



# Identification of *Vibrio parahaemolyticus* Pandemic Group-Specific DNA Sequence by Genomic Subtraction

Okura, Masatoshi

Osawa, Ro

Arakawa, Eiji

Terajima, Jun

Watanabe, Haruo

---

(Citation)

Journal of Clinical Microbiology, 43(7):3533-3536

(Issue Date)

2005-07

(Resource Type)

journal article

(Version)

Version of Record

(URL)

<https://hdl.handle.net/20.500.14094/90000323>



## Identification of *Vibrio parahaemolyticus* Pandemic Group-Specific DNA Sequence by Genomic Subtraction

Masatoshi Okura,<sup>1</sup> Ro Osawa,<sup>1\*</sup> Eiji Arakawa,<sup>2</sup> Jun Terajima,<sup>2</sup> and Haruo Watanabe<sup>2</sup>

Department of Bioscience, Graduate School of Science and Technology, Kobe University, Rokko-dai 1-1, Nada-ku, Kobe, 657-8501, Japan,<sup>1</sup> and Department of Bacteriology, National Institute of Infectious Diseases, Toyama, Shinjuku-ku, Tokyo, 162-8640, Japan<sup>2</sup>

Received 6 January 2005/Returned for modification 18 February 2005/Accepted 28 February 2005

**A genomic subtraction between a pandemic *Vibrio parahaemolyticus* and a nonpandemic strain that seemed to be clonally related was performed. A subtractive DNA fragment was identified to be a part of a 16-kbp insertion sequence which was present in almost all pandemic strains but not in nonpandemic strains tested.**

*Vibrio parahaemolyticus* is a major cause of seafood-borne gastroenteritis, frequently associated with the consumption of raw or undercooked seafood (2). Although various serovars of the bacterium can cause infections, O3:K6, O4:K68, and several other serovars producing thermostable direct hemolysin (TDH) have been recognized as the predominant group responsible for most outbreaks occurring worldwide since 1996 (6, 8). Past molecular studies (8, 13) based on pulsed-field gel electrophoresis (PFGE) and arbitrarily primed-PCR (AP-PCR) indicated that those pandemic strains showed almost identical fragment patterns, suggesting them to be clonally related, thus forming the “pandemic group.” Nevertheless, we still do not have any substantial evidence to explain the “pandemicity” of these strains.

In this context, Matsumoto et al. (6) reported that the pandemic group exhibited a unique sequence (ToxR/new) within the *toxRS* operon, which encodes transmembrane proteins involved in the regulation of virulence-associated genes conserved in the genus *Vibrio*. Meanwhile, we have recently identified yet another pandemic group-specific sequence which is located in a gene encoding a hypothetical protein approximately 80% homologous to the Mn<sup>2+</sup> and Fe<sup>2+</sup> transporter of the natural resistance-associated macrophage protein family of *Vibrio vulnificus* (10). However, the above examples were limited to a few base replacements from those of nonpandemic strains and thus did not seem to lead to any definite pandemic trait. In the present study, we therefore aimed at identifying pandemic group-specific DNA sequences on a much larger scale, using genomic subtraction, in which whole genome preparations of a pandemic strain (tester strain) are subtracted from that of a nonpandemic strain (driver strain) to leave any group-specific sequences to be identified.

In the present study, a pandemic strain, NIID K7, isolated in 1998 from a Japanese patient and carrying the TDH gene, ToxR/new, and pandemic group-specific sequence was the tester strain. A nonpandemic strain, KE10491, isolated from a Japanese patient in 1983 which possessed ToxR/new without the TDH gene and pandemic group-specific sequence was the

driver strain (Table 1). As described previously (9), the strains resemble each other in their PFGE and AP-PCR profiles (Fig. 1), suggesting that they are clonally related. It was thus hoped that we could identify the unique sequences by genomic subtraction more efficiently than by using remotely related strains.

For genomic subtraction, representational difference analysis was performed, following essentially the procedure described by Calia et al. (3) with use of a commercially available kit (GeneFisher PCR Subtraction System; Takara Shuzo, Tokyo, Japan). Briefly, 1 microgram of each of the DNAs was digested with 10 units of the enzyme BglII. The digests were then extracted with phenol-chloroform, precipitated, and resuspended at a concentration of 0.2 µg/µl. Three sets of complementary oligonucleotide pairs, designated *RBgl*, *JBgl*, and *NBgl*, were provided with the above kit; each of these pairs consists of a 24-mer and a 12-mer which, when annealed, generate a BglII-compatible overhang.

