

## Characterization of clinical and environmental isolates of *Vibrio cidicii* sp. nov., a close relative of *Vibrio navarrensis*

Fabini D. Orata,<sup>1</sup> Yue Xu,<sup>1</sup> Lori M. Gladney,<sup>2,3</sup> Lavanya Rishishwar,<sup>4,5,6</sup> Rebecca J. Case,<sup>1</sup> Yan Boucher,<sup>1</sup> I. King Jordan<sup>4,5,6</sup>† and Cheryl L. Tarr<sup>2</sup>†

### Correspondence

Yan Boucher  
yboucher@ualberta.ca

<sup>1</sup>Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada

<sup>2</sup>Enteric Diseases Laboratory Branch, Centers for Disease Control and Prevention, Atlanta, GA, USA

<sup>3</sup>IHRC, Incorporated, Atlanta, GA, USA

<sup>4</sup>School of Biology, Georgia Institute of Technology, Atlanta, GA, USA

<sup>5</sup>Applied Bioinformatics Laboratory, Atlanta, GA, USA

<sup>6</sup>PanAmerican Bioinformatics Institute, Santiago de Cali, Valle del Cauca, Colombia

Four *Vibrio* spp. isolates from the historical culture collection at the Centers for Disease Control and Prevention, obtained from human blood specimens ( $n=3$ ) and river water ( $n=1$ ), show characteristics distinct from those of isolates of the most closely related species, *Vibrio navarrensis* and *Vibrio vulnificus*, based on phenotypic and genotypic tests. They are specifically adapted to survival in both freshwater and seawater, being able to grow in rich media without added salts as well as salinities above that of seawater. Phenotypically, these isolates resemble *V. navarrensis*, their closest known relative with a validly published name, but the group of isolates is distinguished from *V. navarrensis* by the ability to utilize L-rhamnose. Average nucleotide identity and percent DNA–DNA hybridization values obtained from the pairwise comparisons of whole-genome sequences of these isolates to *V. navarrensis* range from 95.4–95.8% and 61.9–64.3%, respectively, suggesting that the group represents a different species. Phylogenetic analysis of the core genome, including four protein-coding housekeeping genes (*pyrH*, *recA*, *rpoA* and *rpoB*), places these four isolates into their own monophyletic clade, distinct from *V. navarrensis* and *V. vulnificus*. Based on these differences, we propose these isolates represent a novel species of the genus *Vibrio*, for which the name *Vibrio cidicii* sp. nov. is proposed; strain LMG 29267<sup>T</sup> (=CIP 111013<sup>T</sup>=2756-81<sup>T</sup>), isolated from river water, is the type strain.

The genus *Vibrio* consists of over 100 species of bacteria autochthonous to the aquatic environment (Gomez-Gil

*et al.*, 2014). Species of clinical significance, such as *Vibrio cholerae*, *Vibrio parahaemolyticus* and *Vibrio vulnificus*, have

†These authors contributed equally to this work.

**Abbreviations:** ANI, average nucleotide identity; API, Analytical Profile Index; CDC, Centers for Disease Control and Prevention; COG, Clusters of Orthologous Groups of proteins; DDH, DNA-DNA hybridization; GGDC, Genome-to-Genome Distance Calculator; MLSA, multilocus sequence analysis; PM1, Phenotype MicroArray 1; TCBS, thiosulfate citrate bile salts sucrose; TSB, tryptic soy broth.

The GenBank/EMBL/DDBJ accession numbers for the whole-genome sequences of *Vibrio cidicii* LMG 29267<sup>T</sup>, 1048-83, 2423-01 and 2538-88 are LOMK00000000, LOBP00000000, LOBQ00000000 and LOBR00000000, respectively, under BioProject accession number PRJNA304180. Accession numbers for the *rpoB* sequences of the same strains are KU593643, KU593646, KU593645 and KU593644, respectively, and for the *rpoB* sequences of *Vibrio navarrensis* LMG 15976<sup>T</sup>, 2232, 0053-83 and 08-2462 are KU593635, KU593636, KU593629 and KU593637, respectively. The accession number for the 16S rRNA gene sequence of *V. cidicii* LMG 29267<sup>T</sup> is KJ807108.

Seven supplementary tables and two supplementary figures are available with the online Supplementary Material.

been studied in depth, as clear identification of pathogens is recognized as essential for the treatment of the disease and epidemiologic surveillance. *Vibrio navarrensis*, a species not previously associated with human illness, has received little attention since the original description of isolates from sewage in 1991 (Urdaci *et al.*, 1991). However, this species has recently been the focus of investigation after the identification of *V. navarrensis* by the Centers for Disease Control and Prevention (CDC, Atlanta, GA, USA) among clinical isolates submitted for routine characterization (Gladney *et al.*, 2014; Gladney & Tarr, 2014). Constructing an evolutionary framework that included contemporary *V. navarrensis* and phenotypically similar historical isolates led to the discovery of four isolates that could represent a novel species of *Vibrio* closely related to *V. navarrensis*. These isolates, three of which were recovered from human clinical specimens, are genetically distinct from *V. navarrensis* and *V. vulnificus* based on phylogenetic analysis of housekeeping gene sequences (Gladney & Tarr, 2014). The closest known relative of *V. navarrensis* with a validly published name is currently *V. vulnificus* (Gomez-Gil *et al.*, 2014; Thompson *et al.*, 2005). The latter is an opportunistic pathogen mainly associated with deaths related to seafood consumption, and it causes a fatality rate of over 50 % in patients with septicaemia (Jones & Oliver, 2009).

In this study, we employed a polyphasic approach to describe the novel species, such as extensive metabolic profiling of the four isolates, comparative genomic analysis to determine DNA–DNA relatedness, and multilocus sequence analysis (MLSA) of core genes. The name *Vibrio cidicii* sp. nov. is proposed for the new species.

