



Comparative Genomic Analysis of *Vibrio diabolicus* and Six Taxonomic Synonyms: A First Look at the Distribution and Diversity of the Expanded Species

Jeffrey W. Turner^{1*}, James J. Tallman¹, Amanda Macias¹, Lee J. Pinnell¹, Nicole C. Elledge¹, Danial Nasr Azadani¹, William B. Nilsson², Rohinee N. Paranjpye², E. V. Armbrust³ and Mark S. Strom²

OPEN ACCESS

Edited by:

Frank T. Robb,
University of Maryland, Baltimore,
United States

Reviewed by:

Charles Lovell,
University of South Carolina,
United States
Iddya Karunasagar,
Nitte University, India
Christopher John Grim,
United States Food and Drug
Administration, United States

*Correspondence:

Jeffrey W. Turner
jeffrey.turner@tamucc.edu

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 11 April 2018

Accepted: 27 July 2018

Published: 15 August 2018

Citation:

Turner JW, Tallman JJ, Macias A,
Pinnell LJ, Elledge NC,
Nasr Azadani D, Nilsson WB,
Paranjpye RN, Armbrust EV and
Strom MS (2018) Comparative
Genomic Analysis of *Vibrio diabolicus*
and Six Taxonomic Synonyms: A First
Look at the Distribution and Diversity
of the Expanded Species.
Front. Microbiol. 9:1893.
doi: 10.3389/fmicb.2018.01893

¹ Department of Life Sciences, Texas A&M University-Corpus Christi, Corpus Christi, TX, United States, ² Division of Environmental and Fisheries Sciences, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA, United States, ³ Center for Environmental Genomics, School of Oceanography, University of Washington, Seattle, WA, United States

Vibrio is a diverse genus of Gammaproteobacteria autochthonous to marine environments worldwide. *Vibrio diabolicus* and *V. antiquarius* were originally isolated from deep-sea hydrothermal fields in the East Pacific Rise. These species are closely related to members of the Harveyi clade (e.g., *V. alginolyticus* and *V. parahaemolyticus*) that are commonly isolated from coastal systems. This study reports the discovery and draft genome sequence of a novel isolate (*Vibrio* sp. 939) cultured from Pacific oysters (*Crassostrea gigas*). Questions surrounding the identity of *Vibrio* sp. 939 motivated a genome-scale taxonomic analysis of the Harveyi clade. A 49-genome phylogeny based on 1,109 conserved coding sequences and a comparison of average nucleotide identity (ANI) values revealed a clear case of synonymy between *Vibrio* sp. 939, *V. diabolicus* Art-Gut C1 and CNCM I-1629, *V. antiquarius* EX25 and four *V. alginolyticus* strains (E0666, FF273, TS13, and V2). This discovery expands the *V. diabolicus* species and makes available six additional genomes for comparative genomic analyses. The distribution of the expanded species is thought to be global given the range of isolation sources (horse mackerel, seawater, sediment, dentex, oyster, artemia and polychaete) and origins (China, India, Greece, United States, East Pacific Rise, and Chile). A subsequent comparative genomic analysis of this new eight-genome subclade revealed a high degree of individual genome plasticity and a large repertoire of genes related to virulence and defense. These findings represent a significant revision to the understanding of *V. diabolicus* and *V. antiquarius* as both have long been regarded as distinct species. This first look at the expanded *V. diabolicus* subclade suggests that the distribution and diversity of this species mirrors that of other Harveyi clade species, which are notable for their ubiquity and diversity.

Keywords: *Vibrio diabolicus*, *Vibrio antiquarius*, *Vibrio alginolyticus*, deep-sea, taxonomy, genomics, diversity

INTRODUCTION

Vibrio is a diverse genus of Gram-negative bacteria (Thompson et al., 2004) comprising some 139 species¹. Since Filippo Pacini first isolated *Vibrio cholerae* in 1854 (Pacini, 1854), the genus has been studied extensively as many species are readily culturable and at least 12 species are human pathogens (Chakraborty et al., 1997). *V. cholerae* and *V. parahaemolyticus* are notable in that both have been associated with pandemic disease (Karaolis et al., 1995; Matsumoto et al., 2000), and *V. vulnificus* is notable for its ability to cause primary sepsis in immunocompromised persons (Strom and Paranjpye, 2000). The genus also includes species associated with disease in commercially important fisheries and foundational species. For example, *V. harveyi* and *V. anguillarum* are pathogens of penaeid shrimp and fish, respectively (Austin and Zhang, 2006; Frans et al., 2011), and *V. coralliilyticus* and *V. shiloi* are pathogens of corals (Rosenberg et al., 2007). Yet the genus is conditionally rare (Shade et al., 2014) and many species are benign (Thompson and Polz, 2006; Thompson et al., 2006).

The geographic range of the *Vibrio* genus is extensive; representative species are commonly found in riverine, estuarine and coastal marine ecosystems worldwide (Thompson et al., 2004). Sea surface temperature can be a limiting factor, restricting growth to warmer latitudes, as many species are mesophilic (Takemura et al., 2014). However, two *Vibrio* species were originally isolated from habitats far removed from coastal ecosystems and the coincidental human host. *V. diabolicus* (strain CNCM I-1629) was first isolated from a polychaete associated with a deep-sea hydrothermal vent in the East Pacific Rise (Raguénès et al., 1997) where temperatures would support mesophilic growth. Similarly, *V. antiquarius* (strain EX25) was first isolated from seawater associated with a sulfide chimney in the East Pacific Rise (Hasan et al., 2015).

Previous reports have proposed that *V. diabolicus* CNCM I-1629 and *V. antiquarius* EX25 harbor specific genes that promote survival and persistence in hydrothermal vent ecosystems. For example, *V. diabolicus* CNCM I-1629 produces an acidic exopolysaccharide (EPS) proposed to confer resistance to high concentrations of the metallic sulfides (e.g., pyrite, chalcopyrite and sphalerite) associated with sulfide chimneys (Raguénès et al., 1997; Goudenège et al., 2014), and the *V. antiquarius* EX25 genome encodes an alkyl hydroperoxide reductase proposed to scavenge endogenous hydrogen peroxide in the deep-sea (Hasan et al., 2015) as this enzyme is also found in the deep-sea tubeworm endosymbiont *Riftia pachyptila* (Markert et al., 2007). Paradoxically, these two species also harbor hemolysins and type III secretion systems (T3SS) that are commonly regarded as virulence-associated factors in pathogenic *Vibrio* species (Goudenège et al., 2014; Hasan et al., 2015). The maintenance of these features in the deep-sea has been cited as evidence that secreted toxins are evolutionarily ancient features that may play a larger role in environmental fitness (Hasan et al., 2015).

What remains unknown is the distribution, diversity and potential virulence of these deep-sea *Vibrio* species. It is possible

that both *V. diabolicus* and *V. antiquarius* exhibit a cosmopolitan distribution. A more global distribution would complement the adaptability of closely related *Vibrio* species (e.g., *V. alginolyticus* and *V. parahaemolyticus*), which have been characterized as opportunitrophs as they are capable of exploiting resources that are both spatially and temporally limited (Polz et al., 2006). Indeed, the detection of *V. antiquarius* EX25 open reading frames (ORFs) in 89 shotgun metagenomic datasets (Hasan et al., 2015) and the recent isolation of twelve *V. diabolicus* isolates in a mid-Atlantic estuary (North Inlet-Winyah Bay National Estuarine Research Reserve, SC, United States) (Klein et al., 2014) supports the hypothesis that these deep-sea *Vibrio* are more widely distributed.

During the course of a long-term, multifaceted study of *V. parahaemolyticus* (Johnson et al., 2012; Paranjpye et al., 2012; Turner et al., 2013, 2016), *Vibrio* sp. 939 was isolated from Pacific oysters (*Crassostrea gigas*) harvested from Puget Sound in the Pacific Northwest region of the United States. This study describes the discovery and draft genome sequencing of *Vibrio* sp. 939 and details efforts to identify the isolate through a comprehensive taxonomic analysis of 49 Harveyi clade genomes. Findings clearly show that *V. antiquarius* and *V. diabolicus* constitute the same species and we recommend using *V. diabolicus* as the antecedent. Findings also show that four misidentified *V. alginolyticus* strains (E0666, FF273, TS13, and V2) are synonyms of *V. diabolicus*. Moreover, the following comparative genomic analysis answers important questions about the expanded species' distribution and diversity.

