2007

Identification of *Vibrio* Isolates by a Multiplex PCR Assay and *rpoB* Sequence Determination

Cheryl L. Tarr
National Center for Infectious Diseases, crt6@cdc.gov

Jayna S. Patel
National Center for Infectious Diseases

Nancy Puhr
National Center for Infectious Diseases

Evangeline Sowers
National Center for Infectious Diseases

Cheryl A. Bopp
National Center for Infectious Diseases

See next page for additional authors

JOURNAL OF CLINICAL MICROBIOLOGY, Jan. 2007, p. 134–140 Vol. 45, No. 1
0095-1137/07/$08.00+0 doi:10.1128/JCM.01544-06

Identification of *Vibrio* Isolates by a Multiplex PCR Assay and *rpoB* Sequence Determination▽

Cheryl L. Tarr,[1,2]* Jayna S. Patel,[1] Nancy D. Puhr,[1] Evangeline G. Sowers,[1] Cheryl A. Bopp,[1] and Nancy A. Strockbine[1]

Foodborne and Diarrheal Diseases Branch, Division of Bacterial and Mycotic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia 30333,[1] and Battelle Memorial Institute, Battelle Science and Technology International, Chemical and Environmental Technologies, Atlanta Analytical Chemistry Group, 2971 Flowers Road South, Atlanta, Georgia 30341[2]

Received 25 July 2006/Returned for modification 1 September 2006/Accepted 2 November 2006

Vibrio*, a diverse genus of aquatic bacteria, currently includes 72 species, 12 of which occur in human clinical samples. Of these 12, three species—*Vibrio cholerae*, *Vibrio parahaemolyticus*, and *Vibrio vulnificus*—account for the majority of *Vibrio* infections in humans. Rapid and accurate identification of *Vibrio* species has been problematic because phenotypic characteristics are variable within species and biochemical identification requires 2 or more days to complete. To facilitate the identification of human-pathogenic species, we developed a multiplex PCR that uses species-specific primers to amplify gene regions in four species (*V. cholerae*, *V. parahaemolyticus*, *V. vulnificus*, and *V. mimicus*). The assay was tested on a sample of 309 *Vibrio* isolates representing 26 named species (including 12 human pathogens) that had been characterized by biochemical methods. A total of 190 isolates that had been identified as one of the four target species all yielded results consistent with the previous classification. The assay identified an additional four *V. parahaemolyticus* isolates among the other 119 isolates. Sequence analysis based on *rpoB* was used to validate the multiplex results for these four isolates, and all clustered with other *V. parahaemolyticus* sequences. The *rpoB* sequences for 12 of 15 previously unidentified isolates clustered with other *Vibrio* species in a phylogenetic analysis, and three isolates appeared to represent unnamed *Vibrio* species. The PCR assay provides a simple, rapid, and reliable tool for identification of the major *Vibrio* pathogens in clinical samples, and *rpoB* sequencing provides an additional identification tool for other species in the genus *Vibrio*.

The genus *Vibrio* is a highly diverse group of gram-negative bacteria that contains approximately 72 species (www.bacterio.net). The group includes symbionts and commensals that are found in or on marine animals, as well as many species that are pathogenic to animals (21). There are 12 species that are routinely isolated from human clinical samples, and the diseases in which they are implicated include diarrheal disease, septicemia, and wound infections (6).

Three species account for the majority of human *Vibrio* infections. Toxigenic *Vibrio cholerae* is the causative agent of the disease cholera and is acquired through ingestion of contaminated food or water. Infection may lead to a profuse, watery diarrhea that can lead to severe dehydration and death if left untreated. Worldwide, large outbreaks are caused by toxigenic strains of serogroups O1 and O139 that produce the cholera toxin, but in the United States, nontoxigenic strains predominate among the cases reported to the Centers for Disease Control and Prevention (CDC) (4). A total of 48 cases of cholera occurred in the United States in 2004 and were reported to the CDC (4), including eight cases infected with toxigenic *V. cholerae*, four of which were associated with travel, and 40 cases infected with nontoxigenic cholera. *Vibrio cholerae* accounted for 9.8% of all *Vibrio* isolates reported in the United States to the CDC in 2004 (4).

Vibrio parahaemolyticus is a halophilic bacterium that naturally inhabits coastal waters. Infection can result in acute gastroenteritis (6), and wound infections can occur as a result of exposing an open sore to warm seawater. Infection is rarely fatal, and the majority of cases arise from ingestion of undercooked seafood, especially raw oysters. A total of 240 cases (49.3% of *Vibrio* cases in the United States) were reported to the CDC in 2004 (4). *Vibrio vulnificus* is another pathogenic species that is naturally present in warm seawater. The typical route of infection is ingestion of raw oysters, but infection can occur through wounds that are exposed to seawater or infected organisms. The bacterium can cause wound infections, gastroenteritis, or septicemia (6). The species is an opportunistic pathogen that causes severe infections in individuals who are immunocompromised and in those with chronic liver disease; in about 50% of such cases fatal complications result from infection. A total of 92 isolates (18.9% of *Vibrio* isolates in the United States) were reported to the CDC in 2004 (4), of which 33% caused fatal infections. Although *V. parahaemolyticus* accounted for the majority of *Vibrio* infections in the United States, *V. vulnificus* was responsible for 82% of the fatalities in which a *Vibrio* species was implicated. Other *Vibrio* species that are routinely isolated from human clinical samples include *V. alginolyticus*, *V. fluvialis*, *V. mimicus*, *Vibrio* (*Grimontia*) *hollisae*, *Vibrio* (*Photobacterium*) *damsela*, *V. furnissii*, *V. cincinnatiensis*, *V. harveyi*, and *V. metschnikovii*. Of these species, *V. mimicus* was of special interest to us: it is closely related to *V.*

* Corresponding author. Mailing address: Foodborne and Diarrheal Diseases Branch, Division of Bacterial and Mycotic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, 1600 Clifton Road NE, Mailstop C-03, Atlanta, GA 30333. Phone: (404) 639-2011. Fax: (404) 639-3333. E-mail: crt6@cdc.gov.