Isolates of *V. cidicii* sp. nov. and *V. navarrensis* used in this study were obtained from the CDC (Table 1). Phenotypic characterization was performed on the four isolates of *V. cidicii* sp. nov. and three isolates of *V. navarrensis* (Tables 2 and S1, available in the online Supplementary Material). The isolates were streaked on tryptic soy broth (TSB; Becton Dickinson) with an added 1.0 % NaCl (BDH), yielding a final concentration of 1.5 % NaCl, and 1.5 % agar (Becton Dickinson). Alternatively, the isolates were streaked on thiosulfate citrate bile salts sucrose (TCBS) agar (Becton Dickinson). The cultures were then incubated overnight (TSB agar) or for two days (TCBS agar) at 30 °C. Single colonies from the TSB agar cultures were tested using the Analytical Profile Index (API) 20 NE (bioMérieux) and the Phenotype MicroArray 1 (PM1) MicroPlate (Biolog) according to the instructions of the manufacturers. A minor modification of the PM1 test was the addition of 1.0 % NaCl to the inoculating fluid (Biolog) to obtain a final concentration of 1.5 % NaCl. The API 20 NE strips and PM1 plates were incubated for 42 or 18 h, respectively, at 30 °C. Additional standard phenotypic tests for the routine identification of *Vibrio* not covered by API 20 NE and PM1 were also performed, including: fermentation using various substrates, citrate (in Simmons agar), DNase (at 25 °C), H<sub>2</sub>S production (in peptone iron agar and triple sugar iron agar), malonate utilization, methyl red, lysine and ornithine decarboxylase (in Moeller medium), motility (swimming and

swarming), ONPG (o-nitrophenyl  $\beta$ -D-galactopyranoside), oxidase, phenylalanine deaminase, tyrosine clearing, and Voges–Proskauer (Farmer *et al.*, 2005; Tarr *et al.*, 2015). Permissive growth temperatures were determined in TSB with a final concentration of 1.5 % NaCl and incubation at a range of 4–45 °C, whereas permissive salinity concentrations were determined in TSB at 30 °C in a range of 0–10 % NaCl. TSB without NaCl was prepared with 17.0 g l<sup>-1</sup> pancreatic digest of casein (Becton Dickinson), 3.0 g l<sup>-1</sup> papaic digest of soybean (Becton Dickinson), 2.5 g l<sup>-1</sup> dextrose (Fisher Scientific) and 2.5 g l<sup>-1</sup> dipotassium phosphate (BDH). Gram staining was conducted on the isolates of *V. cidicii* sp. nov. following the protocol of Claus (1992) and viewing under a light microscope (Carl Zeiss) at  $\times$ 1000 magnification.

For genotypic characterization, genomic DNA was extracted from overnight TSB cultures of the isolates of *V. cidicii* sp. nov. with the ArchivePure DNA Cell/Tissue Kit (5 PRIME). Whole-genome sequencing and assembly of the environmental isolate LMG 29267<sup>T</sup> (=CIP 111013<sup>T</sup>=2756-81<sup>T</sup>) were performed with the PacBio RS and the SMRT (Single-Molecule, Real-Time) Analysis software 2.0 (Pacific Biosciences), respectively, as previously described (Gladney *et al.*, 2014). For the clinical isolates 1048-83 and 2538-88, 150-bp paired-end reads were generated on the MiSeq platform (Illumina), as previously described (Gladney *et al.*, 2014). The genome for clinical isolate 2423-01 was sequenced using 454 sequencing on the Genome Sequencer FLX System (454 Life Sciences) and also on the Genome Analyzer IIX platform (Illumina), generating 70-bp single-end reads. *De novo* assemblies of the sequences from the clinical isolates were performed using the CG-Pipeline 0.4.1 (Kislyuk *et al.*, 2010).

From the *V. cidicii* sp. nov. and *V. navarrensis* whole-genome sequences (Table 1), the G+C content was determined using Geneious 8.1.2 (Kearse *et al.*, 2012). Pairwise average nucleotide identity (ANI) was calculated using the dnadiff program in MUMmer 3.0 (ANIm; Kurtz *et al.*, 2004). Pairwise percent DNA–DNA hybridization (DDH) was also calculated *in silico* using the Genome-to-Genome Distance Calculator 2.0 (GGDC; <http://ggdc.dsmz.de/distcalc2.php>; Meier-Kolthoff *et al.*, 2013). The genome sequences were annotated with RAST 2.0 (Rapid Annotation Using Subsystem Technology; Aziz *et al.*, 2008) and Prodigal 1.2 (Prokaryotic Dynamic Programming Gene-finding Algorithm; Hyatt *et al.*, 2010). Orthologous protein-coding gene families were determined from the annotated genomes by pairwise bidirectional BLASTP (Altschul *et al.*, 1990) using the OrthoMCL pipeline 2.0 (Li *et al.*, 2003) with 30 % identity cut-off (Rost, 1999). The gene families unique to *V. cidicii* sp. nov. or *V. navarrensis* were subsequently determined using Intella 1.7.0 (<https://www.vound-software.com>). The predicted functions of these gene families were determined based on the Clusters of Orthologous Groups of proteins (COG) database (Tatusov *et al.*, 2000) and by sequence similarity search in the GenBank database (National Center for Biotechnology Information) using BLASTP (Altschul *et al.*, 1990).

**Table 1.** Source and year of isolation of the strains of *V. cidicii* sp. nov. and *V. navarrensis* used in this study

Species and strain	Source of isolation	Geographical origin	Year of isolation	Genome accession number*	Reference(s)
<i>Vibrio cidicii</i> sp. nov.					
LMG 29267 <sup>T</sup> (=CIP 111013 <sup>T</sup> = 2756-81 <sup>T</sup> )	River water	Not known	1981	LOMK00000000	This study
1048-83	Human blood	USA	1983	LOBP00000000	This study
2423-01	Human blood	USA	2001	LOBQ00000000	This study
2538-88	Human blood	USA	1988	LOBR00000000	This study
<i>Vibrio navarrensis</i>					
LMG 15976 <sup>T</sup> (=ATCC 51183 <sup>T</sup> =BCRC 15896 <sup>T</sup> =CAIM 609 <sup>T</sup> =CCRC 15896 <sup>T</sup> =CCUG 28805 <sup>T</sup> =CIP 103381 <sup>T</sup> =DSM 21557 <sup>T</sup> =NCIMB 13120 <sup>T</sup> =1397-6 <sup>T</sup> =2540-90 <sup>T</sup> )	Sewage	Spain	1982	JMCG00000000	Gladney <i>et al.</i> (2014); Urdaci <i>et al.</i> (1991)
2232 (=2541-90)	Sewage	Spain	1983	JMCH00000000	This study; Urdaci <i>et al.</i> (1991)
0053-83	Human wound	USA	1983	JMCF00000000	Gladney <i>et al.</i> (2014)
08-2462	Human blood	USA	2008	JMCI00000000	Gladney <i>et al.</i> (2014)

\*BioProject accession numbers: PRJNA304180 (*V. cidicii* sp. nov.) and PRJNA242769 (*V. navarrensis*).