MATERIALS AND METHODS

Sample Collection

Vibrio sp. 939 was isolated from Pacific oysters (*C. gigas*) harvested from Hood Canal, which is a basin of Puget Sound in Washington State (United States). The isolate was one of hundreds collected by the NOAA Northwest Fisheries Science Center (Seattle, WA, United States) during a multifaceted *V. parahaemolyticus* study (Johnson et al., 2012; Paranjpye et al., 2012; Turner et al., 2013, 2016). Briefly, oysters were scrubbed, shucked and homogenized, and presumptive *Vibrio* species were isolated by direct plating on thiosulfate-citrate-bile salts-sucrose (TCBS) agar (Becton, Dickinson and Company, Franklin Lakes, NJ, United States) overnight at 30°C. *Vibrio* sp. 939 was presumed to be *V. parahaemolyticus* based on the detection of the species-associated thermolabile hemolysin (*tlh*) gene (Bej et al., 1999), but the isolate was omitted from our previous *V. parahaemolyticus* multilocus sequence typing (MLST) study (Turner et al., 2013) as many of the MLST loci proved recalcitrant to amplification. The preliminary assignment was also called into question by two recent studies showing that the *V. parahaemolyticus*-specific *tlh* primers could also amplify sequence variants of the *tlh* gene in closely related species (Klein et al., 2014; Yáñez et al., 2015).

Culture Conditions

Starting with a frozen cell culture (25% glycerol, v/v, -80°C), *Vibrio* sp. 939 was streaked for isolation on lysogeny broth (LB)

¹<http://www.bacterio.net/vibrio.html>

(Fisher Scientific, Fair Lawn, NJ, United States) supplemented with 1.5% Bacto agar (Becton, Dickinson and Company, Franklin Lakes, NJ, United States) and grown overnight (18 h) at 30°C. A single isolated colony was transferred to 5 mL LB in a 15 mL BD Falcon tube (Becton, Dickinson and Company, Franklin Lakes, NJ, United States) and grown overnight (18 h) at 30°C with shaking (120 rpm) in an Excella E24 shaking incubator (Eppendorf, Hamburg, Germany).

DNA Isolation

Bacteria in the overnight culture (1 mL) were pelleted by centrifugation ($9,400 \times g$, 5 min) in an Eppendorf 5424E centrifuge (Hamburg, Germany) and washed twice with an equal volume of phosphate buffered saline (PBS). The DNA was isolated from the pelleted cells using a ChargeSwitch gDNA Mini Bacteria Kit (Invitrogen, Carlsbad, CA, United States) following the manufacturer's protocols. The DNA was quantified and assayed for quality (A_{260}/A_{280}) using a BioPhotometer D30 (Eppendorf, Hamburg, Germany) and stored at -20°C .

Genome Sequencing

Genomic DNA was sequenced with an Illumina MiSeq instrument at the New York University Genome Technology Center using paired-end chemistry (2×300 bp). The sequencing library was prepared using the PCR-free version of a KAPA DNA Library Preparation Kit (Kapa Biosystems, Wilmington, MA, United States). Overlapping paired reads were merged using FLASH version 1.2.11 (Magoc and Salzberg, 2011). The merged reads were trimmed of adapter sequences and low quality bases with Trim Galore! version 0.4.4², which is a wrapper script for Cutadapt (Martin, 2011) and FastQC (Andrews, 2010). The optimal k -mer size was estimated with KmerGenie version 1.7 (Chikhi and Medvedev, 2014) and the draft genome was assembled *de novo* with Velvet version 1.2.10 (Zerbino and Birney, 2008). The assembled genome was initially annotated and inspected with the web-based RAST annotation service and SEED Viewer (Aziz et al., 2008; Overbeek et al., 2014). The final annotation was completed with the National Center for Biotechnology Information's (NCBI) Prokaryotic Genome Annotation Pipeline (PGAP) (Klimke et al., 2009).

Phylogenetics

The relatedness of *Vibrio* sp. 939 and 48 closely related Harveyi clade species (e.g., *V. diabollicus*, *V. antiquarius*, *V. parahaemolyticus*, *V. alginolyticus*, and *V. harveyi*) (Supplementary Table S1) was inferred by constructing a maximum-likelihood tree using a set of single-copy homologs present in all the 49 Harveyi clade genomes. The genomes were downloaded from NCBI. Single-copy homologous genes were clustered using *get_homologues* (options $-M -t 49 -r \text{EX25.gbk} -e$) (Contreras-Moreira and Vinuesa, 2013). The program options dictated that the homolog search was carried out with the OrthoMCL algorithm (Li et al., 2003) using the default E -value cutoff ($1e-05$) and that clustering was limited to single-copy homologs present in all 49 genomes. Additionally, the closed

genome of *V. antiquarius* EX25 was used as the reference genome and clusters containing in paralogs were excluded. Homologous clusters were aligned individually with MUSCLE version 3.6 (Edgar, 2004) and alignments were trimmed to the length of the shortest sequence with trimAl version 1.4.15 (Capella-Gutierrez et al., 2009). Trimmed alignments were then filtered by length (250 bp cutoff), concatenated with the fasta manipulation tool in Galaxy (Giardine et al., 2005) and a maximum-likelihood tree was constructed with IQ-TREE version 1.5.5 (Nguyen et al., 2014) with 1,000 ultrafast bootstraps (Minh et al., 2013) using the best-fit model (GTR+I+G4) as determined by ModelFinder (Kalyaanamoorthy et al., 2017). The phylogenetic tree was annotated with FigTree version 1.4.3³.

Phylogenetic Network

To visualize conflicting phylogenetic signal owing to reticulate evolutionary processes, a Neighbor-Net phylogenetic network (Bryant and Moulton, 2004) was constructed for the concatenated multilocus alignment (above) with SplitsTree4 version 4.14.4 (Huson and Bryant, 2006) using default settings and 1,000 bootstraps.

Genome Similarity

The genomes of *V. diabollicus* Art-Gut C1 and CNCM I-1629, *Vibrio* sp. 939, *V. antiquarius* EX25 and *V. alginolyticus* 12G01, 40B, E0666, FF273, K01M1, NBRC 15630, TS13, and V2 were compared with an all-versus-all alignment using JSpeciesWS (Richter et al., 2016). JSpeciesWS is a web server implementation of the popular JSpecies Taxonomic Thresholds Program (Richter and Rosselló-Móra, 2009) that calculates the average nucleotide identity (ANI) between a query genome and a reference genome where 95–96% similarity is considered the threshold for identifying a prokaryotic species. To illustrate the sequence similarity of the eight *V. diabollicus* subclade genomes (defined in this study as *V. diabollicus* Art-Gut C1 and CNCM I-1629, *Vibrio* sp. 939, *V. antiquarius* EX25, and *V. alginolyticus* E0666, FF273, TS13, and V2), a circular blast map was constructed using the BLAST Ring Image Generator (BRIG) (Alikhan et al., 2011). For this purpose, the lower and upper identity thresholds were set at 92 and 96% (respectively) and the closed genome of *V. antiquarius* EX25 was used as the reference genome. For added contrast, the *V. alginolyticus* NBRC 15630 genome was included in the comparison. Genome attributes (e.g., genome size, %GC content, total genes, protein coding genes, RNA genes) were also calculated to compare these eight genomes.