▽ Published ahead of print on 8 November 2006.

TABLE 1. Sequence of primers and final concentration and product size for each of five primer pairs in the multiplex PCR

Target species	Primer	Sequence (5′ to 3′)	Concn (μM)	Amplicon size (bp)
V. cholerae	*Vc.sodB*-F	AAG ACC TCA ACT GGC GGT A	0.5	248
	Vc.sodB-R	GAA GTG TTA GTG ATC GCC AGA GT		
V. mimicus	*Vm.sodB*-F	CAT TCG GTT CTT TCG CTG AT	0.75	121
	Vm.sodB-R2	GAA GTG TTA GTG ATT GCT AGA GAT		
V. parahaemolyticus	*Vp.flaE*-79F	GCA GCT GAT CAA AAC GTT GAG T	1.0	897
	Vp.flaE-934R	ATT ATC GAT CGT GCC ACT CAC		
V. vulnificus	*Vv.hsp*-326F	GTC TTA AAG CGG TTG CTG C	0.25	410
	Vv.hsp-697R	CGC TTC AAG TGC TGG TAG AAG		
All *Vibrio* spp.	*V*.16S-700F	CGG TGA AAT GCG TAG AGA T	0.05	663
	V.16S-1325R	TTA CTA GCG ATT CCG AGT TC		

cholerae, and so it can be difficult to differentiate the two species because they share many phenotypic characteristics (5). *V. mimicus* can also carry the cholera toxin gene as well as other virulence-associated genes that are used to identify *V. cholerae* (3).

Accurate phenotypic identification of *Vibrio* species is problematic, largely because of the great variability in biochemical characteristics (21). O'Hara et al. (16) evaluated six commercial systems for the ability to identify the 12 species of *Vibrio* found in clinical samples. They found that many isolates are not accurately identified by commercial methods, with the accuracy of systems ranging from 63.9% to 80.9% (16). Most members of the genus are halophilic, and the addition of NaCl is often required for enzymatic activity; however, the concentration of NaCl can affect the biochemical profile and lead to erroneous identification with at least one system (API 20E [13]). The characteristics shared by *Aeromonas* and *Vibrio* also result in classification of isolates in the wrong genus (1, 18). Additional drawbacks of biochemical methods are that they are time-extensive (i.e., reactions must incubate for an extended time period, often requiring 2 to 7 days to complete) and that the interpretation of results requires specialized training that may not be available to all laboratories.

Molecular methods that utilize the PCR and nucleotide sequence determination overcome many of the limitations of phenotypic methods. Molecular techniques, particularly nucleotide sequence determination, provide data that are objectively scored to provide an unambiguous identification. Nucleotide sequences can be placed into a phylogenetic framework, which allows an assessment of the genetic relationships among isolates and can reveal isolates that represent undescribed species. Most importantly, methods that utilize the PCR can lead to identification of an isolate within hours as opposed to days and can be used on small quantities of cells, including those that are not viable or are otherwise unculturable.

A rapid and accurate molecular method that identifies multiple species in one assay would be a useful tool for clinical laboratories and would enhance *Vibrio* surveillance and diagnosis. A simple molecular tool could also lead to better estimates of the incidence of *Vibrio* infections in the United States. While a total of 487 *Vibrio* isolates were reported to the CDC in 2004 (4), the incidence of *Vibrio* infections is likely to be much greater, particularly because only infections due to toxigenic *V. cholerae* serogroup O1 or O139 were nationally notifiable prior to 2006.

In this report, we present a two-step approach to identify *Vibrio* isolates that are pathogenic to humans: a multiplex PCR assay to identify the most commonly encountered *Vibrio* isolates (*V. cholerae*, *V. parahaemolyticus*, *V. vulnificus*, and *V. mimicus*) and a robust sequence-based approach using *rpoB* sequencing and phylogenetic analysis that can cluster isolates into other *Vibrio* species and can clearly differentiate *Vibrio* from other closely related genera, such as *Aeromonas*. The multiplex PCR primarily targets variation in conserved housekeeping genes. We also generated a *Vibrio rpoB* reference database and then determined nucleotide sequence for an 870-bp portion of the *rpoB* gene to identify *Vibrio* isolates that could not be classified on the basis of phenotype or the multiplex PCR. Sequence analysis of the *rpoB* gene has been proposed as a method for bacterial identification (15) and has been used to identify bacterial isolates in a number of taxa, including *Pseudomonas* (2), *Corynebacterium* (8), and the staphylococci (14). The combination of the multiplex PCR and *rpoB* sequencing allowed us to quickly and accurately identify the predominant *Vibrio* species that are pathogenic to humans, unambiguously place unidentified *Vibrio* isolates into species within a phylogenetic framework, and identify isolates that appear to represent as-yet-unnamed *Vibrio* species.

MATERIALS AND METHODS

A total of 316 bacterial isolates were used in the study: 309 *Vibrio* isolates, representing 26 named species, as well as isolates in the genera *Photorhabdus* (one), *Salmonella* (one), *Yersinia* (one), and *Aeromonas* (four). We included *Vibrio* (*Grimontia*) *hollisae* and *Vibrio* (*Photobacterium*) *damsela* among the 12 species that are pathogenic to humans. All isolates were obtained from the frozen culture collection in the Foodborne and Diarrheal Diseases Branch of the CDC. Isolates were stored at −70°C in tryptic soy broth with 20% glycerol.