With additional genome sequences from closely related species of the genus *Vibrio* obtained from the GenBank database (Table S2), single-copy, protein-coding core gene families were determined using OrthoMCL 2.0 (Li *et al.*, 2003). The sequences were aligned using ClustalW 2.1 (Larkin *et al.*, 2007), and the alignments were concatenated, stripping columns with at least one gap, using Geneious 8.1.2 (Kearse *et al.*, 2012). This resulted in a single alignment with a total length of 446 032 bp, which was used to reconstruct a maximum-likelihood tree with RAxML 8.2.8 (Stamatakis, 2014) using the GTR (general time reversible) nucleotide substitution model and gamma distribution pattern. Robustness of branching was estimated with 100 bootstrap replicates. Moreover, a subset of four housekeeping genes was selected for MLSA – *pyrH*, *recA*, *rpoA* and *rpoB* (Gladney & Tarr, 2014; Tarr *et al.*, 2007; Thompson *et al.*, 2005). From the partial DNA sequences, a concatenated alignment of 2313 bp was obtained and used to reconstruct a maximum-likelihood tree, as described above. Patristic distances between species, the sum of the lengths of the branches that link two terminal nodes in a tree, were calculated from the latter tree using Geneious 8.1.2 (Kearse *et al.*, 2012). In addition, whole-genome phylogeny was also reconstructed based on genomic similarity (ANI) between each pair of genome sequences. First, pairwise ANIm was computed using MUMmer 3.0 (Kurtz *et al.*, 2004). The similarity obtained was then converted into average nucleotide distances (=100–ANI). The resulting distance matrix was utilized to quantify all pairwise species distances as well as in the reconstruction of a

neighbour-joining tree (Saitou & Nei, 1987) using the software MEGA 7.0 (Kumar *et al.*, 2016).

All the isolates of *V. cidicii* sp. nov. and *V. navarrensis* studied exhibited growth in TSB without NaCl (Tables 2 and S1). This is contrary to a previous report of these isolates not exhibiting growth in nutrient broth without NaCl (Gladney & Tarr, 2014). This is possibly due to differences in the media used in both studies, as test conditions will dictate salt requirement (Farmer *et al.*, 2005). Urdaci *et al.* (1991) reported seven out of ten isolates of *V. navarrensis* grew weakly in peptone water without NaCl. Growth in medium without NaCl was previously reported for a few species of the genus *Vibrio*, including *V. cholerae* and *Vibrio mimicus* (Farmer *et al.*, 2005; Gomez-Gil *et al.*, 2014), two species of great clinical significance. The ability of bacteria to survive in freshwater makes it more likely to come in contact with humans through ingestion (Boucher *et al.*, 2015). The isolation of strains of *V. cidicii* sp. nov. and *V. navarrensis* from river water and sewage, respectively, suggests these species are also able to survive in low salt environments. Furthermore, both species are also able to survive at 40 °C, a trait observed mostly in pathogenic vibrios that can survive inside the human body (Farmer *et al.*, 2005; Gomez-Gil *et al.*, 2014).

*Vibrio cidicii* sp. nov. resembles *V. navarrensis* in the majority of phenotypic characteristics tested (123 of 158 tests or 78 %; Table 2 and Table S1). However, a single phenotypic feature distinguished *V. cidicii* sp. nov. from *V. navarrensis*:

**Table 2.** Summary of phenotypic test results for *V. cidicii* sp. nov., *V. navarrensis* and *V. vulnificus*

Strains: 1, *V. cidicii* sp. nov. LMG 29267<sup>T</sup>; 2, *V. cidicii* sp. nov. 1048-83; 3, *V. cidicii* sp. nov. 2423-01; 4, *V. cidicii* sp. nov. 2538-88; 5, *V. navarrensis* LMG 15976<sup>T</sup>; 6, *V. navarrensis* 0053-83; 7, *V. navarrensis* 08-2462; 8, *V. vulnificus*. +, Growth/positive test result; –, no growth/negative test result; v, variable results between tests; ND, not determined. All strains were positive for indole production, methyl red, phenylalanine deaminase, swimming motility (37 °C), gelatin hydrolysis, aesculin hydrolysis, reduction of nitrate to nitrite, oxidase, DNase (25 °C), acid production from D-glucose, growth in TSB with 6.5 % NaCl (6 % for *V. vulnificus*), and growth in TSB at 40 °C. All strains were negative for Voges–Proskauer, H<sub>2</sub>S production, urea hydrolysis, arginine dihydrolase, gas production from D-glucose, acid production from L-arabinose, glycerol and D-xylose, and growth in TSB at 4 °C.