Pangenome Analysis

To assess the diversity of the *V. diabollicus* subclade (defined above), the pangenome of this subclade was determined by two methods. First, the pangenome was estimated with *get_homologues* using option $t = 0$ to report all homologous genes clusters and the accompanying *parse_pangenome_matrix.pl* script was used to call the subset of genes that constitute the core (genes present in all eight strains) and accessory (genes present in fewer than eight strains) pangenome compartments. Seeing

²<https://github.com/FelixKrueger/TrimGalore>

³<http://tree.bio.ed.ac.uk/software/figtree/>

that the *get_homologues* analysis excludes clusters with fewer than three sequences as well as non-coding sequences, a second pangenome analysis was carried out with Panseq (Laing et al., 2010) using default parameters. The Panseq software estimates the pangenome using an all-versus-all alignment of genomic regions and reports the core, accessory and pangenome. For this analysis, the core genome was defined as the genomic regions present in all eight genomes. Additionally, to capture the presence of singletons (i.e., genomic regions present in only one genome), Panseq's Novel Region Finder was used to discover genomic regions novel to individual genomes. The novel regions were then filtered by length (10 Kb cutoff) to identify probable genomic islands (GIs). These GIs were then inspected in the web-based SEED Viewer (blastn *E*-value cutoff $1e-05$) (Overbeek et al., 2014) for features linked with survival or persistence in specific habitats (e.g., deep-sea versus coastal).

Subclade-Associated Genes

The output from the above pangenome analysis was parsed with the *parse_pangenome_matrix.pl* script to identify genes ubiquitous in the *V. diabollicus* subclade ($N = 8$ strains) but absent from the other 41 strains. Subclade-associated genes and their protein products were aligned individually with MUSCLE version 3.6 (Edgar, 2004) and the average sequence identity for each alignment was calculated with trimAl version 1.4.15 (Capella-Gutierrez et al., 2009).

Metagenome Survey

To investigate the presence of *V. diabollicus* in publicly available environmental metagenomic data, the protein sequences encoded by the most conserved non-hypothetical subclade-associated genes (above) were queried against the NCBI Metagenome Protein Database (*env_nr*) using default blastp parameters. Database hits were declared significant if the sequence identity was comparable to the above subclade protein alignments (i.e., greater than 95%).

Potential Virulence

The occurrence and diversity of genes related to virulence and defense was investigated using SEED subsystem feature counts implemented in the web-based SEED Viewer (above). For this purpose, the eight genomes belonging to the *V. diabollicus* subclade were inspected for proteins with significant matches (blastp with *E*-value cutoff $1e-05$) to five subsystems: (1) Adhesion, (2) Toxins and Superantigens, (3) Bacteriocins and Ribosomally Synthesized Antibacterial Peptides, (4) Resistance to Antibiotics and Toxic Compounds and (5) Invasion and Intracellular Resistance. To specifically investigate enzymes conferring resistance to antibiotics, the β -lactamase proteins were aligned with MUSCLE version 3.6 (Edgar, 2004), trimmed with trimAl version 1.4.15 (Capella-Gutierrez et al., 2009) and a maximum-likelihood tree was constructed with IQ-TREE version 1.5.5 (Nguyen et al., 2014) with 1,000 ultrafast bootstraps (Minh et al., 2013) using the best-fit model (WAG) as determined by ModelFinder (Kalyaanamoorthy et al., 2017). Additionally, the β -lactamase proteins (present in all genomes) and the fosfomycin resistance proteins (present in *V. alginolyticus* TS13 and V2)

were searched against the Antibiotic Resistance Genes Database (ARDB) using default parameters (blastp with *E*-value cutoff $1e-05$) (Liu and Pop, 2009).

Antibiotic Susceptibility

The production of β -lactamase was tested using Cefinase disks (Becton, Dickinson and Company, Franklin Lakes, NJ, United States) according to the manufacturer's instructions. Susceptibility to select β -lactam antibiotics was determined as described previously (Letchumanan et al., 2015). *V. antequarius* 939 was grown overnight (18 h) on Mueller-Hinton agar plates (Becton, Dickinson and Company, Franklin Lakes, NJ, United States) (supplemented with 1.5% NaCl) at 37°C. Overnight growth was resuspended in a filter-sterilized 0.85% NaCl solution, normalized to approximate a 0.5 McFarland standard, and fresh Mueller-Hinton agar plates were seeded with a bacterial lawn using a sterile cotton swab. Penicillin (10 units), ampicillin (10 μ g), cephalothin (30 μ g), and carbenicillin (100 μ g) antibiotic disks (Becton, Dickinson and Company, Franklin Lakes, NJ, United States) were dispensed on the lawn and plates were incubated overnight (18 h) at 37°C. The zones of inhibition were measured and interpreted according to the Clinical and Laboratory Standards Institute (CLSI) guidelines (Clinical and Laboratory Standards Institute [CLSI], 2010). The pandemic type strain of *V. parahaemolyticus* (RIMD2210633) (Nasu et al., 2000) was tested for comparative purposes.

RESULTS

Genome Sequencing

The draft genome of *Vibrio* sp. 939 was comprised of 46 contigs totaling 5,430,661 bp in length with a 44.6% GC content. The N50 of the assembly was 855,346 bp and the lengths of the smallest and largest contigs were 570 and 1,035,854 bp, respectively. The PGAP annotation detected 5,049 genes, 4,876 proteins, 112 RNAs (13 rRNA, 95 tRNA, and 4 other RNA) and 61 pseudogenes. Based on the initial evaluation of the draft genome in the SEED Viewer (the View Closest Neighbors feature), the whole-genome shotgun project was deposited as *V. antequarius* 939 at DDBJ/ENA/GenBank under the accession number AOJB00000000.

Phylogenetic Tree

To construct a phylogeny of the Harveyi clade, a search for homologous genes returned a total of 1,125 gene clusters shared by all 49 genomes. Sixteen of the clusters were removed by the 250 bp length filter to yield a total of 1,109 single-copy homologs. The concatenation of the homologs produced a multilocus alignment totaling 1,034,083 bp. The concatenated alignment was comprised of 595,014 constant sites and 404,946 parsimony informative sites with 217,816 distinct site patterns. The best-fit model according to the Bayesian information criterion (BIC) scores and weights was calculated as GTR+I+G4.

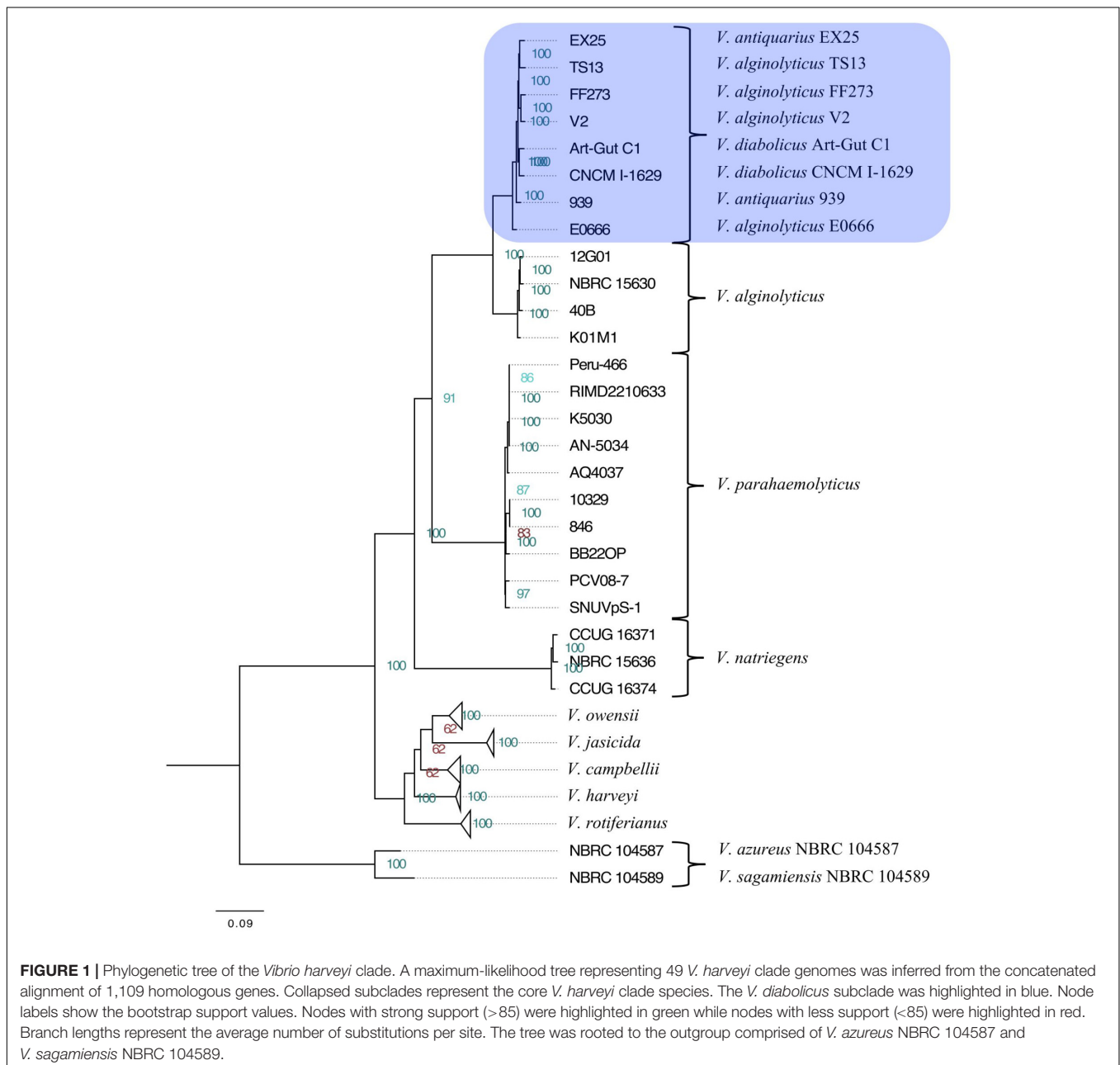
The maximum-likelihood phylogeny resolved the Harveyi clade into eleven subclades: *V. alginolyticus*, *V. azureus*, *V. campbellii*, *V. harveyi*, *V. jasicida*, *V. natriegens*, *V. owensii*,

