 1e−10.
^c^ rRNA genes were predicted by RNAmmer version 1.2.
^d^ tRNA genes were predicted by tRNAscan version 1.23.
showed the profile curves for the core gene number. **Vibrionaceae** species strains were added successively into the analysis by the distance between them and **V. fluvialis** in the phylogenetic tree. The distributions of the clusters of orthologous groups (COG) in the core genes were examined. Chi-square tests showed that the numbers of core genes in the same COG groups did not significantly differ between **V. fluvialis** and **V. furnissii**, which further demonstrated their close relationship. When **V. fluvialis** was pairwise compared with the other nine **Vibrionaceae** species, the significant differences in gene number were found mainly in three COG groups: (i) transcription; (ii) translation, ribosomal structure, and biogenesis; and (iii) signal transduction mechanisms.

**Genome comparison of V. fluvialis and V. furnissii.** In this study, we found that the mean BLASTp identity of the homologous genes between **V. fluvialis** 85003 and **V. furnissii** NCTC 11218 was 86.8%. The complete genome sequence of the **V. furnissii** sp. nov. strain NCTC 11218 was reported (15). It consists of two circular chromosomes (3.2 Mb and 1.6 Mb). The GC content of chromosome I is 50.7%, and that of chromosome II is 50.5%. Chromosome I contains 3,013 open reading frames (ORFs) and 95 tRNA sequences, and chromosome II has 1,448 ORFs and 5 tRNA sequences. Using **V. furnissii** as the reference, the **V. fluvialis** 85003 genome contains 141 more genes and shares 3,948 homologous genes (see Fig. S2 in the supplemental material). The GC content of the **V. fluvialis** 85003 genome is 50.0%. Approximately 25.0% (1,148/4,596) of the chromosomal ORFs lacked a COG match in nature, and these genes account for the majority of genes that are unique to the **V. fluvialis** 85003 genome. For the remaining functionally known genes, we examined the distributions in different COG groups (see Table S1 and Fig. S1 in the supplemental material). Amino acid transport and metabolism and energy production and conversion accounted for 8.9% (308/3,448) and 6.4% (219/3,448), respectively, of the genes in this major COG group. Compared with **V. furnissii**, carbohydrate transport and
metabolism and cell motility accounted for a small fraction of the genes in *V. fluvialis* 85003. COG function classification identified individual genes involved in chromatin structure and dynamics, as well as RNA processing.

**Genetic bases of metabolic analysis to differentiate** *V. fluvialis* **from other closely related** *Vibrio* **spp.** (i) *V. fluvialis* and *V. furnissii*. Because they are two closely related *Vibrio* species, the biochemical phenotype of *V. fluvialis* is very similar to that of *V. furnissii*, except for the gas production during the D-glucose fermentation test. *V. furnissii* produces gas from the fermentation of D-glucose, while *V. fluvialis* does not (12). The gas production from glucose is associated with activity of formate dehydrogenase (34). The pathway comparison between *V. fluvialis* 85003 and *V. furnissii* NCTC 11218 revealed that the vfu_A02711 gene of *V. furnissii* NCTC 11218, annotated as producing an NADP-dependent formate dehydrogenase alpha subunit, was absent in *V. fluvialis* strain 85003 (Table 2 and Fig. 4A). Formate dehydrogenase (NADP⁺) (EC 1.2.1.43) catalyzes the chemical reaction with for-

<table>
<thead>
<tr>
<th>Biochemical tests</th>
<th>Gene(s)</th>
<th>Strain species</th>
<th>Test result for:</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Glucose</td>
<td>vfu_A02711</td>
<td><em>V. furnissii</em></td>
<td><em>V. fluvialis</em></td>
</tr>
<tr>
<td>Lysine</td>
<td>cadAB operon</td>
<td><em>V. cholerae</em></td>
<td><em>V. fluvialis</em></td>
</tr>
<tr>
<td></td>
<td>VC0278</td>
<td><em>V. cholerae</em></td>
<td><em>V. fluvialis</em></td>
</tr>
<tr>
<td>Ornithine</td>
<td>speF-potE operon</td>
<td><em>V. cholerae</em></td>
<td><em>V. fluvialis</em></td>
</tr>
<tr>
<td>Arginine</td>
<td>arcABCD operon</td>
<td><em>V. fluvialis</em></td>
<td><em>V. cholerae</em></td>
</tr>
<tr>
<td></td>
<td>argR</td>
<td><em>V. fluvialis</em></td>
<td><em>V. cholerae</em></td>
</tr>
</tbody>
</table>