Phenotypic test/substrate tested	1	2	3	4	5	6	7	8†
Citrate (Simmons agar)	–	+	–	–	+	+	+	+
Lysine decarboxylase (Moeller medium)	–	–	–	–	–	–	–	+
Ornithine decarboxylase (Moeller medium)	–	–	–	–	–	–	–	+
Swarming (marine agar, 25 °C)	–	+	–	+	–	–	–	–
Acid production from:								
L-Rhamnose	+	+	+	+	–	–	–	–
Salicin	–	–	+	+	–	–	–	+
Assimilation of:								
L-Arabinose*	v	v	v	v	v	v	v	–
Glycerol	+	+	+	+	+	+	+	ND
L-Rhamnose	+	+	+	+	–	–	–	–
D-Xylose	+	+	+	+	+	+	+	ND
Growth in TSB (at 30 °C) with:								
0 % NaCl	+	+	+	+	+	+	+	–
1.5 % NaCl	+	+	+	+	+	+	+	– (1 %)
Growth in TSB (with 1.5 % NaCl) at:								
30 °C	+	+	+	+	+	+	+	–

\*Assimilation of L-arabinose (for *V. cidicii* sp. nov. and *V. navarrensis* isolates): positive with PM1, negative with API 20 NE.

†Results for *V. vulnificus* were obtained from Farmer *et al.* (2005) and Gomez-Gil *et al.* (2014).

it tested positive for the utilization of L-rhamnose as the sole carbon and energy source in both fermentation and assimilation tests. Although previous reports support our result of an L-rhamnose-negative *V. navarrensis* (Farmer *et al.*, 2005; Gomez-Gil *et al.*, 2014), a recent study reported one isolate of *V. navarrensis* also capable of utilizing L-rhamnose (Gladney & Tarr, 2014). To our knowledge, this is the only reported isolate of *V. navarrensis* that is L-rhamnose-positive. *Vibrio vulnificus* is also not able to utilize this substrate (Farmer *et al.*, 2005; Gomez-Gil *et al.*, 2014). Other species of the genus *Vibrio* that are capable of utilizing this substrate are *Vibrio hispanicus*, *Vibrio natriegens* and *Vibrio pectenicida*, all of which are very distantly related to *V. cidicii* sp. nov. and *V. navarrensis* (Gomez-Gil *et al.*, 2014; Thompson *et al.*, 2005). Comparison of the annotated genomes of isolates of *V. cidicii* sp. nov. and *V. navarrensis* revealed four genes found only in the former, encoding proteins involved in L-rhamnose transport and metabolism: L-rhamnose isomerase, L-rhamnose mutarotase, L-rhamnose-proton symporter and rhamnulose-1-phosphate aldolase (Table S3; Ryu *et al.*, 2004; Sawada & Takagi, 1964; Wilson & Ajl, 1957). L-Rhamnose is produced in high levels by diatoms (Brown, 1991), a group of phytoplankton found in both marine and freshwater environments, where *V. cidicii* sp.

nov. was also found. This suggests a physiological differentiation with *V. navarrensis*, in which *V. cidicii* sp. nov. is adapted to living on or near algae and exploiting their carbon exudates. An additional 24 genes with predicted functions were found in *V. cidicii* sp. nov. but were absent from *V. navarrensis* (Table S3). These could encode other distinguishing characteristics for the species. However, no physiological tests were available to test the phenotypes they are predicted to encode.

Three additional characteristics distinguish *V. cidicii* sp. nov. from its closest relatives, although not universal among isolates (Tables 2 and S1). First, two of the four isolates of *V. cidicii* sp. nov. exhibited swarming on marine agar. Many vibrios have been reported to exhibit swarming motility, including *Vibrio cincinnatiensis* and *Vibrio proteolyticus*. However, *V. navarrensis* and *V. vulnificus* are negative for the phenotype (Farmer *et al.*, 2005; Gomez-Gil *et al.*, 2014). Second, two of the four isolates of *V. cidicii* sp. nov. tested positive for salicin fermentation, while none of the *V. navarrensis* isolates did. This trait is also a key differentiating characteristic between *V. navarrensis* and *V. vulnificus*, with the latter also able to utilize the substrate (Farmer *et al.*, 2005). Lastly, three of the four isolates of *V. cidicii* sp. nov. are negative for the utilization of sodium citrate (citrate

test), whereas both *V. navarrensis* and *V. vulnificus* are able to utilize the substrate (Farmer *et al.*, 2005). Both *V. cidicii* sp. nov. and *V. navarrensis* can be differentiated from *V. vulnificus* by two characteristics; they test negative for lysine and ornithine decarboxylase (Farmer *et al.*, 2005; Gomez-Gil *et al.*, 2014).

We observed contradicting results between the PM1 and API 20 NE tests we conducted for the assimilation of L-arabinose, where all isolates of *V. cidicii* sp. nov. and *V. navarrensis* tested positive with the PM1 system and negative with the API 20 NE system. This difference can be attributed to the differences in methods between the tests (e.g. incubation period, NADH production and redox dye chemistry detection versus turbidity detection). We conclude that this test is not reliable for the identification of the species *V. cidicii* sp. nov. and *V. navarrensis*. Additionally, we observed a difference between our assimilation and fermentation tests (positive for all isolates with the former and negative with the latter) with two other substrates, glycerol and D-xylose. This difference is due to the different attributes being measured (i.e. NADH production versus acid production), suggesting that both species are capable of utilizing the substrates aerobically but not via fermentation.