V. parahaemolyticus, *V. rotiferianus*, *V. sagamiensis*, and a mixed subclade comprised of two *V. diabolicus* strains (Art-Gut C1 and CNCM I-1629), two *V. antiquarius* strains (939 and EX25) and four *V. alginolyticus* strains (E0666, FF273, TS13, and V2) (**Figure 1**). The entirety of the tree, without collapsed branches, was included in the **Supplementary Figure S1**. This phylogeny showed strong agreement with previous whole-genome Harveyi clade phylogenies (Thompson et al., 2009; Lin et al., 2010; Urbanczyk et al., 2013, 2014, 2016; Ke et al., 2017), but unlike previous phylogenies, the aim of this phylogeny was the resolution of the non-core Harveyi subclades (e.g., *V. alginolyticus*, *V. antiquarius*, *V. diabolicus*, and *V. parahaemolyticus*). To that end, the discovery of

a well-supported *V. diabolicus* subclade comprised of two *V. diabolicus* strains (Art-Gut C1 and CNCM I-1629), two *V. antiquarius* strains (939 and EX25) and four *V. alginolyticus* strains (E0666, FF273, TS13, and V2) was novel.

Phylogenetic Network

To infer the role of reticulate evolutionary processes, a Neighbor-Net network was constructed using the same 1,109 multilocus alignment (above). The network (**Figure 2**) showed complete agreement with the phylogenetic tree. A more reticulate network was evident between the *V. diabolicus* and *V. alginolyticus* subclades. Regardless, the bootstrap analysis of the phylogenetic network showed strong support for a distinct



V. diabolica subclade comprised of two *V. diabolica* strains (Art-Gut C1 and CNCM I-1629), two *V. antiquarius* strains (939 and EX25) and four *V. alginolyticus* strains (E0666, FF273, TS13, and V2).

Genome Similarity

The relatedness of the *V. diabolica* subclade genomes and representative genomes from its nearest phylogenetic neighbor (*V. alginolyticus*) was further analyzed by calculating pairwise ANI values (Table 1). Results showed that the eight genomes assigned to the *V. diabolica* subclade (*V. diabolica* Art-Gut C1 and CNCM I-1629, *V. antiquarius* 939 and EX25, and *V. alginolyticus* E0666, FF273, TS13, and V2) shared greater than 97% ANI. In contrast, members of this subclade shared less than 92% ANI with the representative *V. alginolyticus* genomes (i.e., *V. alginolyticus* 12G01, 40B, K0M1, and NBRC 15630).

To visualize the synonymy in ANI values between the eight *V. diabolica* subclade genomes, a circular blastn map was constructed (Figure 3). The map clearly illustrated the high degree of synonymy between the two *V. diabolica* genomes (Art-Gut C1 and CNCM I-1629), the *V. antiquarius* 939 genome, the four *V. alginolyticus* genomes (E0666, FF273, TS13, and V2)

and the reference genome (*V. antiquarius* EX25). In contrast, the non-synonymy between these genomes and *V. alginolyticus* NBRC 15630 was apparent. Notable hypervariable genomic regions labeled on the map include a capsular polysaccharide region, a super integron, a type VI secretion system (T6SS) and a glucuronic acid utilization cluster.

A more general comparison of the eight *V. diabolica* subclade isolates (Supplementary Table S1) revealed that they were isolated from a range of sources (i.e., horse mackerel, seawater, sediment, dentex, oyster, artemia, and polychaete) collected from locations around the world (i.e., China, India, Greece, United States, East Pacific Rise, and Chile). The genome sizes ranged from 5,048,917 to 5,430,661 bp (*V. alginolyticus* E0666 and *V. antiquarius* 939, respectively) while the number of genes ranged from 4,572 to 5,251 (*V. alginolyticus* FF273 and *V. diabolica* CNCM I-1629, respectively), and the GC content ranged from 44.6 to 44.9% (*V. alginolyticus* FF273 and *V. antiquarius* EX25, respectively) (Table 2).

Genome Diversity

To assess the genome diversity of the *V. diabolica* subclade, the pangenome of the eight *V. diabolica* subclade genomes