For representation, the pair of *RBgl* oligonucleotides were ligated onto the ends of BglII-digested DNAs. After an initial extension reaction at 72°C for 5 min to fill in the complementary strand to the ligated 24-mer, PCR was used to amplify these fragments with the 24-mer oligonucleotide as the primer. Following amplification, the *RBgl* oligonucleotides were digested off of the amplified fragments with BglII and the fragments from the tester strain were gel purified. The second set of *JBgl* was then ligated onto the purified fragments of the tester DNA as above.

The hybridization reaction was set up using 40 µg of driver DNA and 0.4 µg of tester DNA (100:1 driver/tester ratio) in a final volume of 8 µl of hybridization buffer. The DNA was denatured at 98°C for 5 min. The hybridization mix was then amplified with the *JBgl* 24-mer as the PCR primer. The first round of amplification was for 10 cycles of 1 min at 95°C and 3 min at 70°C, with the last cycle followed by an extension at 72°C for 10 min, followed by digestion of the products by mung bean nuclease. The nuclease-treated product was then amplified for an additional 20 cycles. The resulting amplicons, which are called the first difference products, were digested off with BglII, and the fragments were gel purified as above.

The third pair of oligonucleotides, *NBgl*, was ligated onto the pool of amplified difference product fragments, and this pool was mixed with a large excess of driver DNA fragments.

\* Corresponding author. Mailing address: Department of Bioscience, Graduate School of Science and Technology, Kobe University, Rokko-dai 1-1, Nada-ku, Kobe City, Japan 657-8501. Phone: 81-78-803-5804. Fax: 81-78-803-5804. E-mail: osawa@ans.kobe-u.ac.jp.

TABLE 1. List of pandemic and nonpandemic strains used for the first-screen PCR tests targeting the genomic subtraction-derived gene sequences

Strain	Serovar	Yr of isolation	Country of isolation	Source	Presence of gene encoding:				Subtraction-derived gene sequence <sup>c</sup>		
					TDH <sup>a</sup>	TRH <sup>a</sup>	ToxRS/new <sup>a</sup>	PGSS <sup>b</sup>	VP2905	VPA1367	VPA1371
Pandemic strains											
NIIDK7 (tester)	O3:K6	1998	Japan	Human	+	—	+	+	+	+	+
KE10495	O3:K6	1996	Japan	Human	+	—	+	+	+	+	+
AN-2416	O3:K6	1998	Bangladesh	Human	+	—	+	+	+	+	+
AN-8373	O3:K6	1998	Bangladesh	Human	+	—	+	+	+	+	+
AN-2189	O4:K68	1998	Bangladesh	Human	+	—	+	+	+	+	+
NIID242-200	O4:K68	2000	Korea	Human	+	—	+	+	+	+	+
Nonpandemic strains											
KE10491 (driver)	O3:K6	1983	Japan	Human	—	—	+	—	—	—	—
KE10465	O3:K6	1985	Japan	Human	—	—	+	—	—	—	—
KE9967	O3:K6	1981	Japan	Human	+	—	—	—	—	+	+
KE9984	O3:K6	1981	Japan	Human	—	+	—	—	—	—	+
KE10538	O4:K8	1999	Thailand	Human	+	—	—	—	—	+	+
KE10579	O1:K1	2000	Japan	Sea water	+	+	—	—	—	—	—

<sup>a</sup> Described previously (9).<sup>b</sup> Described previously (10).<sup>c</sup> GenBank accession numbers from Makino et al. (5).

A second round of subtractive/kinetic enrichment PCR was carried out, with the *NBgl* 24-mer oligonucleotide as a primer. For the final round of the enrichment, the *NBgl* pair of oligonucleotides was digested off, the *JBgl* oligonucleotide pair was ligated back onto the tester DNA fragments, and subtractive/kinetic enrichment PCR was repeated.

After three rounds of subtractive hybridization of the adaptor-ligated DNAs, secondary identified PCR products were cloned into the plasmid pGEM-T Easy vector (Promega, Madison, WI.). Subtractive PCR inserts were then amplified from the clones (30 cycles of 94°C for 30 s, 57°C for 30 s, and 72°C for 1 min) by using T7 and SP6 primers followed by nested PCR with the primers provided with the subtraction kit (20 cycles of 95°C for 1 min and 72°C for 3 min). Subsequently, the

nested PCR products were purified by GENECLEAN II kit (Q-BIO Gene, Montreal, Canada).