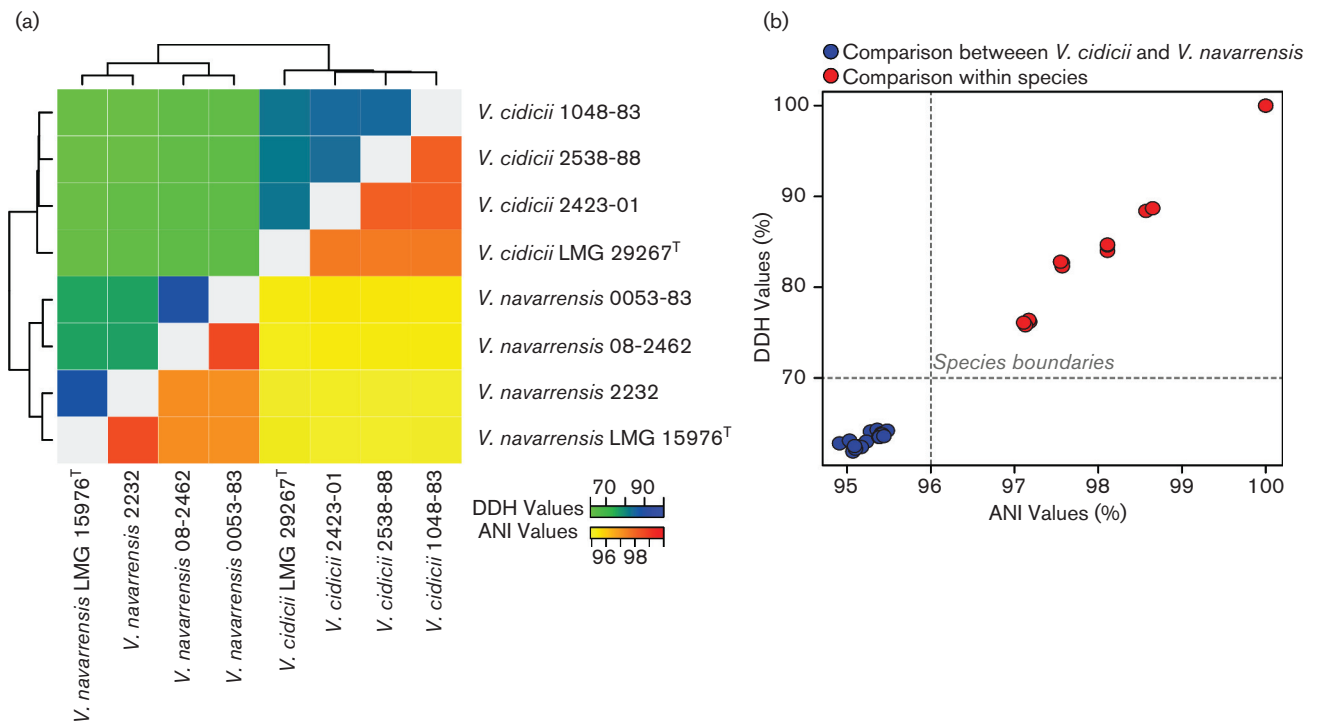
Based on their whole-genome sequences, the G+C content of the four isolates of *V. cidicii* sp. nov. range from 47.9–48.2 mol%, which is within the known range for the genus *Vibrio* (38.0–51.0 mol%; Farmer *et al.*, 2005). This eliminates the assignment of the isolates to other genera in the family *Vibrionaceae* such as *Aliivibrio* (38.0–42.0 mol%; Urbanczyk *et al.*, 2007), *Photobacterium* (39.0–44.0 mol%; Thyssen & Ollevier, 2005), and *Salinivibrio* (49.4–50.5 mol%; Ventosa, 2005), and most genera in the family *Enterobacteriaceae* (50.0–67.0 mol%; Brenner & Farmer, 2005).

Various tools to measure DNA–DNA relatedness *in silico* are available to replace the traditional method of DDH (Goris *et al.*, 2007; Konstantinidis & Tiedje, 2005; Meier-Kolthoff *et al.*, 2013; Richter & Rosselló-Móra, 2009). Here, we determined relatedness of organisms using ANI and DDH by pairwise comparisons of whole-genome sequences. The ANI between isolates within the species *V. cidicii* sp. nov. or *V. navarrensis* range from 97.4–100.0% (Fig. 1 and Table S4). In contrast, the ANI between *V. cidicii* sp. nov. and *V. navarrensis* range from 95.4–95.8%. Since the results are close to the cut-off of 96% ANI for two genomes to belong to the same species (Richter & Rosselló-Móra, 2009), we complemented our ANI results with percent DDH data. The GGDC package was used to calculate percent DDH *in silico* to mimic wet lab-based DDH (Meier-Kolthoff *et al.*, 2013). Percent DDH within the species *V. cidicii* sp. nov. or *V. navarrensis* range from 75.8–88.7%, whereas they range from 61.9–64.3% between the two groups (Fig. 1 and Table S5). The determined same-species cut-off for DDH is 70% (Goris *et al.*, 2007), supporting our

ANI results and suggesting the two groups to be distinct from each other.

MLSA further supports our proposal of a novel species. Single-copy, protein-coding core genes are used as alternatives to 16S rRNA gene sequences for the identification and phylogenetic analysis of various species of the genus *Vibrio*, since there is a lack of species-level resolution using 16S rRNA gene sequences (Gladney & Tarr, 2014; Thompson *et al.*, 2005). A core genome tree was reconstructed from 586 single-copy core genes that are shared by all strains used in this study (Vernikos *et al.*, 2015). The four isolates of *V. cidicii* sp. nov. form a monophyletic clade that is distinct from the *V. navarrensis* and *V. vulnificus* clades, with 100% bootstrap support (Fig. S1). Since recombination is also apparent within the core genome and can occur at a high rate for very closely related species (Orata *et al.*, 2015), we examined a subset of four housekeeping genes (*pyrH*, *recA*, *rpoA* and *rpoB*) that do not exhibit recombination among the isolates of *V. cidicii* sp. nov., *V. navarrensis* and *V. vulnificus*. These genes have been shown to be reliable for the taxonomic characterization of vibrios (Gladney & Tarr, 2014; Tarr *et al.*, 2007; Thompson *et al.*, 2005). Phylogenetic analysis using the four housekeeping genes also distinguishes *V. cidicii* sp. nov. from *V. navarrensis* and *V. vulnificus* (Fig. 2). The average patristic distance calculated from this tree between the *V. cidicii* sp. nov. isolates and the *V. navarrensis* isolates is 0.066, while lower average distances of 0.005 and 0.007 are obtained when comparing isolates within the species *V. cidicii* sp. nov. or *V. navarrensis*, respectively (Table S6). To further demonstrate this distinction, a phylogeny was reconstructed based on whole-genome comparisons to account for whole-genome variation between isolates that would otherwise be excluded from the core genome (Fig. S2). This phylogeny also shows the distinct clustering of the isolates of *V. cidicii* sp. nov. from *V. navarrensis*. The average nucleotide distances calculated for this tree (Table S7) show that the diversity within *V. cidicii* sp. nov. (1.780) or *V. navarrensis* (2.160) is much lower than the diversity between the two species (4.414). Our phylogenetic analyses placed the *V. cidicii* sp. nov. lineage into the context of a larger *Vibrio* phylogeny, showing that the novel species is distinct from all *Vibrionaceae* that have been characterized to date (Figs 2, S1 and S2; Gladney & Tarr, 2014). On the other hand, the use of the 16S rRNA gene did not clearly distinguish *V. cidicii* sp. nov. and *V. navarrensis* (Gladney & Tarr, 2014).