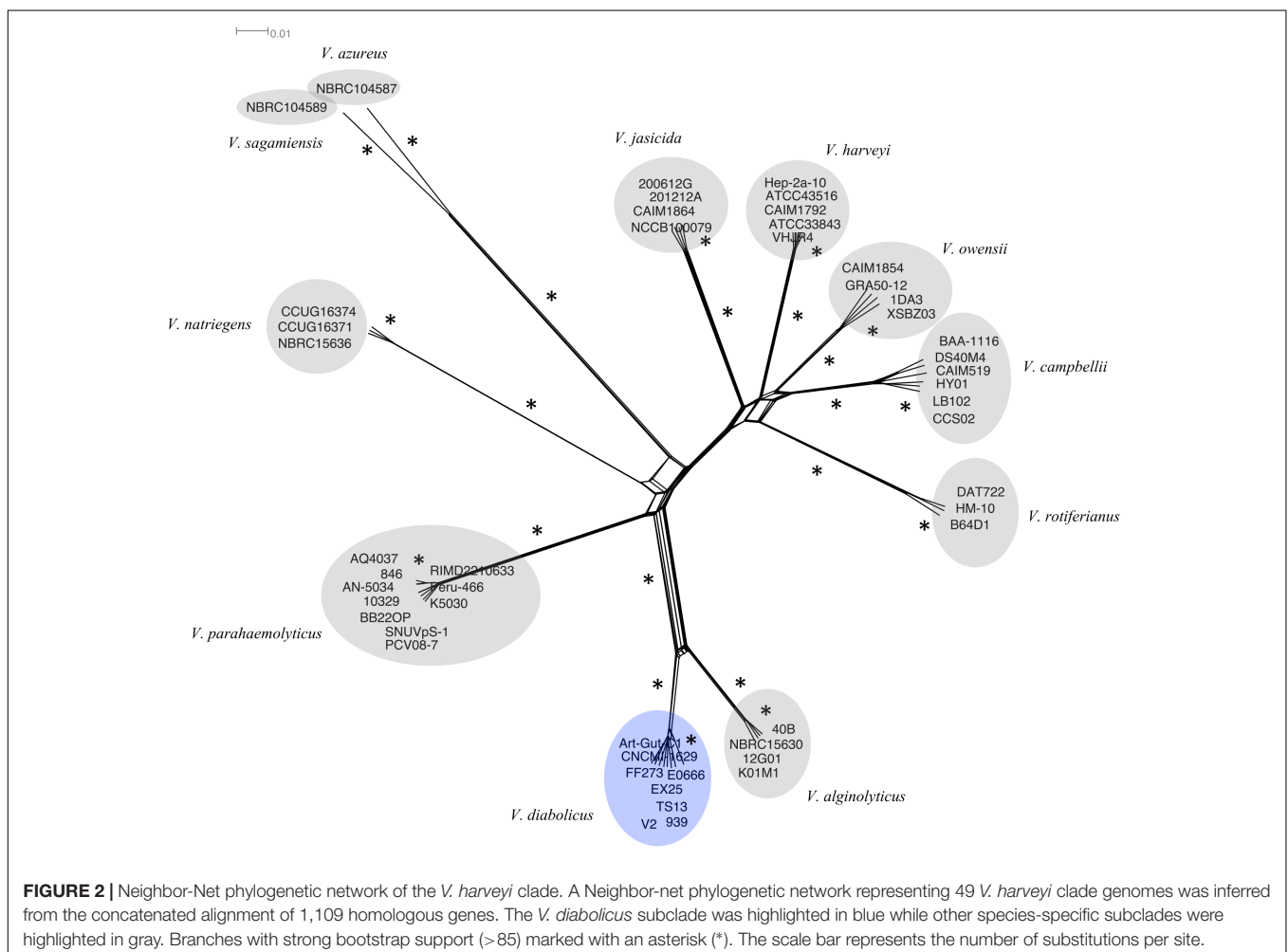


TABLE 1 | The pairwise comparison of average nucleotide identity (ANI) values between *V. diabollicus*, *V. antiquarius*, and *V. alginolyticus* subclade genomes.

Genome	1	2	3	4	5	6	7	8	9	10	11	12
(1) <i>V. diabollicus</i> Art-Gut C1	*	97.76	97.71	97.83	97.48	97.50	97.75	97.59	91.28	91.49	91.44	91.28
(2) <i>V. diabollicus</i> CNCM I-1629	97.89	*	97.66	97.88	97.62	97.68	97.75	97.67	91.38	91.47	91.50	91.47
(3) <i>V. antiquarius</i> 939	97.65	97.52	*	97.72	97.39	97.48	97.64	97.44	91.27	91.41	91.43	91.33
(4) <i>V. antiquarius</i> EX25	97.91	97.93	97.80	*	97.65	97.80	97.83	97.74	91.45	91.52	91.62	91.52
(5) <i>V. alginolyticus</i> E0666	97.6	97.67	97.59	97.68	*	97.49	97.52	97.50	91.68	91.88	91.79	91.82
(6) <i>V. alginolyticus</i> FF273	97.53	97.63	97.58	97.69	97.39	*	97.54	98.08	91.62	91.63	91.69	91.66
(7) <i>V. alginolyticus</i> TS13	97.71	97.67	97.62	97.71	97.37	97.49	*	97.40	91.39	91.50	91.47	91.48
(8) <i>V. alginolyticus</i> V2	97.68	97.68	97.57	97.74	97.44	98.17	97.54	*	91.62	91.67	91.69	91.71
(9) <i>V. alginolyticus</i> 12G01	91.42	91.46	91.38	91.43	91.74	91.69	91.48	91.69	*	98.34	98.33	98.52
(10) <i>V. alginolyticus</i> 40B	91.59	91.53	91.49	91.47	91.86	91.77	91.57	91.72	98.26	*	98.21	98.31
(11) <i>V. alginolyticus</i> KOM1	91.57	91.58	91.57	91.56	91.77	91.80	91.59	91.76	98.25	98.21	*	98.26
(12) <i>V. alginolyticus</i> NBRC 15630	91.43	91.52	91.42	91.48	91.85	91.80	91.59	91.79	98.48	98.30	98.27	*

The values in bold font exceed the 95–96% threshold for species assignment. The dashed lines separate the *V. diabollicus* and *V. alginolyticus* subclades. The * indicates 100% similarity as the genomes were compared against themselves.

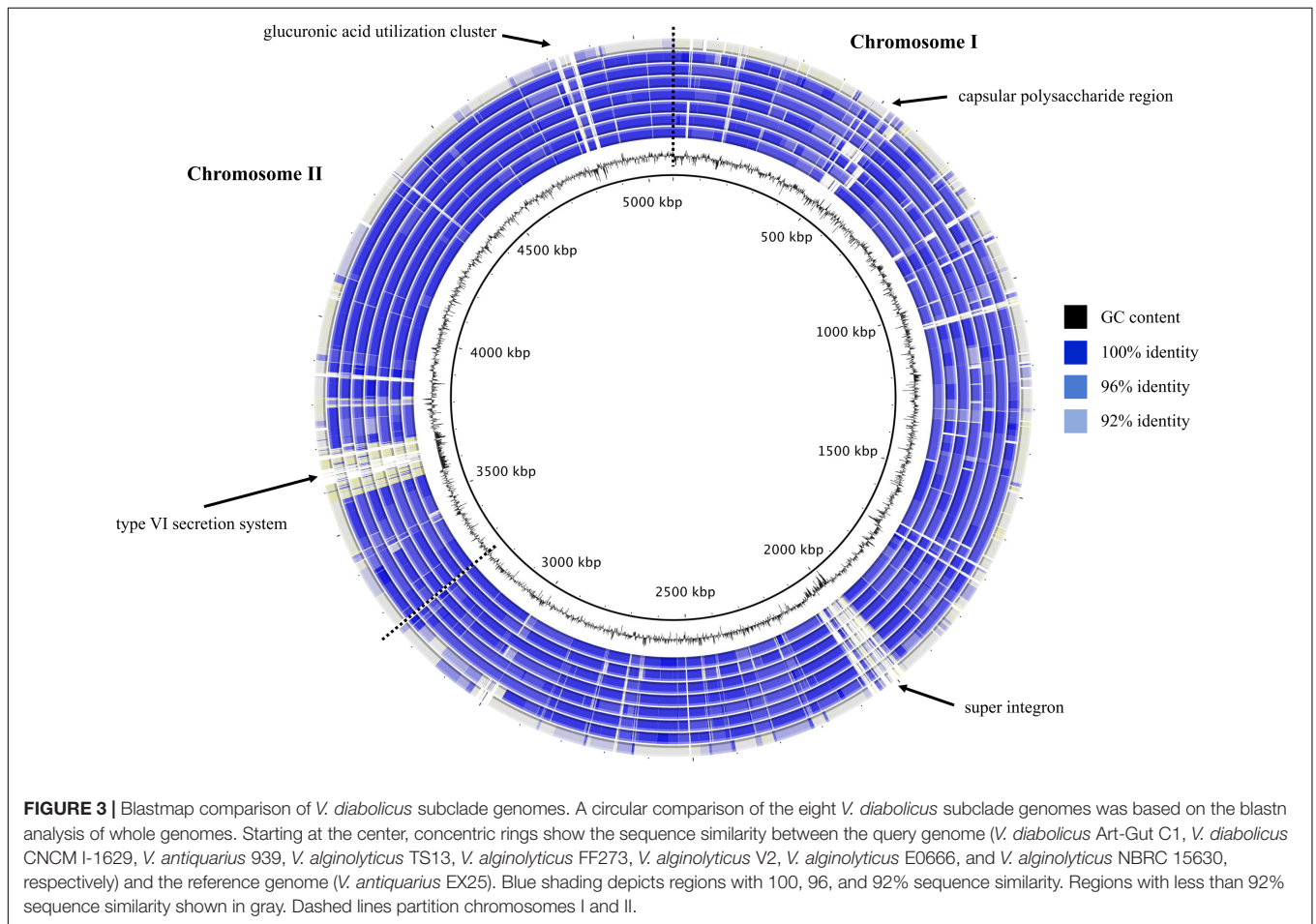


FIGURE 3 | Blastmap comparison of *V. diabollicus* subclade genomes. A circular comparison of the eight *V. diabollicus* subclade genomes was based on the blast analysis of whole genomes. Starting at the center, concentric rings show the sequence similarity between the query genome (*V. diabollicus* Art-Gut C1, *V. diabollicus* CNCM I-1629, *V. antiquarius* 939, *V. alginolyticus* TS13, *V. alginolyticus* FF273, *V. alginolyticus* V2, *V. alginolyticus* E0666, and *V. alginolyticus* NBRC 15630, respectively) and the reference genome (*V. antiquarius* EX25). Blue shading depicts regions with 100, 96, and 92% sequence similarity. Regions with less than 92% sequence similarity shown in gray. Dashed lines partition chromosomes I and II.