PCR template preparation. Crude lysates were prepared for PCR as follows. Species pathogenic to humans were grown overnight at 35°C on tryptic soy agar plates with 5% sheep blood. Marine species were grown on marine agar plates and incubated at either room temperature or 20°C (*V. logei*) for 48 h. After incubation, a single colony was scraped from the plate's surface and suspended in 200 μl of 1× Tris-EDTA, pH 8.0. The suspension was then heated at 95°C for 10 min and centrifuged for 2 min to pellet cellular debris.

Development of the multiplex assay. We initially used information in GenBank (http://www.ncbi.nlm.nih.gov/GenBank/) to select target genes for the multiplex assay. *Vibrio* species were well represented for two variable genes in the database: *hsp60* (11) and *sodB* (Yoshiyuki Yamada, unpublished data). Sequences for the two genes were downloaded from GenBank, and each was aligned with ClustalX (22). We visually searched for sequence regions that potentially could differentiate individual pathogenic species from other *Vibrio* spp., and candidate primer sites were identified for *V. vulnificus*, *V. mimicus*, and *V. cholerae*. Intraspecific variation within each site was then assessed for *V. vulnificus* in *hsp60* (18 isolates; primers from reference 23) and for *V. cholerae* and *V. mimicus* in

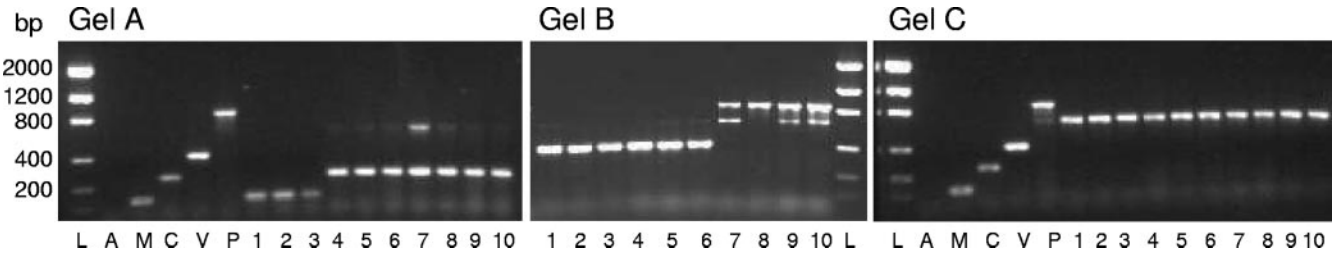

FIG. 1. Agarose gels showing results of the multiplex PCR. L, low-molecular-mass ladder. Controls are labeled as follows: A, no-DNA control; M, *V. mimicus*; C, *V. cholerae*; V, *V. vulnificus*; P, *V. parahaemolyticus*. Gel A: samples 1 to 3, *V. mimicus*; samples 4 to 10, *V. cholerae* serogroup O1 Ogawa, Inaba, non-O1, O139, O141, O22, and O155, respectively. Gel B: samples 1 to 6, *V. vulnificus*; samples 7 to 10, *V. parahaemolyticus* O4:K68, O3:K6, O3:K6, and O3:K48, respectively. Gel C: sample 1, *V. alginolyticus*; sample 2, *V. cincinnatiensis*; sample 3, *P. damsela*; sample 4, *V. fluvialis*; sample 5, *V. furnissii*; sample 6, *V. harveyi*; sample 7, *V. hollisae*; sample 8, *V. metschnikovii*; sample 9, *V. splendidus*; sample 10, *V. campbellii*. Note that the 16S rRNA control amplicon is usually outcompeted by the species-specific amplicons, but it is clearly visible in some samples, particularly samples of *V. parahaemolyticus* (e.g., gel A, *V. cholerae* sample 7, and gel B, *V. parahaemolyticus* samples 7, 9, and 10).

sodB (48 *V. cholerae* isolates and 19 *V. mimicus* isolates; primers courtesy of Yoshiyuki Yamada). For the *V. parahaemolyticus* marker, we downloaded four *Vibrio* genome sequences from GenBank (*V. parahaemolyticus*, NC 004603/NC 0046053; *V. cholerae*, NC 002505/NC 002506; *V. fischeri*, NC 006840/NC 006841; and *V. vulnificus*, NC 004459/004460) and then extracted and aligned sequences for the flagellin genes. Primers were designed to match the *flaE* sequence of *V. parahaemolyticus*. A highly conserved, positive internal control for the multiplex PCR was developed on the basis of an alignment of *Vibrio* and *Enterobacteriaceae* 16S rRNA gene sequences.

Optimization was performed for each primer pair individually and then all primers in combination. Final primer concentrations were adjusted to give approximately equal signals for each gene fragment. Each reaction contained 2 units of AmpliTaq Gold (Applied Biosystems, Foster City, CA), and final concentrations of each deoxynucleoside triphosphate (dNTP) and $MgCl_2$ of 0.2 mM and 1.5 mM, respectively, and primer concentrations ranged from 0.05 μM to 1 μM (Table 1). Thus, a typical 20-μl reaction mixture contained 1.5 μl crude lysate, 0.2 μl of AmpliTaq Gold, 2 μl dNTP stock (2 mM each), 2 μl 10× buffer, 1.2 μl $MgCl_2$ stock (25 mM), 2.9 μl double-distilled water, and the various amounts of each 10 μM stock of the primers: *Vc* primers, 1 μl each; *Vm* primers, 1.5 μl each; *Vp* primers, 2 μl each; *Vv* primers, 0.5 μl each; and 16S rRNA primers, 0.1 μl each. The thermal cycling profile was as follows: a 15-min soak at 93°C followed by 35 cycles of 92°C for 40 s, 57°C for 1 min, and 72°C for 1.5 min and a final soak at 72°C for 7 min. PCR amplicons were electrophoresed in 1.5% agarose for 1 h at 100 V and then visualized by ethidium bromide staining and UV illumination. The result for each isolate was determined by comparison to those for control isolates of *V. mimicus*, *V. cholerae*, *V. vulnificus*, and *V. parahaemolyticus*. Any isolate for which the original identification did not match our multiplex results, or which could not be classified based on the PCR or phenotypic characters, was subjected to *rpoB* sequence determination and phylogenetic analysis as described below.