Overall, phylogenetic analyses confirm the position of *V. cidicii* sp. nov. in the genus *Vibrio*, which forms a monophyletic clade distinct from *V. navarrensis* and *V. vulnificus*, supporting its identification as a novel species of the genus *Vibrio*. This distinction is further confirmed by ANI and percent DDH below 96% and 70%, respectively, between species. The ability of *V. cidicii* sp. nov. to utilize L-rhamnose could be a feature that drove its speciation from a common ancestor shared with *V. navarrensis*. Further studies are needed to determine the prevalence of *V.*



**Fig. 1.** Whole-genome comparisons among isolates of *V. cidicii* sp. nov. and *V. navarrensis*. (a) Heat map showing percent DDH and ANI along with the clustering-based dendrogram showing the relationship between *V. cidicii* sp. nov. and *V. navarrensis* genomes. (b) Pairwise percent DDH and ANI comparisons between isolates of *V. cidicii* sp. nov. and *V. navarrensis* below the species boundary cut-offs indicating that they are different species.

*cidicii* sp. nov. in various environments. The isolation of strains from human blood suggests it is capable of infecting humans and can be pathogenic. Two additional isolates were recovered since 2014 after the commencement of this study, underscoring the need to identify and characterize isolates of this pathogen. It will be important to study the pathogenicity and epidemiology of this novel species for control, treatment and prevention of disease.

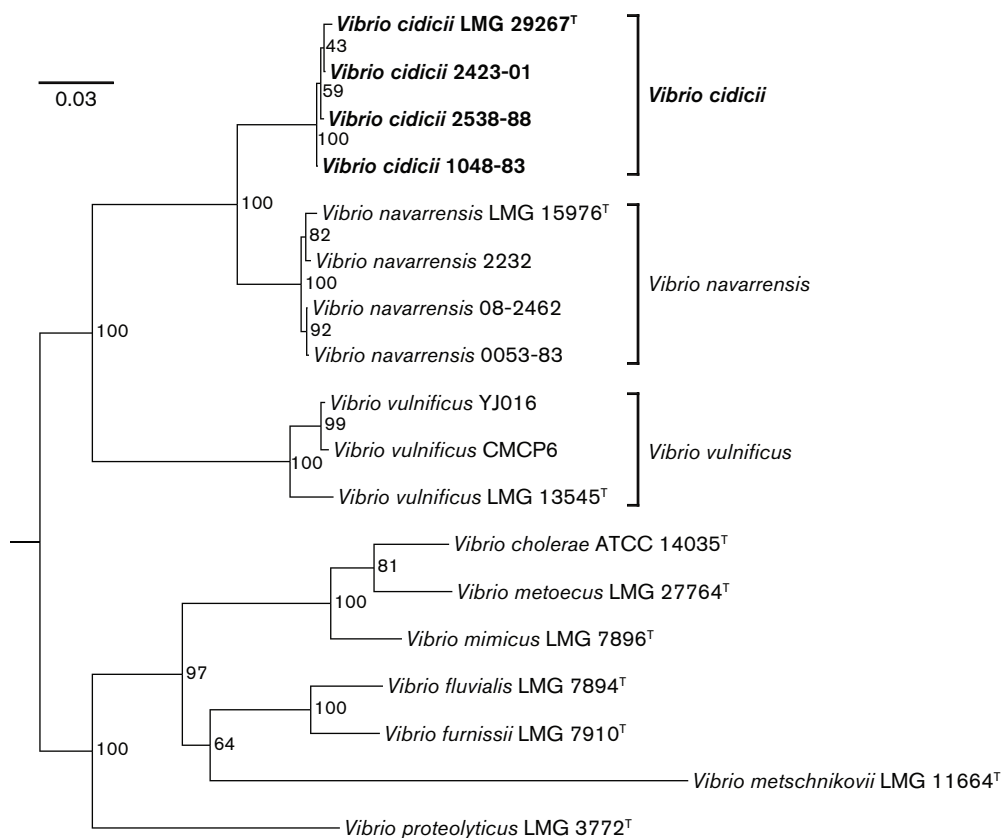
### Description of *Vibrio cidicii* sp. nov.

*Vibrio cidicii* (ci.di'ci.i. N.L. gen. n. *cidicii* from the phonetics of the CDC, the acronym for the Centers for Disease Control and Prevention, where the species was initially identified).

Cells are Gram-negative, curved, motile rods, 0.64–0.78 × 1.48–1.68 μm in size, which produce convex, smooth, circular, entire, cream colonies on TSB agar and yellow colonies (sucrose-fermenting) on TCBS agar. Growth is observed in TSB at 30 °C with salt concentrations in the range of 0–6.5 % NaCl, and up to 8 % for some isolates (three out of four tested); no growth occurs in the presence of 10 % NaCl. Growth is also observed in TSB with 1.5 % total NaCl concentration at a temperature range of 30–40 °C, and no growth occurs at 4 °C and 45 °C. The ability to utilize L-rhamnose as the sole carbon and energy source

distinguishes *V. cidicii* sp. nov. from *V. navarrensis*, its closest relative. In addition, the following characteristics are variable across isolates: swarming on marine agar (two positive out of four tested; negative for *V. navarrensis*), salicin fermentation (two positive out of four tested; negative for *V. navarrensis*), and sodium citrate utilization (three negative out of four tested; positive for *V. navarrensis*). Positive results in tests for: indole production; methyl red test; phenylalanine deaminase; gelatin and aesculin hydrolysis; reduction of nitrate to nitrite; oxidase and DNase; and acid production from D-glucose, cellobiose, maltose, mannitol, mannose, L-rhamnose, sucrose and trehalose. Negative results in tests for: Voges–Proskauer reaction; H<sub>2</sub>S production; urea hydrolysis; arginine dihydrolase, lysine decarboxylase and ornithine decarboxylase; malonate utilization; gas production from D-glucose; and acid production from D-adonitol, L-arabinose, D-arabitol, dulcitol, erythritol, glycerol, *myo*-inositol, lactose, melibiose, mucic acid, raffinose and D-xylose.