(*V. diabollicus* Art-Gut C1 and CNCM I-1629, *V. antiquarius* 939 and EX25, and *V. alginolyticus* E0666, FF273, TS13, and V2) was determined by two methods: the clustering of homologous genes with *get_homologues* and the aligning of homologous genomic regions with Panseq. The *get_homologues* method reported 4,854 total and 3,422 core genes shared by all eight

genomes. Thus, the pangenome consisted of 3,422 core genes (~70.5% of the pangenome) and 1,422 accessory genes (~29.5% of the pangenome). The Panseq method reported a 7,856,358 bp pangenome comprised of a 4,605,287 bp core genome (~58.6% of the pangenome) and a 3,251,071 bp accessory genome (~41.4% of the pangenome).

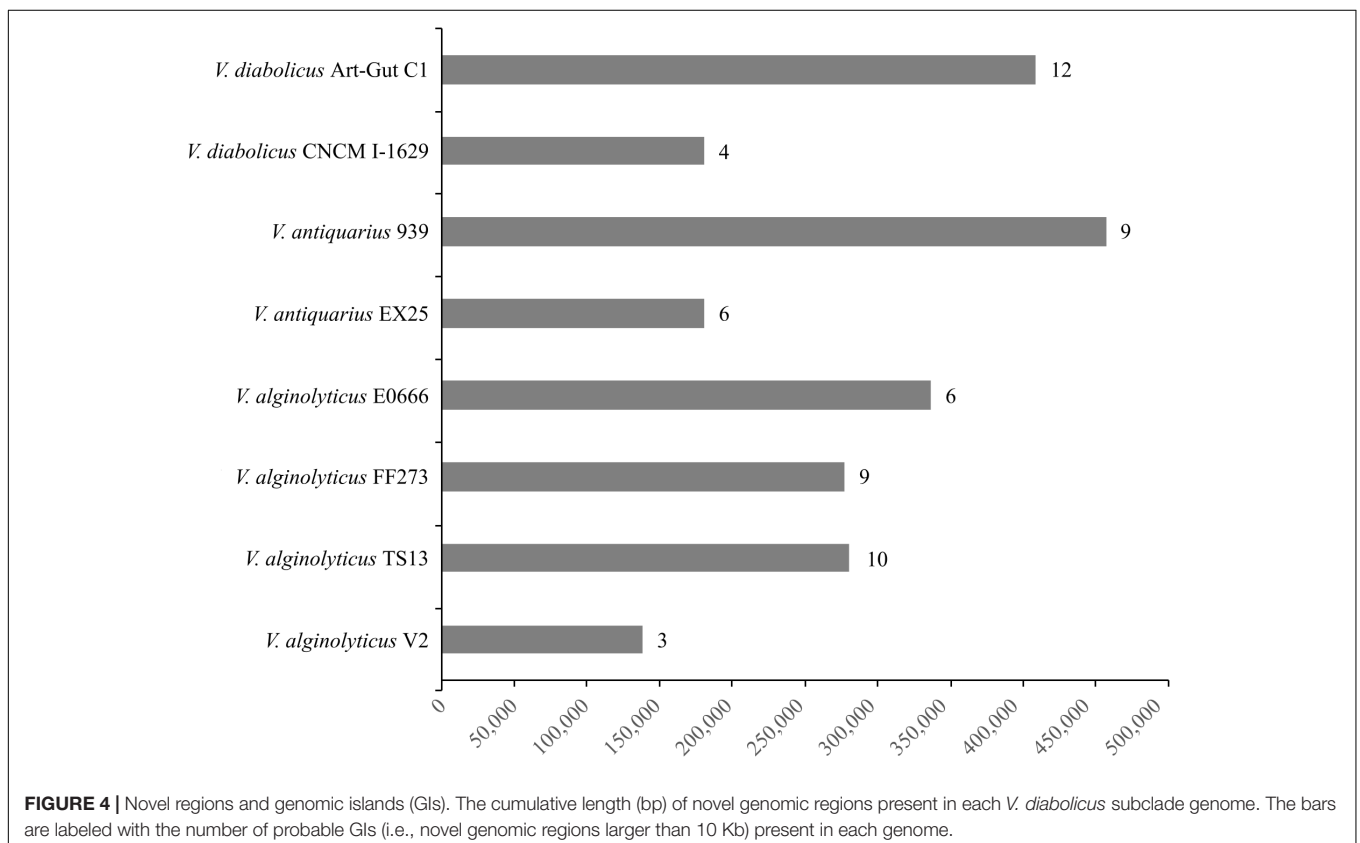
TABLE 2 | A summary of the attributes of the *V. diabolicus* subclade genomes: *V. diabolicus* Art-Gut C1 and CNCM I-1629, *V. antiquarius* 939 and EX25, and *V. alginolyticus* E0666, FF273, TS13 and V2.

Attribute	Art-Gut C1	CNCM I-1629	939	EX25	E0666	FF273	TS13	V2
Genome size (bp)	5,236,997	5,132,517	5,430,661	5,089,025	5,048,917	5,231,685	5,103,549	5,068,299
Level ^a	Draft	Draft	Draft	Closed	Draft	Draft	Draft	Draft
Contigs	60	29	46	2	129	179	40	33
% GC	44.8	44.8	44.6	44.9	44.8	44.6	44.7	44.8
Genes	4,843	5,251	5,049	4,778	4,616	4,705	4,858	4,572
Proteins	4,593	4,855	4,876	4,545	4,407	4,574	4,659	4,448
Pseudogenes	150	297	61	71	110	104	122	42
RNAs	100	99	112	162	99	23	77	82

^aLevel indicates the completeness of the genome assembly.

The Panseq Novel Region Finder revealed that novel genomic regions totaling 2,295,535 bp accounted for the vast majority of the accessory genome: *V. diabolicus* Art-Gut C1 (415,117 bp), *V. diabolicus* CNCM I-1629 (183,383 bp), *V. antiquarius* 939 (464,398 bp), *V. antiquarius* EX25 (183,450 bp), *V. alginolyticus* E0666 (341,836 bp), *V. alginolyticus* FF273 (281,065 bp), *V. alginolyticus* TS13 (285,061 bp), and *V. alginolyticus* V2 (140,215 bp) (Figure 4). Further, among the novel genomic regions, probable genomic islands (GIs) greater than 10 Kb in length were prevalent: *V. diabolicus* Art-Gut C1 (12 GIs), *V. diabolicus* CNCM I-1629 (4 GIs), *V. antiquarius* 939 (9 GIs), *V. antiquarius* EX25 (6 GIs), *V. alginolyticus* E0666 (6 GIs), *V. alginolyticus* FF273 (9 GIs), *V. alginolyticus* TS13 (10 GIs), and

V. alginolyticus V2 (3 GIs) (Figure 4). These GIs encoded a variety of proteins including hypothetical proteins, capsular polysaccharide biosynthesis proteins, lipopolysaccharide biosynthesis proteins, proteins related to carbohydrate and DNA metabolism, membrane transport proteins, phage-related proteins (phage baseplate proteins, tail proteins, tail fiber proteins, capsid proteins, endolysins, and integrases) and conjugative transfer proteins. However, these novel regions and GIs did not encode proteins with obvious habitat-specific functions. Rather, these novel regions encoded features thought to promote fitness in a range of habitats. In particular, a *V. alginolyticus* TS13 GI (26,181 bp) carried a widespread colonization island (WCI) with a tight-adherence (*tad*)



locus containing 13 genes (*tadV-flp-rcpCAB-tadZABCDEFG*). Similarly, the alkyl hydroperoxide reductase proposed to scavenge endogenous hydrogen peroxide in the deep-sea (Hasan et al., 2015) was detected in all eight *V. diabollicus* strains.