***rpoB* amplification and sequence determination.** We developed primers to amplify and sequence a 984-bp portion of the *rpoB* gene by first extracting *rpoB* sequences from *Vibrio* genome sequences and aligning them with *Enterobacteriaceae* sequences reported in the work of Mollet et al. (15). We used primers CM32b (15) and 1110F (5′-GTA GAA ATC TAC CGC ATG ATG-3′) for amplification and two additional primers for sequencing (1661F, 5′-TTY ATG GAY CAR AAC AAC CC-3′; 1783R, 5′-GGA CCT TYA GGN GTT TCG AT-3′). The 50-μl PCR mixtures contained 2 units of AmpliTaq Gold, 1.5 μl of crude lysate, and the following reagents at the final concentrations listed in parentheses: primer (0.5 μM each), dNTPs (0.2 mM each), Betaine (1 mM) (Sigma-Aldrich, St. Louis, MO), and $MgCl_2$ (1.5 mM). The touchdown PCR was preceded by a 15-min soak at 93°C. The thermal cycle was as follows: 92°C for

TABLE 2. Summary of multiplex results for 316 bacterial isolates, listed by original species identification

Original identification	No. of isolates	Multiplex result
V. cholerae[a]	40	*V. cholerae* amplicon
V. parahaemolyticus[b]	63	*V. parahaemolyticus* amplicon
V. vulnificus	60	*V. vulnificus* amplicon
V. mimicus	27	*V. mimicus* amplicon
V. alginolyticus	22	16S rRNA amplicon
V. alginolyticus	1	*V. parahaemolyticus* amplicon[c]
V. fluvialis	28	16S rRNA amplicon
P. damsela	7	16S rRNA amplicon
V. furnissii	1	16S rRNA amplicon
V. harveyi/trachuri	2	16S rRNA amplicon
G. hollisae	10	16S rRNA amplicon
V. metschnikovii	1	16S rRNA amplicon
V. cincinnatiensis	2	16S rRNA amplicon
V. aestuarianus	1	16S rRNA amplicon
V. anguillarum	5	16S rRNA amplicon
V. campbellii	1	16S rRNA amplicon
V. costicola	1	16S rRNA amplicon
V. halioticoli	4	16S rRNA amplicon
V. logei	1	16S rRNA amplicon
V. natriegens	2	16S rRNA amplicon
V. nereis	1	16S rRNA amplicon
V. nigripulchritudo	1	16S rRNA amplicon
V. ordalii	1	16S rRNA amplicon
V. orientalis	3	16S rRNA amplicon
V. pectenicida	2	16S rRNA amplicon
V. pelagius	4	16S rRNA amplicon
V. splendidus	3	16S rRNA amplicon
Vibrio sp.	12	16S rRNA amplicon
Vibrio sp.	3	*V. parahaemolyticus* amplicon[c]
Aeromonas sp.	4	16S rRNA amplicon
Photorhabdus sp.	1	16S rRNA amplicon
Salmonella sp.	1	16S rRNA amplicon
Yersinia sp.	1	16S rRNA amplicon

[a] Included serogroups O1, O139, and O141.
[b] Included serotype O3:K6.
[c] Results confirmed by 16S rRNA and *rpoB* sequence determination.

TABLE 3. Genetic variation within 870 bp of the *rpoB* gene in 11 *Vibrio* species[a]

Species	No. of strains	No. of STs[b]	No. of polymorphic sites	$d_S/100$[c]	$d_N/100$[d]	$d_N - d_S/100$
V. alginolyticus	10	5	6	0.6	0.00	−0.60
V. harveyi	4	4	18	3.2	0.16	−3.05
V. parahaemolyticus	11	6	17	1.8	0.03	−1.73
V. vulnificus	7	7	22	3.7	0.00	−3.69
V. metschnikovii	2	2	1	0.4	0.00	−0.39
V. fluvialis	6	5	12	2.0	0.00	−2.00
V. furnissii	3	3	6	1.6	0.00	−1.57
V. mimicus	5	5	18	2.9	0.07	−2.86
V. cholerae	7	3	6	0.7	0.00	−0.67
V. damsela	3	3	4	1.0	0.00	−1.04
V. hollisae	5	2	17	2.6	0.07	−2.52

[a] Calculations include the isolates identified by *rpoB* sequencing.
[b] ST, sequence type; no. of STs indicates the number of unique sequences.
[c] $d_S/100$, average number of synonymous substitutions per 100 synonymous sites.
[d] $d_N/100$, average number of nonsynonymous substitutions per 100 nonsynonymous sites.