**TABLE 2** Identification of genes responsible for amino acid metabolism whose absence in the recipient strain results in a negative phenotype

![Genome structure of the absent genes identified in the species with negative phenotypes. Panels A to D show the absent genes identified that are associated with the D-glucose fermentation pathway (A), the lysine utilization pathway (B), the ornithine utilization pathway (C), and the arginine utilization pathway (D).](http://aem.asm.org/)
mate and NADP\(^+\) as the substrates and CO\(_2\) and NADPH as products. We introduced vfu_A02711 of *V. furnissii* via a recombinant plasmid into *V. fluvialis* strain 85003 and reproduced the gas production from glucose of *V. furnissii* from glucose (Fig. 5A). The results experimentally demonstrated that the deletion of the NADP-dependent formate dehydrogenase alpha subunit gene accounts for the anaerogenic phenotype of glucose fermentation in *V. fluvialis*.

(ii) *V. fluvialis* and *V. cholerae*. *V. cholerae* is the second closest species to *V. fluvialis*. Species-specific minimal biochemical tests are used to identify *V. fluvialis*; without these tests, *V. fluvialis* may be confused with *V. cholerae* (5). Utilizations of arginine, lysine, and ornithine are the major biochemical tests used to differentiate the species *V. fluvialis* and *V. cholerae*. *V. cholerae* is \(\sim + +\) for utilization of these three amino acids, respectively, but *V. fluvialis* is \(+ --\). We searched for enzymes and genes involved in their metabolism in the genomes of both species. Lysine is used by the enzyme lysine decarboxylase, whose genes are carried by the operon containing VC0281 (cadC) and VC0280 (cadB) in *V. cholerae* N16961. The transcription of this operon is controlled by the upstream membrane-bound transcription activator CadC, which is encoded by VC0278 (35). These three genes are absent in the *V. fluvialis* 85003 genome (Table 2 and Fig. 4B). Introduction of cadBA of *V. cholerae* into *V. fluvialis* 85003 with a plasmid still reproduced a negative result for the lysine decarboxylase test, whereas introduction of both cadBA and cadC of *V. cholerae* into *V. fluvialis* 85003 yielded a positive result, as for *V. cholerae* (Fig. 5B). Therefore, a lack of the three genes results in a defect in the catabolism of lysine by *V. fluvialis*.

Ornithine decarboxylase catalyzes ornithine to form putrescine. The speF-potE operon encoded ornithine decarboxylase and the polypeptide transport PotE protein, which are induced at an acidic pH (36). VCA1063 and VCA1062 encode SpeF and PotE in *V. cholerae* but were not found in the *V. fluvialis* 85003 genome (Table 2 and Fig. 4C). Introduction of the intact speF-potE operon of *V. cholerae* into *V. fluvialis* strain 85003 reproduced the positive result (Fig. 5C).

The arginine deiminase pathway enables *V. fluvialis* to grow anaerobically on arginine. The arc genes appear clustered in an operon-like structure, including arcA (arginine deiminase), arcB (ornithine carbamoyltransferase), arcC (carbamate kinase), and arcD (putative arginine-ornithine antiporter). arcC and arcD were absent in the *V. cholerae* N16961 genome (Table 2 and Fig. 4D). Introduction of the arcCD of *V. fluvialis* into the *V. cholerae* N16961 strain did not reproduce the positive result for the arginine dihydrolase test, and introduction of the whole arcABCD operon into the N16961 strain reproduced a positive but time-delayed result. ArgR is an essential local transcriptional regulator of the arc operon (37). The argR in strain N16961 only shares 41% of its identity with *V. fluvialis*. We then introduced both the argR gene and arcABCD operon of *V. fluvialis* into *V. cholerae* strain N16961; the transformed strain showed a positive result over the time frame examined (Fig. 5D).

Taken together, the above data indicate that the differences in the hexose and amino acid metabolic tests used to differentiate *V. fluvialis* from *V. furnissii* and *V. cholerae* are due to an absence of the genes that are involved in the metabolic pathways. The gene deletions may endow these species with different phenotypes. From the genome comparisons and the complement experiments, the genetic mechanisms of the metabolic differences of these species are revealed. These data also present the genetic evidence to study the differentiation and evolution of these closely related species within *Vibrio*. Based on the gene differences, nucleic acid-based tests could be developed to replace the biochemical tests used to identify these similar species.