Subclade-Associated Genes

The search for genes ubiquitous in the *V. diabollicus* subclade ($N = 8$ strains) but absent from the other 41 strains, returned five genes (Table 3). All five were single-copy and highly conserved, producing gene and protein alignments with greater than 96% sequence identity. Two encoded hypothetical proteins (located on chromosome I) while the remaining three encoded a hydroxyectoine utilization dehydratase, an ornithine cyclodeaminase and a phosphoesterase (located on chromosome II).

Metagenome Survey

Three subclade-associated coding sequences (hydroxyectoine utilization dehydratase, ornithine cyclodeaminase, and phosphoesterase) were selected as *V. diabollicus*-associated markers. When queried against the NCBI Metagenome Protein Database, proteins homologous to each were detected in more than 100 environmental metagenomes and the majority were categorized as marine metagenomes. For the hydroxyectoine utilization dehydratase query, the highest scoring hit showed 82% coverage, 59% identity and a $2e-105$ *E*-value. For the ornithine cyclodeaminase query, the highest scoring hit showed 96% coverage, 49% identity and a $4e-96$ *E*-value. For the phosphoesterase query, the highest scoring hit showed 46% coverage, 23% identity and a $2e-06$ *E*-value. Thus, all database hits were declared non-significant as the sequence identity was not comparable to the subclade protein alignments (i.e., greater than 95%).

Potential Virulence

A comparison of SEED subsystem features revealed a large repertoire of features associated with virulence and defense (Figure 5). Importantly, each of the eight *V. diabollicus* subclade genomes demonstrated a nearly identical subsystem profile. For example, the accessory colonization factor *AcfA*, assigned to the Adhesion subsystem, was present in all eight genomes. Similarly, 10–11 features assigned to the Bacteriocins and Ribosomally Synthesized Antibacterial Peptides subsystem were present in all genomes as part of a colicin V production cluster, and the 12–13 features assigned to the Invasion and Intracellular Resistance subsystem were present in all genomes as part of a suite of proteins with homology to a *Mycobacterium* virulence operon involved in protein synthesis. The exception was the Toxins and Superantigens subsystem where the number of features ranged from zero to six; however, when present, proteins assigned to this subsystem included the transcriptional activator *ToxR*, the transmembrane regulatory protein *ToxS*, one to two copies of the *Zonula occludens* toxin (*Zot*) and one to two copies of the accessory cholera enterotoxin (*Ace*).

Evaluation of subsystem features also revealed that 74–79 features were assigned to the Resistance to Antimicrobial Compounds subsystem. Further, when that subsystem

was expanded into more specific subsystems (e.g., cobalt-zinc-cadmium resistance, tetracycline resistance and multidrug efflux pumps), all eight shared a nearly identical resistome (Figure 5). The resistome included features associated with resistance to toxic metals and metalloids (e.g., copper, cobalt, zinc, cadmium, chromium, and arsenic) and features associated with resistance to antibiotics (i.e., fluoroquinolones, fosfomycin, tetracycline, and penicillin). Additionally, the resistome included 22–24 efflux pumps per genome. These efflux pumps included pumps belonging to the multidrug and toxin extrusion (MATE) and resistance-nodulation-cell division (RND) families.

A phylogenetic analysis of proteins assigned to the β -lactamase subsystem (1–3 proteins per genome) revealed the presence of two distinct β -lactamases (Supplementary Figure S2). According to the comprehensive ARDB, these enzymes were identified as class A and class C β -lactamases associated with resistance to carbenicillin and cephalosporin, respectively. Similarly, the two *FosA* enzymes conferring resistance to fosfomycin, present as single copy genes in *V. alginolyticus* TS13 and V2, were identified as identical glutathione transferases, which confer resistance by catalyzing the addition of glutathione to fosfomycin.

Antibiotic Susceptibility

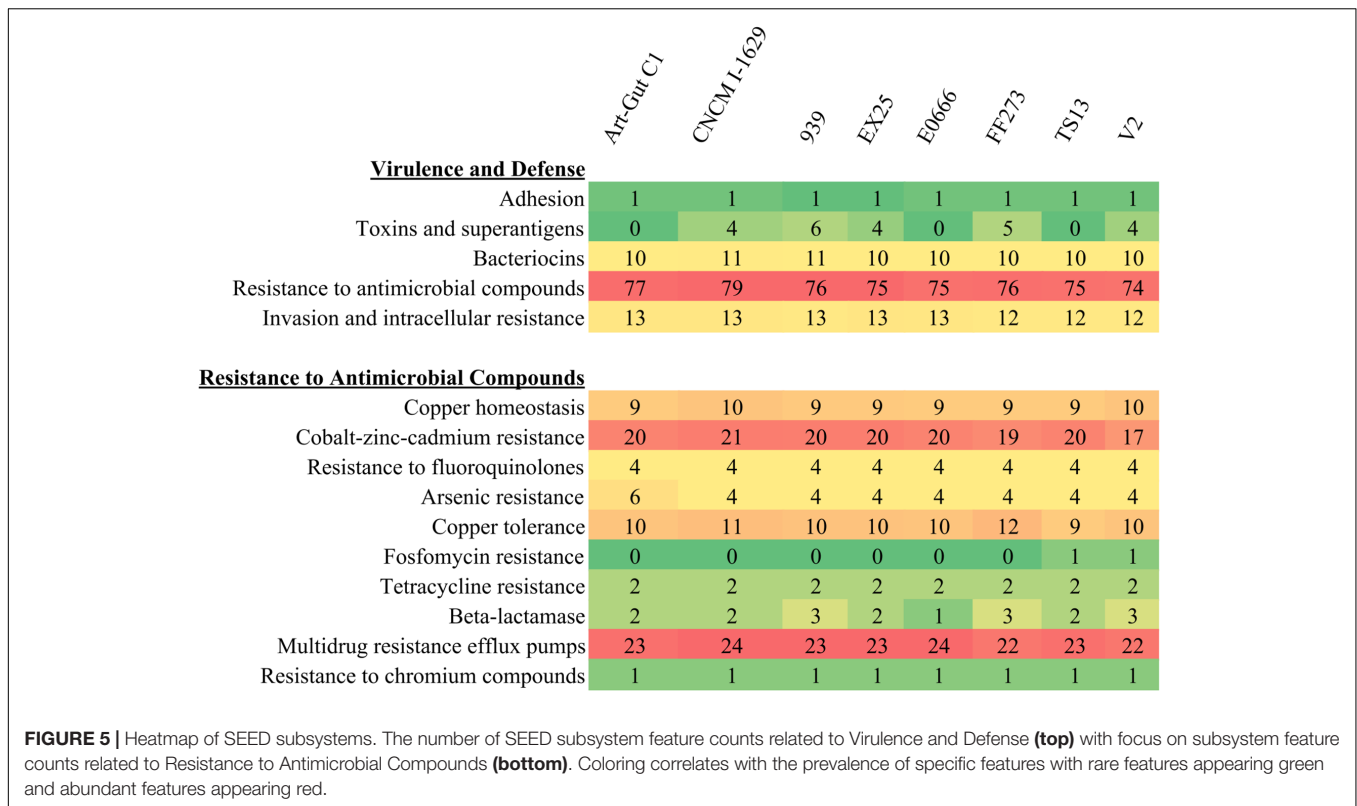
A positive Cefinase test indicated that *V. antiquarius* 939 and *V. parahaemolyticus* RIMD2210633A produced β -lactamase. *V. antiquarius* 939 exhibited resistance to penicillin, ampicillin, and cephalothin, and susceptibility to carbenicillin. *V. parahaemolyticus* RIMD2210633 exhibited resistance to penicillin and cephalothin, intermediate susceptibility to ampicillin, and susceptibility to carbenicillin.