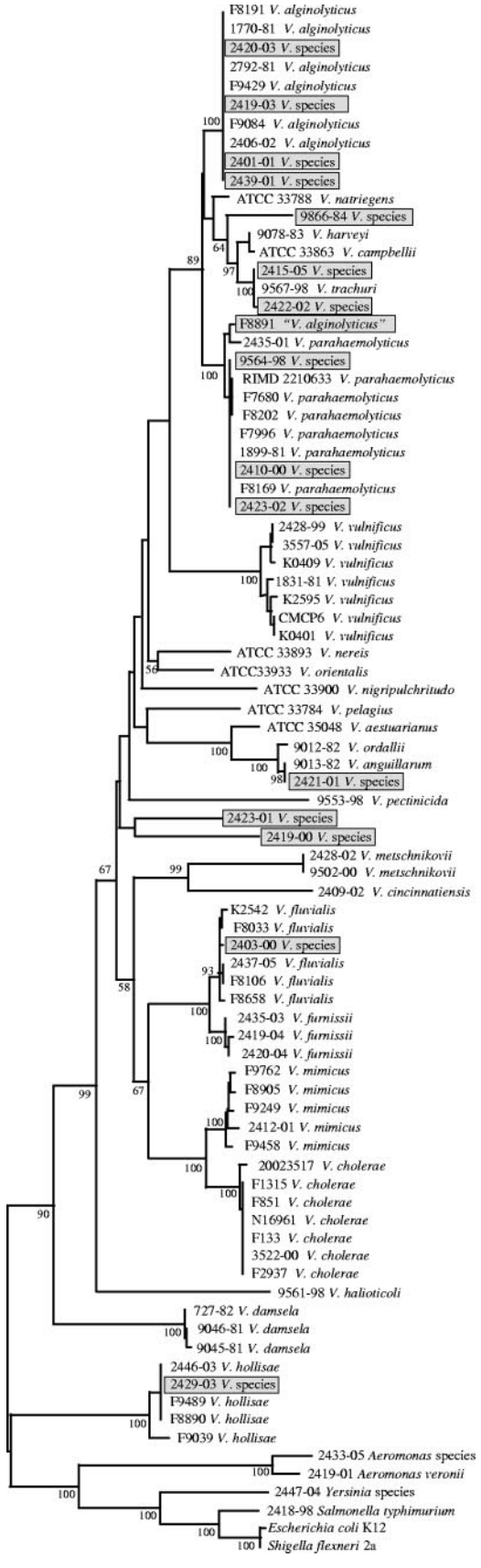

FIG. 2. Neighbor-joining tree constructed from Kimura two-parameter distances calculated from 870 bp of *rpoB* sequence. The numbers at each node indicate the percentage of bootstrap replications in which a particular node appears (shown for nodes where value was >50%). Isolates that were identified by *rpoB* sequence analysis are shown highlighted in gray and are designated by the original identification. The *rpoB* sequences for *Escherichia coli* K-12 and *Shigella flexneri* 2a were extracted from complete genome sequences downloaded from GenBank (accession numbers NC 000913 and NC 004337, respectively). GenBank accession numbers for sequences generated in this study are EF 064367 to EF 064446.

40 s, 60°C to 50°C for 1 min (decreasing 1°C each cycle, then 50°C for 30 cycles), 72°C for 1.5 min, and a final soak at 72°C for 7 min. Agarose gel electrophoresis and ethidium bromide staining were used to check the quality and quantity of PCR product, and amplicons were then purified with the QIAquick PCR purification kit (QIAGEN, Valencia, CA). The product was quantified by comparison to a known quantity of the low-molecular-mass ladder (Invitrogen, Carlsbad, CA) on an ethidium bromide-stained gel. We used the DTCS Quickstart kit (Beckman Coulter, Fullerton, CA), following manufacturer's instructions with the following specifications: we used 1 μl of a 20 μM primer stock solution, 50 fmol of template, an annealing temperature of 50°C, and Betaine at a final concentration of 1 mM. Sequence determination was conducted on a Beckman CEQ 8000 Genetic Analysis system.

Chromatograms were imported into SEQMAN in the package DNAStar (Lasergene), assembled into contigs, edited, and exported as FASTA files. The *rpoB* sequences were aligned with ClustalX, and pairwise genetic distances were calculated using the Kimura two-parameter method in the program MEGA (10). The *rpoB* sequences from three published *Vibrio* genome sequences were also included in the analysis (7, 9, 12). A neighbor-joining tree was constructed from the genetic distance matrix, and the bootstrap confidence intervals for each node were calculated over 1,000 replicate trees.

Nucleotide sequence accession numbers. GenBank accession numbers for sequences generated in this study are EF 064367 to EF 064446.

RESULTS AND DISCUSSION

Multiplex results. The multiplex assay was tested on a collection of 316 bacterial isolates that included 309 *Vibrio* isolates. Among these 309 isolates were 190 isolates representing the four target species, isolates representing an additional 22 named species (see Table 2), and a total of 15 isolates that had been identified only to genus level. Representative gels are shown in Fig. 1. All 190 isolates of the four target species yielded the expected species-specific amplicon (Table 2). For three of the target species, the 16S rRNA amplicon was usually outcompeted by the species-specific amplicon, but both amplicons were usually visible for *V. parahaemolyticus* (Fig. 1). Only the 16S rRNA fragment was amplified from the isolates representing the remaining 22 species, as expected. Of the 15 isolates that were previously identified only to genus level, three (9564-98, 2410-00, and 2423-02) yielded the *V. parahaemolyticus* amplicon. Also, the *V. parahaemolyticus* amplicon was obtained for one isolate that was originally reported to the CDC as *V. alginolyticus* (F8891), a result that is not too surprising given the similar biochemical profiles of the two species (6). The classification of these four isolates as *V. parahaemolyticus* was confirmed using nucleotide sequencing and evolutionary analysis of a segment of the *rpoB* gene as described below. Thus, we found that the multiplex PCR rapidly and accurately discriminated four *Vibrio* species that account for the majority of *Vibrio* infections in the United States each year (4).