Using the purified product as probes, we then performed dot blot hybridization against the tester and driver genomic DNAs or the plasmid DNAs in order to confirm whether the subtractive clones were specific to the tester DNA. Labeling probes, prehybridization, and hybridization were performed following the manufacturer's instructions. (Alkphos Direct manual: Amersham Biosciences Corp., Piscataway, NJ). Genomic or plasmid DNAs were transferred to a Hybond N+ nylon membrane (Amersham Biosciences) and hybridized with an alkaline phosphatase-labeled probe derived from a total of 28 subtractive PCR fragments. Positive hybridization was detected by coloration of sample spots on the membrane with nitroblue tetrazolium-5-bromo-4-chloro-3-indolylphosphate (NBT-BCIP) (Kirkegaard & Perry Laboratories, Gaithersburg, MD). The hybridization tests indicated that three (F2-1, F3-1, and F3-4) of the fragments were unique to the tester DNA (data not shown).

Subsequently, sequencing each fragment was initiated using two primers (SP6 promoter primer and T7 promoter primer) complementary to the sequence of the plasmid vector pGEM-T Easy (Promega). Both DNA strands were sequenced using a BigDye terminator cycle sequencing kit v3.0 (Applied Biosystems, Foster City, Calif.) on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems) following the manufacturer's protocol. Comparison with the sequences available in GenBank and EMBL by using the BLAST network service ([www.ncbi.nlm.nih.gov/BLAST/BLAST.cgi](http://www.ncbi.nlm.nih.gov/BLAST/BLAST.cgi)) (1) revealed that the 433-bp sequence of subtractive fragment F2-1 was identical to a partial gene sequence of *V. parahaemolyticus* (chromosome 1, positions 90860 to 91292, GenBank accession no. VP2905) encoding a hypothetical protein of unknown function (5). The 213-bp sequence of F3-1 and the 188-bp sequence of F3-4 were identified as partial gene sequences of *V. parahaemolyticus* chromosome 2, encoding a putative type III secretion system lipoprotein precursor, EprK (positions 235181 to 235393, GenBank accession no. VPA1367), and a hypothetical

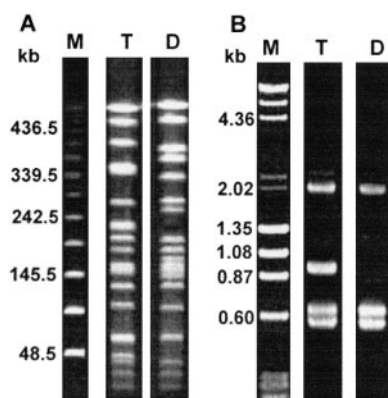


FIG. 1. Clonal similarity in PFGE (A) and AP-PCR (B) profiles between the pandemic *V. parahaemolyticus* strain KIID K7 and the nonpandemic strain KE 10491, used as a tester and a driver strain, respectively, for the genomic subtraction. Lane T, NIID K7 O3:K6; pandemic strain. Lane D, KE 10491 O3:K6; nonpandemic strain. Lane M, molecular size markers (bacteriophage lambda ladder; PFG Maker; New England BioLabs, Beverly, Mass. [A]), or a mixture of phage lambda DNA digested with HindIII and phage phiX174 DNA digested with HaeIII [B]).

TABLE 2. Results for 153 *V. parahaemolyticus* strains by PCR targeting the subtraction-derived VP2905 sequence, a partial sequence of the 16-kbp HU insertion sequence

Strain group <sup>a</sup>	Serovars included	No. of strains tested	No. of strains with positive PCR result <sup>b</sup>
Pandemic group	O3:K6, O4:K68, O1:K25, O1:K26, O1:KUT	55	54
Nonpandemic group	O1:K1, O1:K2, O1:K12, O1:K32, O1:K56, O1:KUT, O2:K3, O2:K28, O2:KUT, O3:K4, O3:K5, O3:K6, O3:K7, O3:K20, O3:K29, O3:K46, O3:K48, O3:K56, O3:K57, O4:K4, O4:K5, O4:K8, O4:K9, O4:K10, O4:K11, O4:K34, O4:K44, O5:K17, O5:KUT, O6:K18, O6:K38, O6:K46, O7:K19, O8:K41, O9:K23, O10:K9, O10:K19, O10:K24, O10:KUT, O11:K11, O11:K40, O11:KUT, O12:K40, O13:K65, O13:K72	98	None

<sup>a</sup> Determined by pandemic group-specific PCR tests described previously (9, 10).<sup>b</sup> 385-bp amplicon was produced.

protein of unknown function (positions 237821 to 238008, GenBank accession no. VPA1371), respectively (5).