Utilizes the following substrates as sole carbon and energy sources: acetic acid, acetoacetic acid, *N*-acetyl-D-glucosamine, adenosine, D-alanine, L-alanine, L-alanyl glycine, L-asparagine, L-aspartic acid, bromosuccinic acid, cellobiose, citric acid, 2'-deoxyadenosine, deoxyribonucleic acid, aesculin, formic acid, D-fructose, fructose 6-phosphate,



**Fig. 2.** Phylogenetic relationship of *V. cidicii* sp. nov. and its closest relatives. The tree is reconstructed from the concatenated alignment of partial DNA sequences of four protein-coding housekeeping genes (*pyrH*, *recA*, *rpoA* and *rpoB*) with a total length of 2313 bp. Bootstrap support is indicated on the nodes. Bar, 0.03 nucleotide substitutions per site.

fumaric acid, gelatin, D-gluconic acid, D-glucose, glucose 1-phosphate, glucose 6-phosphate, L-glutamic acid, L-glutamine, glycerol, DL- $\alpha$ -glycerol phosphate, glycyl L-aspartic acid, glycyl L-glutamic acid, glycyl L-proline,  $\alpha$ -hydroxybutyric acid,  $\alpha$ -hydroxyglutaric acid- $\gamma$ -lactone, inosine,  $\alpha$ -ketobutyric acid,  $\alpha$ -ketoglutaric acid, L-lactic acid, L-lyxose, D-malic acid, L-malic acid, DL-malic acid, D-maltose, maltotriose, D-mannitol, D-mannose, methyl pyruvate, methyl  $\beta$ -D-glucoside, monomethyl succinate, phenylalanine, potassium gluconate, L-proline, propionic acid, D-psiocose, pyruvic acid, L-rhamnose, D-ribose, L-serine, succinic acid, sucrose, L-threonine, thymidine, trehalose, L-tryptophan, Tween 40, Tween 80, uridine and D-xylose. The following substrates are not utilized (in four tested isolates): adipic acid, 2-aminoethanol, D-arabitol, L-arginine, D-aspartic acid, capric acid, erythritol, L-fucose, L-galactonic acid- $\gamma$ -lactone, D-galacturonic acid, glucuronamide, D-glucosaminic acid, glyoxylic acid, m-hydroxyphenylacetic acid, p-hydroxyphenylacetic acid, *myo*-inositol, lysine, malonate,  $\alpha$ -methyl-D-galactoside, ornithine, phenylacetic acid, phenylethylamine, 1,2-propanediol, raffinose, D-serine, m-tartaric acid, D-threonine, tricarballic acid, trisodium citrate, tyramine, and urea.

The type strain is LMG 29267<sup>T</sup> (=CIP 111013<sup>T</sup>=2756-81<sup>T</sup>) isolated from river water in 1981 (country of origin unknown). The type strain displays all of the properties given above for the species. The G+C content of the type strain is 47.9 mol% based on whole-genome sequencing.

## Acknowledgements

F. D. O. was supported by Alberta Innovates – Technology Futures. R. J. C. and Y. B. were supported by the Natural Sciences and Engineering Research Council of Canada. Y. B. was also supported by the Canadian Institute for Advanced Research. L. R. and I. K. J. were supported by the Georgia Institute of Technology (GIT) Bioinformatics Graduate Program and the IHRC-GIT Applied Bioinformatics Laboratory.

## References

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* 215, 403–410.
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formosa, K., Gerdes, S., Glass, E. M. & other authors (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9, 75.