TABLE 4. Average nucleotide distance between each pair of *Vibrio* species calculated for the *rpoB* gene[a]

Species	Nucleotide distance											
	1	2	3	4	5	6	7	8	9	10	11	12
1. *V. alginolyticus*		35.8	27.6	78.1	140.5	145.8	106.3	109.0	111.1	117.4	145.9	170.6
2. *V. harveyi*	4.2		44.6	103.2	149.0	150.5	108.5	110.8	108.3	116.0	151.0	180.3
3. *V. parahaemolyticus*	3.2	5.3		81.0	141.0	146.9	111.2	111.5	111.0	117.5	149.2	175.5
4. *V. vulnificus*	9.6	13.0	10.0		126.5	130.6	108.5	116.2	116.3	117.4	147.3	164.2
5. *V. metschnikovii*	18.4	19.7	18.5	16.4		116.5	138.3	142.8	132.7	124.2	173.5	198.9
6. *V. cincinnatiensis*	19.2	19.9	19.4	17.0	14.9		135.3	131.7	121.4	120.4	188.0	197.2
7. *V. fluvialis*	13.4	13.7	14.1	13.8	18.2	17.7		18.3	81.5	83.4	163.3	173.2
8. *V. furnissii*	13.8	14.1	14.2	14.9	18.9	17.1	2.1		83.4	83.7	166.7	176.3
9. *V. mimicus*	14.1	13.7	14.1	14.9	17.4	15.7	10.0	10.3		35.1	170.2	177.1
10. *V. cholerae*	15.0	14.8	15.0	15.1	16.1	15.6	10.3	10.4	4.2		170.6	173.5
11. *V. damsela*	19.0	19.8	19.5	19.3	23.4	25.7	21.7	22.2	22.8	22.8		139.0
12. *V. hollisae*	23.0	24.5	23.7	22.0	27.8	27.3	23.3	23.8	24.0	23.4	18.2	

[a] Upper right triangle, number of nucleotide differences; lower left triangle, genetic distance based on Kimura two-parameter model.

***rpoB* sequencing results.** We obtained 870 bp of sequence from a total of 60 *Vibrio* isolates that had been classified by biochemical methods, 15 isolates that could not be classified on the basis of phenotype, and four isolates of other enteric bacteria that were included to root the phylogenetic tree. The *rpoB* gene was polymorphic within all 11 of the pathogenic *Vibrio* species for which we sequenced more than one isolate (Table 3). The synonymous substitution rate (d_S) was greater than the nonsynonymous substitution rate (d_N), as was expected for a conserved housekeeping gene (Table 3).

All *Vibrio* species included in the analysis were distinguishable from isolates of other species on the basis of their *rpoB* sequence (Fig. 2). The level of bootstrap support for each of the pathogenic species clusters was very high, ranging from 93% for *V. fluvialis* isolates to 100% for most of the other *Vibrio* species (Fig. 2). The closest species pair was *V. harveyi* and *V. campbellii*, which differed at two nucleotide sites (0.23%). Genetic distances for other closely related species pairs were 2.1% (*V. fluvialis* and *V. furnissii*), 3.2% (*V. alginolyticus* and *V. parahaemolyticus*), and 4.2% (*V. cholerae* and *V. mimicus*) (Table 4). The species clusters based on the *rpoB* phylogeny were consistent with the results based on sequencing of other loci (19).

We sequenced a total of 15 isolates that were not classified by phenotypic methods because they differed in one or more biochemical characteristics from other *Vibrio* isolates. Information on the isolates is shown in Table 5. The sequences from 10 of the isolates fell within the core *Vibrio* group: four clustered with *V. alginolyticus*, three with *V. parahaemolyticus*, and three with the *V. harveyi*-related group. The sequences from the remaining five isolates were scattered in the tree, with one isolate clustering with *V. anguillarum*, one with *V. fluvialis*, and one with *V. hollisae*. Two of the *Vibrio* sequences did not cluster with any of the other groups in the tree. An isolate (F8891) that had been reported to the CDC as *V. alginolyticus* did not cluster with other sequences for that species but instead clustered with *V. parahaemolyticus*, which was consistent with the results from the multiplex PCR assay.

The cluster of isolates including *V. harveyi* raised questions about classification of these closely related isolates. Thompson

TABLE 5. Characteristics of 15 isolates that were identified by *rpoB* sequence determination and phylogenetic analysis

Isolate	Source	Yr[a]	Locality	Closest match[b]	Molecular identification (*rpoB*)	Atypical phenotype[c]
9866-84	Environment	1984	Aegean Sea	*V. harveyi* group 2	*V. harveyi* related	NA
9564-98	Unknown	1998	Texas	*V. harveyi*	*V. parahaemolyticus*	Ind^- Orn^- Lip^-
2403-00	Blood	2000	Texas	*V. fluvialis*	*V. fluvialis*	Oxi^-
2410-00	Stool	2000	North Carolina	*V. alginolyticus*	*V. parahaemolyticus*	Cit^+ Suc^+ Sal^+ $ONPG^+$
2419-00	Stool	2000	Colorado	No match	No close relationship	NA
2401-01	Gall bladder	2001	Massachusetts	*V. harveyi*	*V. alginolyticus*	$VP\alpha^-$ Mot^-
2421-01	Oyster (outbreak)	2001	Nevada	No match	*V. anguillarum*	NA
2423-01	Blood	2001	Virginia	*V. vulnificus*	No close relationship	NA
2439-01	Stool	2001	Maryland	*V. alginolyticus*	*V. alginolyticus*	Cit^+ Sal^+
2422-02	Wound	2002	Hawaii	*V. alginolyticus*	*V. harveyi* (*trachuri*)	MR^- Lip^+ Gal^+
2423-02	Stool	2002	Louisiana	*V. vulnificus*	*V. parahaemolyticus*	Cit^+ Suc^+ Sal^+ Cel^+
2419-03	Nasal sinus	2003	Indiana	*V. alginolyticus*	*V. alginolyticus*	$VP\alpha^-$ Cit^+
2420-03	Nasal sinus	2003	Indiana	*V. alginolyticus*	*V. alginolyticus*	$VP\alpha^-$ Cit^+ $Acet^+$
2429-03	Stool	2003	Louisiana	*V. vulnificus*	*V. hollisae*	MR^+ Gly^+
2415-05	Blood	2005	Hawaii	*V. harveyi*	*V. harveyi* (*trachuri*)	Orn^+ Tyr^+ Gal^+