From this sequence information, primer pairs specific to each gene were designed by using Primer Premier version 4.04 software (Premier Biosoft International, Palo Alto, Calif.). The primer sets were as follow: 5'-GGCTGCTATAACATTG AGCAC-3' (F2-1-F) and 5'-GAGGACTTGTGAAATCCCA TG-3' (F2-1-R) for partial amplification of VP2905, 5'-TTAG TTGTTTCAGGTTGCCG-3' (F3-1-F) and 5'-TCACACCAC CTGAACGAAGTA-3' (F3-1-R) for partial amplification of VPA1367, and 5'-GTGCGCCTAAATCGAGTTCT-3' (F3-4-F) and 5'-GCGACAGAAAGCCTGAACAT-3' (F3-4-R) for partial amplification of VPA1371.

With these primer sets, screening PCR tests were then performed on a panel of six pandemic strains (inclusive of the tester strain) and six nonpandemic strains (inclusive of the driver strain) listed in Table 1. The PCR amplification was performed in a total volume of 20 µl. Two microliters of each genomic DNA preparation (1 ng DNA/µl Tris-EDTA buffer) was added to the PCR master mix, which consisted of 2 µl of 10x PCR buffer (with 15 mM Mg<sup>2+</sup>; Promega), 0.25 µl of a deoxynucleoside triphosphate mixture (0.125 mM each deoxynucleoside triphosphate), 0.125 µl of each primer (0.125 µM of each primer), and 0.125 µl (0.625 U) of *Taq* DNA polymerase (Promega), with the remaining volume consisting of distilled water.

A GeneAmp PCR System 2700 thermal cycler (Applied Biosystems) was used for PCR amplification, consisting of 30 cycles of denaturation at 94°C for 30 s, primer annealing at 55°C for 30 s, and extension at 72°C for 30 s. Five microliters of the PCR products were then electrophoresed on 2% agarose gels, stained with ethidium bromide (0.25 µg/ml), and photographed under UV light for positive PCR products (385 bp for VP2905, 504 bp for VPA1367, and 584 bp for VPA1371). These results revealed that the gene sequence VP2905 was specific only to the six pandemic strains (Table 1).

To further confirm that this gene sequence was specific to the pandemic group, we performed additional PCR tests with the primer set of F2-1-F and F2-1-R on a total of 153 *V. parahaemolyticus* strains of various serotypes (Table 2). These included 55 pandemic strains and 98 nonpandemic strains that were grouped by two separate pandemic group-specific PCR tests as described previously (9, 10). The test results revealed that all pandemic strains except one, KIH 03-57 (a clinical isolate of O4:K68), had the gene sequence VP2905, while all

nonpandemic group strains did not (Table 2). This gene is located within an approximately 16-kb sequence which is inserted in the open reading frame of a histone-like DNA-binding protein, HU-α (5). Coincidentally, this finding was suggested by those of Williams et al. (12), in which pandemic *V. parahaemolyticus* has a histone-like DNA-binding protein, HU-α, with few amino acid differences from that of a nonpandemic strain due to a 16-kbp insertion sequence at the 3' end of the open reading frame of HU-α.

Histone-like proteins of bacterial species are small and heat stable and bind to single- and double-stranded DNA, thereby altering DNA structure and topology (4). In enteric bacteria, mutations in a histone-like protein nucleoid-structuring protein result in modification in their drug resistance (7). In streptococci, Stinson et al. (11) also demonstrated that an extracellular histone-like protein has a potential role associated with tissue inflammation in their pathogenicity. We can speculate that the modification of HU-α by several terminal amino acids might endow *V. parahaemolyticus* with a novel pathogenic factor, thus enabling it to be a pandemic clone.

The 16-kbp insertion sequence includes at least 10 genes all encoding hypothetical proteins of unknown functions. These proteins may be responsible for novel phenotypic or pathogenic properties yet to be elucidated for the pandemic group. It should also be noted that the insertion sequence includes a gene encoding a putative phage protein (5) and has an identical sequence (TTCTTCAG) immediately after its 5' and 3' ends, suggesting that the insertion sequence was of phage origin.