- Boucher, Y., Orata, F. D. & Alam, M. (2015).** The out-of-the-delta hypothesis: dense human populations in low-lying river deltas served as agents for the evolution of a deadly pathogen. *Front Microbiol* **6**, 1120.
- Brenner, D. J. & Farmer, J. J., III. (2005).** Family I. *Enterobacteriaceae*. In *Bergey's Manual of Systematic Bacteriology—Volume Two: The Proteobacteria, Part B: The Gammaproteobacteria*, pp. 587–850. Edited by D. J. Brenner, N. R. Krieg, J. T. Staley, G. M. Garrity, D. R. Boone, P. De Vos, F. A. Rainey, M. Goodfellow & K. H. Schleifer. New York, NY: Springer.
- Brown, M. R. (1991).** The amino-acid and sugar composition of 16 species of microalgae used in mariculture. *J Exp Mar Biol Ecol* **145**, 79–99.
- Claus, D. (1992).** A standardized Gram staining procedure. *World J Microbiol Biotechnol* **8**, 451–452.
- Farmer, J. J., III, Janda, M., Brenner, F. W., Cameron, D. N. & Birkhead, K. M. (2005).** Genus I. *Vibrio*. In *Bergey's Manual of Systematic Bacteriology—Volume Two: The Proteobacteria, Part B: The Gammaproteobacteria*, pp. 494–546. Edited by D. J. Brenner, N. R. Krieg, J. T. Staley, G. M. Garrity, D. R. Boone, P. De Vos, M. Goodfellow, F. A. Rainey & K. H. Schleifer. New York, NY: Springer.
- Gladney, L. M. & Tarr, C. L. (2014).** Molecular and phenotypic characterization of *Vibrio navarrensis* isolates associated with human illness. *J Clin Microbiol* **52**, 4070–4074.
- Gladney, L. M., Katz, L. S., Knipe, K. M., Rowe, L. A., Conley, A. B., Rishishwar, L., Mariño-Ramírez, L., Jordan, I. K. & Tarr, C. L. (2014).** Genome sequences of *Vibrio navarrensis*, a potential human pathogen. *Genome Announc* **2**, e01188–14.
- Gomez-Gil, B., Thompson, C. C., Matsumura, Y., Sawabe, T., Iida, T., Christen, R., Thompson, F. & Sawabe, T. (2014).** The family *Vibrionaceae*. In *The Prokaryotes—Gammaproteobacteria*, pp. 659–747. Edited by E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt & F. Thompson. Berlin, Germany: Springer.
- Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P. & Tiedje, J. M. (2007).** DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**, 81–91.
- Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W. & Hauser, L. J. (2010).** Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119.
- Jones, M. K. & Oliver, J. D. (2009).** *Vibrio vulnificus*: disease and pathogenesis. *Infect Immun* **77**, 1723–1733.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S. & other authors (2012).** Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649.
- Kislyuk, A. O., Katz, L. S., Agrawal, S., Hagen, M. S., Conley, A. B., Jayaraman, P., Nelakuditi, V., Humphrey, J. C., Sammons, S. A. & other authors (2010).** A computational genomics pipeline for prokaryotic sequencing projects. *Bioinformatics* **26**, 1819–1826.
- Konstantinidis, K. T. & Tiedje, J. M. (2005).** Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* **102**, 2567–2572.
- Kumar, S., Stecher, G. & Tamura, K. (2016).** MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* **33**, 1870–1874.
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C. & Salzberg, S. L. (2004).** Versatile and open software for comparing large genomes. *Genome Biol* **5**, R12.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A. & other authors (2007).** CLUSTAL W and CLUSTAL X version 2.0. *Bioinformatics* **23**, 2947–2948.
- Li, L., Stoeckert, C. J. & Roos, D. S. (2003).** OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* **13**, 2178–2189.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P. & Göker, M. (2013).** Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**, 60.
- Orata, F. D., Kirchberger, P. C., Méheust, R., Barlow, E. J., Tarr, C. L. & Boucher, Y. (2015).** The dynamics of genetic interactions between *Vibrio metoecus* and *Vibrio cholerae*, two close relatives co-occurring in the environment. *Genome Biol Evol* **7**, 2941–2954.
- Richter, M. & Rosselló-Móra, R. (2009).** Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* **106**, 19126–19131.
- Rost, B. (1999).** Twilight zone of protein sequence alignments. *Protein Eng* **12**, 85–94.
- Ryu, K. S., Kim, C., Kim, I., Yoo, S., Choi, B. S. & Park, C. (2004).** NMR application probes a novel and ubiquitous family of enzymes that alter monosaccharide configuration. *J Biol Chem* **279**, 25544–25548.
- Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Sawada, H. & Takagi, Y. (1964).** The metabolism of L-rhamnose in *Escherichia coli*: III. L-rhamnulose-phosphate aldolase. *Biochim Biophys Acta* **92**, 26–32.
- Stamatakis, A. (2014).** RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.
- Tarr, C. L., Patel, J. S., Pühr, N. D., Sowers, E. G., Bopp, C. A. & Strockbine, N. A. (2007).** Identification of *Vibrio* isolates by a multiplex PCR assay and *rpoB* sequence determination. *J Clin Microbiol* **45**, 134–140.
- Tarr, C. L., Bopp, C. A. & Farmer, J. J., III. (2015).** *Vibrio* and related organisms. In *Manual of Clinical Microbiology*, pp. 762–772. Edited by J. H. Jorgensen, M. A. Pfaller, K. C. Carroll, G. Funke, M. L. Landry, S. S. Richter & D. W. Warnock. Washington, DC: ASM Press.
- Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V. (2000).** The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* **28**, 33–36.
- Thompson, F. L., Gevers, D., Thompson, C. C., Dawyndt, P., Naser, S., Hoste, B., Munn, C. B. & Swings, J. (2005).** Phylogeny and molecular identification of vibrios on the basis of multilocus sequence analysis. *Appl Environ Microbiol* **71**, 5107–5115.
- Thyssen, A. & Ollevier, F. (2005).** Genus II. *Photobacterium*. In *Bergey's Manual of Systematic Bacteriology—Volume Two: The Proteobacteria, Part B: The Gammaproteobacteria*, pp. 546–552. Edited by D. J. Brenner, N. R. Krieg, J. T. Staley, G. M. Garrity, D. R. Boone, P. De Vos, M. Goodfellow, F. A. Rainey & K. H. Schleifer. New York, NY: Springer.
- Urbanczyk, H., Ast, J. C., Higgins, M. J., Carson, J. & Dunlap, P. V. (2007).** Reclassification of *Vibrio fischeri*, *Vibrio logei*, *Vibrio salmonicida* and *Vibrio wodanis* as *Aliivibrio fischeri* gen. nov., comb. nov., *Aliivibrio logei* comb. nov., *Aliivibrio salmonicida* comb. nov. and *Aliivibrio wodanis* comb. nov. *Int J Syst Evol Microbiol* **57**, 2823–2829.
- Urdaci, M. C., Marchand, M., Ageron, E., Arcos, J. M., Sesma, B. & Grimont, P. A. (1991).** *Vibrio navarrensis* sp. nov., a species from sewage. *Int J Syst Bacteriol* **41**, 290–294.
- Ventosa, A. (2005).** Genus III. *Salinivibrio*. In *Bergey's Manual of Systematic Bacteriology—Volume Two: The Proteobacteria, Part B: The Gammaproteobacteria*, pp. 552–555. Edited by D. J. Brenner, N. R. Krieg, J. T. Staley, G. M. Garrity, D. R. Boone, P. De Vos, M. Goodfellow, F. A. Rainey & K. H. Schleifer. New York, NY: Springer.
- Vernikos, G., Medini, D., Riley, D. R. & Tettelin, H. (2015).** Ten years of pan-genome analyses. *Curr Opin Microbiol* **23**, 148–154.
- Wilson, D. M. & Aji, S. (1957).** Metabolism of L-rhamnose by *Escherichia coli*. I. L-rhamnose isomerase. *J Bacteriol* **73**, 410–414.