[a] Year in which the isolate was received at the Centers for Disease Control and Prevention.
[b] The species identification with the highest probability score based on phenotype matching.
[c] An atypical phenotype is one that is seen in 14% or fewer of the isolates of a given species (Table 3) (6). Results are not given for isolates that did not fall within any of the 12 *Vibrio* species that are pathogenic to humans. Abbreviations for phenotypic tests: Acet, sodium acetate; Cel, cellobiose; Cit, Simmons citrate; Gal, galactose; Gly, glycerol; Ind, indole; Lip, corn oil lipase; Mot, motility; MR, methyl red; ONPG, *o*-nitrophenyl-β-D-galactopyranoside; Orn, ornithine decarboxylase; Oxi, oxidase; Sal, salicin; Suc, sucrose; VPα, Voges-Proskauer. NA, not applicable.

et al. (20) suggested that *V. trachuri* is synonymous with *V. harveyi*. The *V. campbellii* and *V. harveyi* isolates included here had *rpoB* sequences that differed by only two nucleotides. These two isolates differed from the cluster of *"V. trachuri"* and two unknowns by 15 to 19 nucleotides (1.8 to 2.2%), a level of divergence that is similar to that observed between *V. fluvialis* and *V. furnissii* and also between two distinct clusters of *V. vulnificus* (Fig. 2; Table 4). In this case, divergence alone cannot dictate a species boundary; however, based on the phylogeny shown here, principles of phylogenetic classification would require either that *V. harveyi* also include *V. campbellii* to avoid paraphyly in the species or that *V. trachuri* be a separate species along with *V. campbellii* and *V. harveyi*. Further study is needed to resolve the relationships between the isolates in the *V. harveyi* cluster.

The *V. harveyi* cluster also included isolate 9866-84, an isolate that was included because it could not be classified by phenotypic methods. The average genetic distance between 9866-84 and other isolates in the cluster was 4.2%, a value that was consistent with distances between closely related species (*V. cholerae* and *V. mimicus*, 4.2%; *V. parahaemolyticus* and *V. alginolyticus*, 3.2%). We suggest that this isolate is a species closely related to, but separate from, *V. harveyi*.

The *rpoB* sequences of two other clinical isolates (2419-00, from stool, and 2423-01, from blood) were divergent from all other *Vibrio* sequences included here ($d > 13\%$) and did not cluster with any other sequences in the phylogenetic tree. Because our *rpoB* database does not yet include all described *Vibrio* species, we needed additional information to determine if these two isolates represented previously named *Vibrio* species. Thus, we sequenced a portion of the 16S rRNA gene and conducted BLAST searches of GenBank to assess the sequences' relatedness to other *Vibrio* isolates (data not shown). The two isolates did not closely match any named species on the basis of the 16S rRNA gene sequence: the closest named species to isolate 2419-00 was *V. harveyi* (980/986-bp match, 99% similarity) and to isolate 2423-01 was *V. cholerae* (1,515/1,545-bp match, 98% similarity). On the basis of the *rpoB* and 16S rRNA gene sequence analysis, we conclude that these two isolates likely represent unnamed *Vibrio* species.

The two-step approach that we have presented here allows rapid and accurate discrimination of the *Vibrio* species that are implicated in human infections. The advantages of our multiplex are that conserved housekeeping genes are used as a source of markers so that the targeted gene is present in all isolates of a species and multiple species can be discriminated with a single PCR assay. The assay is designed as a rapid classification tool for clinical laboratories, but the markers themselves could be incorporated into other PCR and DNA microarray assays that use virulence genes as markers for pathogenic species (e.g., reference 17). We note that this assay used alone would not assess the pathogenic potential of bacteria in environmental samples, where virulence determination is important in assessing the risk of human illness. We are currently developing molecular probes for the remaining *Vibrio* human pathogens and adapting the assay to a Luminex platform (Luminex Corporation, Austin, TX) to provide a single assay that can rapidly detect and identify the 12 species that are pathogenic to humans.

The *rpoB* gene also provides a useful identification tool; its advantages over the 16S rRNA gene are that it exists as a single-copy gene and it has sufficient phylogenetic signal to discriminate among all of the *Vibrio* species that we have sequenced to date. Thus, our *rpoB* reference database that we are assembling will augment the identification of *Vibrio* species by PCR assays and by sequence-based approaches under way in other laboratories (19).