Accordingly, we examined further the exceptional pandemic strain KIH 03-57 by PCRs targeting several different regions within or adjacent to the insertion sequence and found that the strain had an intact HU-α sequence, completely devoid of the insertion sequence (data not shown). The evidence suggests that the strain might have been cured of the putative phage. Further investigation is in progress to detect any phenotypic difference between the pandemic strains and this strain in order to evaluate any involvement of the insertion sequence in the emergence of pandemic *V. parahaemolyticus*.

We are grateful to R. A. Whiley of the Department of Oral Microbiology, St. Bartholomew's and Royal London School of Medicine and Dentistry, for his valuable comments on an earlier draft of this paper.

## REFERENCES

1. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
2. Blake, P. A., R. E. Weaver, and D. G. Hollis. 1980. Disease of humans (other than cholera) caused by vibrios. *Annu. Rev. Microbiol.* **34**:341–367.
3. Calia, K. E., M. K. Waldor, and S. B. Calderwood. 1998. Use of representational difference analysis to identify genomic differences between pathogenic strains of *Vibrio cholerae*. *Infect. Immun.* **66**:849–852.
4. Dorman, C. J., and P. Deighan. 2003. Regulation of gene expression by histone-like proteins in bacteria. *Curr. Opin. Genet. Dev.* **13**:179–184.
5. Makino, K., K. Oshima, K. Kurokawa, K. Yokoyama, T. Uda, K. Tagomori, Y. Iijima, N. Najima, M. Nakano, A. Yamashita, Y. Kubota, S. Kimura, Y. Yasunaga, T. Honda, H. Shinagawa, M. Hattori, and T. Iida. 2003. Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholerae*. *Lancet* **361**:743–749.
6. Matsumoto, C., J. Okuda, M. Ishibashi, M. Iwanaga, P. Garg, T. Ramamurthy, H. -C. Wong, A. DePaola, Y. B. Kim, M. J. Albert, and M. Nishibuchi. 2000. Pandemic spread of an O3:K6 clone of *Vibrio parahaemolyticus* and emergence of related strains evidenced by arbitrarily primed PCR and *toxRS* sequence analyses. *J. Clin. Microbiol.* **38**:578–585.
7. Nishino, K., and A. Yamaguchi. 2004. Role of histone-like protein H-NS in multidrug resistance of *Escherichia coli*. *J. Bacteriol.* **186**:1423–1429.
8. Okuda, J., M. Ishibashi, E. Hayakawa, T. Nishino, Y. Takeda, A. K. Mukhopadhyay, S. Garg, S. K. Bhattacharya, G. B. Nair, and M. Nishibuchi. 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**:3150–3155.
9. Okura, M., R. Osawa, A. Iguchi, E. Arakawa, J. Terajima, and H. Watanabe. 2003. Genotypic analyses of *Vibrio parahaemolyticus* and development of a pandemic group-specific multiplex PCR assay. *J. Clin. Microbiol.* **41**:4676–4682.
10. Okura, M., R. Osawa, A. Iguchi, M. Takagi, E. Arakawa, J. Terajima, and H. Watanabe. 2004. PCR-based identification of pandemic group *Vibrio parahaemolyticus* with a novel group-specific primer pair. *Microbiol. Immunol.* **48**:787–790.
11. Stinson, M. W., R. McLaughlin, S. H. Choi, Z. E. Juarez, and J. Barnard. 1998. Streptococcal histone-like protein: primary structure of *hlpA* and protein binding to lipoteichoic acid and epithelial cells. *Infect. Immun.* **66**:259–265.
12. Williams, T. L., S. M. Musser, J. L. Nordstrom, A. DePaola, and S. R. Monday. 2004. Identification of a protein biomarker unique to the pandemic O3:K6 clone of *Vibrio parahaemolyticus*. *J. Clin. Microbiol.* **42**:1657–1665.
13. Wong, H.-C., S.-H. Liu, T.-K. Wang, C.-L. Lee, C.-S. Chiou, D.-P. Liu, M. Nishibuchi, and B.-K. Lee. 2000. Characteristics of *Vibrio parahaemolyticus* O3:K6 from Asia. *Appl. Environ. Microbiol.* **66**:3981–3986.