Virulence-related genes identified in the genome of *V. fluvialis* 85003. (i) Hemolysin. *V. fluvialis* is an enteric pathogen; however, its pathogenesis has not been assigned to one or several virulence factors, such as the clear role of cholera toxin in the pathogenesis of *V. cholerae*. *V. fluvialis* reportedly produces an extracellular hemolysin (38). In this study, several hemolysin-related genes, which showed high similarity to the genes identified
TABLE 3 Homologs of T6SS-associated genes identified in the *V. fluvialis* 85003 isolate and BLASTp identity results

<table>
<thead>
<tr>
<th><em>V. fluvialis</em> gene</th>
<th>Homolog in other <em>Vibrio</em> spp.</th>
<th>Product by BLASTx KEGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>VFL_001426</td>
<td>vfu_B00795 (82)</td>
<td>ImpA</td>
</tr>
<tr>
<td>VFL_001427</td>
<td>vfu_B00794 (100)</td>
<td>Hcp</td>
</tr>
<tr>
<td>VFL_001428</td>
<td>vfu_B00793 (99)</td>
<td>VCA0107 (35) ImpB</td>
</tr>
<tr>
<td>VFL_001429</td>
<td>vfu_B00792 (97)</td>
<td>VCA0108 (34) ImpC</td>
</tr>
<tr>
<td>VFL_001430</td>
<td>vfu_B00791 (98)</td>
<td>VCA0108 (25) ImpD</td>
</tr>
<tr>
<td>VFL_001431</td>
<td>vfu_B00790 (100)</td>
<td>ImpF</td>
</tr>
<tr>
<td>VFL_001432</td>
<td>vfu_B00789 (98)</td>
<td>VCA0110 (24) ImpG</td>
</tr>
<tr>
<td>VFL_001433</td>
<td>vfu_B00788 (99)</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>VFL_001434</td>
<td>vfu_B00786 (99)</td>
<td>VCA0116 (42) VsaG</td>
</tr>
<tr>
<td>VFL_001435</td>
<td>vfu_B00785 (94)</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>VFL_001436</td>
<td>vfu_B00784 (92)</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>VFL_001437</td>
<td>vfu_B00783 (93)</td>
<td>VsaD</td>
</tr>
<tr>
<td>VFL_001438</td>
<td>vfu_B00782 (99)</td>
<td>VCA0114 (27) ImpJ</td>
</tr>
<tr>
<td>VFL_001439</td>
<td>vfu_B00781 (81)</td>
<td>ImpK</td>
</tr>
<tr>
<td>VFL_001440</td>
<td>vfu_B00780 (90)</td>
<td>VCA0120 (20) ImpL</td>
</tr>
<tr>
<td>VFL_002824</td>
<td>vfu_B01101 (89)</td>
<td>VCA0018 (72) VgrG</td>
</tr>
<tr>
<td>VFL_002825</td>
<td>vfu_B01957 (98)</td>
<td>VCA0018 (75) VgrG</td>
</tr>
<tr>
<td>VFL_002826</td>
<td>vfu_B01190 (93)</td>
<td>VCA0121 (56) Vsl</td>
</tr>
<tr>
<td>VFL_002827</td>
<td>vfu_B01189 (98)</td>
<td>VCA0120 (84) ImpL</td>
</tr>
<tr>
<td>VFL_002828</td>
<td>vfu_B01188 (90)</td>
<td>VCA0119 (73) Vsi</td>
</tr>
<tr>
<td>VFL_002829</td>
<td>vfu_B01187 (86)</td>
<td>VCA0118 (61) VsaI</td>
</tr>
<tr>
<td>VFL_002830</td>
<td>vfu_B01186 (95)</td>
<td>VCA0117 (75) or3-dependent transcriptional regulator</td>
</tr>
<tr>
<td>VFL_002831</td>
<td>vfu_B01185 (93)</td>
<td>VCA0116 (86) VsaG</td>
</tr>
<tr>
<td>VFL_002832</td>
<td>vfu_B01184 (94)</td>
<td>VCA0115 (85) ImpK</td>
</tr>
<tr>
<td>VFL_002833</td>
<td>vfu_B01183 (98)</td>
<td>VCA0114 (90) ImpJ</td>
</tr>
<tr>
<td>VFL_002834</td>
<td>vfu_B01182 (91)</td>
<td>VCA0113 (84) VsaD</td>
</tr>
<tr>
<td>VFL_002835</td>
<td>vfu_B01181 (82)</td>
<td>VCA0112 (64) ImpL</td>
</tr>
<tr>
<td>VFL_002836</td>
<td>vfu_B01180 (97)</td>
<td>VCA0111 (81) ImpH</td>
</tr>
<tr>
<td>VFL_002837</td>
<td>vfu_B01179 (96)</td>
<td>VCA0110 (84) ImpG</td>
</tr>
<tr>
<td>VFL_002838</td>
<td>vfu_B01178 (100)</td>
<td>VCA0109 (89) Hypothetical protein</td>
</tr>
<tr>
<td>VFL_002839</td>
<td>vfu_B01177 (98)</td>
<td>VCA0108 (92) ImpC</td>
</tr>
<tr>
<td>VFL_002840</td>
<td>vfu_B01176 (89)</td>
<td>VCA0107 (87) ImpB</td>
</tr>
</tbody>
</table>