REFERENCES

1. **Abbott, S. L., L. S. Seli, M. Catino, Jr., M. A. Hartley, and J. M. Janda.** 1998. Misidentification of unusual *Aeromonas* species as members of the genus *Vibrio*: a continuing problem. J. Clin. Microbiol. **36:**1103–1104.
2. **Ait Tayeb, L., E. Ageron, F. Grimont, and P. A. Grimont.** 2005. Molecular phylogeny of the genus *Pseudomonas* based on *rpoB* sequences and application for the identification of isolates. Res. Microbiol. **156:**763–773.
3. **Boyd, E. F., K. E. Moyer, L. Shi, and M. K. Waldor.** 2000. Infectious CTXΦ and the vibrio pathogenicity island prophage in *Vibrio mimicus*: evidence for recent horizontal transfer between *V. mimicus* and *V. cholerae*. Infect. Immun. **68:**1507–1513.
4. **Centers for Disease Control and Prevention.** 2005. Summary of human *Vibrio* isolates reported to CDC, 2004. Centers for Disease Control and Prevention, Atlanta, GA. http://www.cdc.gov/foodborneoutbreaks/vibrio_sum/CSTEVibrio2004.pdf.
5. **Davis, B. R., G. R. Fanning, J. M. Madden, A. G. Steigerwalt, H. B. Bradford, Jr., H. L. Smith, Jr., and D. J. Brenner.** 1981. Characterization of biochemically atypical *Vibrio cholerae* strains and designation of a new pathogenic species, *Vibrio mimicus*. J. Clin. Microbiol. **14:**631–639.
6. **Farmer, J. J. III, J. M. Janda, and K. Birkhead.** 2003. *Vibrio*, p. 706–718. *In* P. R. Murray, E. J. Baron, J. H. Jorgensen, M. A. Pfaller, and R. H. Yolken (ed.), Manual of clinical microbiology, 8th ed., vol. 1. ASM Press, Washington, DC.
7. **Heidelberg, J. F., J. A. Eisen, W. C. Nelson, R. A. Clayton, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, L. Umayam, S. R. Gill, K. E. Nelson, T. D. Read, H. Tettelin, D. Richardson, M. D. Ermolaeva, J. Vamathevan, S. Bass, H. Qin, I. Dragoi, P. Sellers, L. McDonald, T. Utterback, R. D. Fleishmann, W. C. Nierman, O. White, S. L. Salzberg, H. O. Smith, R. R. Colwell, J. J. Mekalanos, J. C. Venter, and C. M. Fraser.** 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. Nature **406:**477–483.
8. **Khamis, A., D. Raoult, and B. La Scola.** 2005. Comparison between *rpoB* and 16S rRNA gene sequencing for molecular identification of 168 clinical isolates of *Corynebacterium*. J. Clin. Microbiol. **43:**1934–1936.
9. **Kim, Y. R., S. E. Lee, C. M. Kim, S. Y. Kim, E. K. Shin, D. H. Shin, S. S. Chung, H. E. Choy, A. Progulske-Fox, J. D. Hillman, M. Handfield, and J. H. Rhee.** 2003. Characterization and pathogenic significance of *Vibrio vulnificus* antigens preferentially expressed in septicemic patients. Infect. Immun. **71:**5461–5471.
10. **Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei.** 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. Bioinformatics **17:**1244–1245.
11. **Kwok, A. Y., J. T. Wilson, M. Coulthart, L. K. Ng, L. Mutharia, and A. W. Chow.** 2002. Phylogenetic study and identification of human pathogenic *Vibrio* species based on partial *hsp60* gene sequences. Can. J. Microbiol. **48:**903–910.
12. **Makino, K., K. Oshima, K. Kurokawa, K. Yokoyama, T. Uda, K. Tagomori, Y. Iijima, M. Najima, M. Nakano, A. Yamashita, Y. Kubota, S. Kimura, T. 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.
13. **Martinez-Urtaza, J., A. Lozano-Leon, A. Vina-Feas, J. de Novoa, and O. Garcia-Martin.** 2006. Differences in the API 20E biochemical patterns of clinical and environmental *Vibrio parahaemolyticus* isolates. FEMS Microbiol. Lett. **255:**75–81.
14. **Mellmann, A., K. Becker, C. von Eiff, U. Keckevoet, P. Schumann, and D. Harmsen.** 2006. Sequencing and staphylococci identification. Emerg. Infect. Dis. **12:**333–336.
15. **Mollet, C., M. Drancourt, and D. Raoult.** 1997. *rpoB* sequence analysis as a novel basis for bacterial identification. Mol. Microbiol. **26:**1005–1011.
16. **O'Hara, C. M., E. G. Sowers, C. A. Bopp, S. B. Duda, and N. A. Strockbine.** 2003. Accuracy of six commercially available systems for identification of members of the family *Vibrionaceae*. J. Clin. Microbiol. **41:**5654–5659.
17. **Panicker, G., D. R. Call, N. J. Krug, and A. K. Bej.** 2004. Detection of pathogenic *Vibrio* spp. in shellfish by using multiplex PCR and DNA microarrays. Appl. Environ. Microbiol. **70:**7436–7444.
18. **Park, T. S., S. H. Oh, E. Y. Lee, T. K. Lee, K. H. Park, M. J. Figueras, and C. L. Chang.** 2003. Misidentification of *Aeromonas veronii* biovar sobria as

Vibrio alginolyticus by the Vitek system. Lett. Appl. Microbiol. **37:**349–353.
19. **Thompson, F. L., D. Gevers, C. C. Thompson, P. Dawyndt, S. Naser, B. Hoste, C. B. Munn, and J. Swings.** 2005. Phylogeny and molecular identification of vibrios on the basis of multilocus sequence analysis. Appl. Environ. Microbiol. **71:**5107–5115.
20. **Thompson, F. L., B. Hoste, K. Vandemeulebroecke, K. Engelbeen, R. Denys, and J. Swings.** 2002. *Vibrio trachuri* Iwamoto et al. 1995 is a junior synonym of *Vibrio harveyi* (Johnson and Shunk 1936) Baumann et al. 1981. Int. J. Syst. Evol. Microbiol. **52:**973–976.
21. **Thompson, F. L., T. Iida, and J. Swings.** 2004. Biodiversity of vibrios. Microbiol. Mol. Biol. Rev. **68:**403–431.
22. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24:**4876–4882.
23. **Wong, R. S., and A. W. Chow.** 2002. Identification of enteric pathogens by heat shock protein 60 kDa (HSP60) gene sequences. FEMS Microbiol. Lett. **206:**107–113.