in other *Vibrio* species, were found in the genome of *V. fluvialis* 85003. The gene VFL_002885 of *V. fluvialis* 85003 shares 98% of its identity with that of *V. furnissii* and 78% with the *V. cholerae* El Tor hemolysin A (*hlyA*) gene. Genes VFL_003095 and VFL_003869 are also predicted to be hemolysin genes. VFL_000876 is 91% identical to the thermostable hemolysin gene of *V. parahaemolyticus*. Four other hemolysin genes were identified in the genome of the *V. fluvialis* 85003 strain (Table 3). VasH is the global regulator of the *V. cholerae* T6SS, which is essential to initiate the transcription of T6SS genes and has a universal role in endemic and pandemic *V. cholerae* strains. Genes VFL_002830 was also annotated for its homolog with vasF. T6SS is implicated in interbacterial relationships, biofilm formation, cytotoxicity, and survival in the phagocytic cells of bacteria. These T6SS homologs may serve similar roles in the virulence of *V. fluvialis*.

**Other virulence-related genes.** Based on the virulence factor database (VFDB; http://www.mgc.ac.cn/VFs/), the genes of other virulence factors were annotated in the *V. fluvialis* 85003 genome, including the genes coding for the ferric uptake regulator (Fur), heat shock protein 90 (HtpB), flagellar biosynthetic proteins (FlhA, FlhC, FlhD, FlIE, FlF, FlI, FlIh, FlLm, FlN, FlP, Flq, Flr, and Fls), toxin-coregulated pilus biosynthesis protein 1 (Tcp1), and accessory colonization factor (AcfB). In addition to fur, other genes essential for iron transport were also predicted, including the *fhp* family for ferric iron, the *flu* family for ferrichrome transport, and the *feel* family for ferrous iron. All of these genes are homologues to those of *V. cholerae*. The mannose-sensitive hemagglutinin (MSHA) pilus, which belongs to the family of type IV pili, is directly associated with the adherence of *V. cholerae* to environmental surfaces and biofilm formation. A homologous gene was also annotated in the *V. fluvialis* 85003 genome, VFL_003720 to 003740 were found similar to the MSHA synthesis genes of *V. cholerae*, which suggests the presence of an MSHA-like pilus with a similar role in *V. fluvialis*.

**Conclusion.** In this study, we sequenced the *V. fluvialis* genome and determined its relationships within *Vibrio* spp. The data provide details of the genomic sequences to show the close relationship between *V. fluvialis* and *V. furnissii* and their phylogenetic position in *Vibrio*, which is similar to the relationship revealed by traditional microbiological strategies. The combination of whole-genome sequencing and comparison also provides a rapid and comprehensive approach to identify the mechanisms of the biochemical phenotyping differences, which are used to characterize these closely related species. The approach revealed that gene deletion contributed to their phenotype diversity. From genome sequence searching, multiple virulence-related genes and gene clusters were found, demonstrating the complicated pathogenesis of this enteric pathogen. Therefore, the genome data promote the understanding of the genetic basis of physiology, virulence, evolution, and the environmental survival mechanisms for *V. fluvialis*. Furthermore, for environmental and disease surveillance purposes, the gene markers and even metabolic pathways could be searched for through genome comparisons and used to develop specific detection methods for this marine microbe.

**ACKNOWLEDGMENTS**

We thank Karl-Gustav Rueggeberg for critically reviewing the manuscript.

This work was supported by the Priority Project on Infectious Disease Control and Prevention (2012ZX10004215 and 2008ZX10004-008) of the Ministry of Health, China, and by the National Basic Research Priorities
REFERENCES


Program (2009CB522604) of the Ministry of Scientific Technology, China.