DISCUSSION

The taxonomy of the *Vibrio* genus is a story of rapid expansion. In the past four decades, the number of recognized species has grown from five (Bergey et al., 1974) to 139⁴. The taxonomy of the *Vibrio* genus is also a story of constant revision. In particular, species boundaries between members of the Harveyi clade (e.g., *V. campbellii*, *V. harveyi*, *V. rotiferianus*, *V. natriegens*, and *V. parahaemolyticus*) have proven especially difficult to delineate (Thompson et al., 2006). However, the advent of multilocus (Thompson et al., 2005, 2007; Sawabe et al., 2007, 2013; González-Escalona et al., 2008) and genome-scale phylogenetics (Thompson et al., 2009; Lin et al., 2010; Urbanczyk et al., 2013, 2014, 2016; Ke et al., 2017) has advanced our understanding of this once enigmatic clade. This study builds on previous work by clarifying the relatedness of two Harveyi clade species first isolated from the deep-sea: *V. diabollicus* and *V. antiquarius*.

Ultimately, questions surrounding the identity of *Vibrio* sp. 939 motivated genome sequencing and a subsequent genome-scale taxonomic analysis of 49 core and non-core members of the Harveyi clade. The initial examination of the draft genome in RAST indicated that the isolate was more

⁴<http://www.bacterio.net/vibrio.html>



closely related to *V. antiquarius*, motivating the *V. antiquarius* 939 assignment in GenBank. However, upon closer examination, a maximum-likelihood 49-genome phylogeny based on the concatenation of 1,109 homologous genes assigned the isolate to a monophyletic subclade comprised of two *V. diabolicus* strains (Art-Gut C1 and CNCM I-1629), two *V. antiquarius* strains (939 and EX25) and four *V. alginolyticus* strains (E0666, FF273, TS13, and V2).

Previous studies have proposed that soft species boundaries within the Harveyi clade can result from complex evolutionary scenarios, such as recombination, that obfuscate the phylogenetic signal (Fraser et al., 2007; Thompson et al., 2007). Seeing that complex evolutionary scenarios are difficult to represent in a phylogenetic tree (Huson and Bryant, 2006), a phylogenetic network was constructed to investigate the uncertainty of the species boundaries. The network revealed two groups of

species that are separated by short branching: (1) *V. jasicida*, *V. harveyi*, *V. owensii*, and *V. campbellii*, and (2) *V. alginolyticus* and *V. diabolicus*. Additionally, a more pronounced pattern of reticulation between *V. alginolyticus* and *V. diabolicus* indicated that recombination obscured the phylogenetic signal between these two species. This short branching and pronounced reticulation could explain why the boundary between these two species was heretofore unresolved. Regardless, the subclades produced by the network analysis were well supported and mirrored the subclade assignments produced by the maximum-likelihood phylogeny.

The 95–96% ANI threshold is commonly applied to support species assignments (Richter and Rosselló-Móra, 2009; Varghese et al., 2015). This study showed that *V. diabolicus* (Art-Gut C1 and CNCM I-1629), *V. antiquarius* (939 and EX25), and *V. alginolyticus* (E0666, FF273, TS13, and V2) genomes shared

TABLE 3 | Description of open reading frames (ORFs) ubiquitous in *V. diabolicus* subclade strains (i.e., detected in *V. diabolicus* Art-Gut C1 and CNCM I-1629, *V. antiquarius* 939 and EX25, and *V. alginolyticus* E0666, FF273, TS13, and V2) but absent from the 41 other strains included in this study.

Accession ^a	Open reading frame ^b	Description	Length ^c	Sequence identity ^d
ACY49842.1	CP001805:76402-76626	Hypothetical protein	225 75	96.27 96.31
ACY50295.1	CP001805:552520-552717	Hypothetical protein	198 66	96.53 95.83
ACY53324.1	CP001806:741108-742082	Hydroxyectoine utilization dehydratase	975 325	99.14 99.14
ACY53325.1	CP001806:742097-743056	Ornithine cyclodeaminase	960 320	98.26 99.29
ACY53763.1	CP001806:1226999-1229245	Phosphoesterase	2,247 749	98.39 98.57

^aGenBank accession number in the type strain *V. antiquarius* EX25. ^bCoordinates of the ORF in chromosome I (CP001805) or chromosome II (CP001806). ^cLength of the gene (bp) and protein (aa). ^dAverage sequence identity of the gene and protein alignments between all *V. diabolicus* subclade genomes.

greater than 97% ANI, which is a strong indication that they constitute the same species. In contrast, these genomes shared less than 92% ANI with representative *V. alginolyticus* genomes, confirming that they are distinct in comparison to their nearest phylogenetic neighbor. A similar study used the 97% ANI threshold to support the conclusion that *V. inhibens* is a heterotypic synonym of *V. jasicida* (Urbanczyk et al., 2016). In this study, the analysis of ANI values clearly showed that two *V. antiquarius* strains (939 and EX25) and four *V. alginolyticus* strains (E0666, FF273, TS13, and V2) are synonyms of *V. diabollicus*.

That the eight *V. diabollicus* strains described in this study were isolated from a range of sources (i.e., horse mackerel, seawater, sediment, dentex, oyster, artemia, and polychaete) collected from locations around the world (i.e., China, India, Greece, United States, East Pacific Rise, and Chile) suggests diverse physiological capabilities. Extensive genomic variability is thought to underlie a prokaryotic species' ability to persist in a wide range of habitats (Cordero and Polz, 2014), and the genomic variability of *V. diabollicus*, evidenced in this study by a large accessory genome (relative to the number of genomes analyzed) and the presence of dozens of strain-specific GIs, may correlate with the species' ability to persist in these varied habitats across the globe. For example, the ability to colonize substrates and form biofilms is critical to persistence, and the production of an adhesive Flp (fimbrial low-molecular-weight protein) pilus (encoded by a 13-gene *tad* locus found on a large GI in *V. alginolyticus* TS13) could facilitate the colonization of diverse habitats (Kachlany et al., 2001). However, the species appears to be conditionally rare as species-associated genetic markers, developed in this study, were not detected in the NCBI environmental metagenome database.

That the only known representatives of *V. antiquarius* (939 and EX25) are taxonomic synonyms of *V. diabollicus* represents a significant shift in the systematics of these bacteria as both *V. diabollicus* and *V. antiquarius* were long regarded as distinct species (Raguénès et al., 1997; Goudenège et al., 2014; Hasan et al., 2015). Yet the most striking synonym was *V. alginolyticus* E0666, which was isolated from spoiled horse mackerel associated with food poisoning (Cao et al., 2013) and later shown to be virulent in a murine model (Liu et al., 2013). The presence of virulence factors in *V. antiquarius* EX25, such as Ace and Zot, was described previously (Hasan et al., 2015). This study showed that the same virulence factors were present in other *V. diabollicus* subclade genomes, suggesting that virulence-associated genes could be ubiquitous across the species, which is not surprising given the prevalence of such features in closely related species (e.g., *V. alginolyticus* and *V. parahaemolyticus*). The species-wide conservation of these features supports the hypothesis that virulence-associated genes could play a fundamental role in the ecological setting (Hasan et al., 2015).

In the event that some *V. diabollicus* strains are virulent, intrinsic resistance to antibiotics could complicate treatment as the analysis of the species' resistome revealed the presence of alleles and genes associated with resistance to fluoroquinolone, fosfomycin, tetracycline, and penicillin. The production of β -lactamase and resistance to penicillin, ampicillin, and

cephalosporin was confirmed experimentally in *V. antiquarius* 939. The presence of numerous efflux pumps, including the MATE and RND families, could expand mechanisms of resistance. The ubiquity of these features is concerning given the species' rarity and heretofore-perceived isolation, whereas widespread and progressive antibiotic resistance is a long-standing trend in *V. cholerae* (Garg et al., 2000; Uppal et al., 2017), *V. parahaemolyticus* (Baker-Austin et al., 2008; Devi et al., 2009), and *V. vulnificus* (Baker-Austin et al., 2009; Shaw et al., 2014).

CONCLUSION

The results of this taxonomic and comparative genomic analysis revealed that *V. diabollicus* and *V. antiquarius* should be classified as the same species. As *V. diabollicus* was discovered and published (Raguénès et al., 1997; Goudenège et al., 2014) prior to *V. antiquarius* (Hasan et al., 2015), the latter is a taxonomic synonym. *V. alginolyticus* remains a distinct species, but four misidentified strains (E0666, FF273, TS13, and V2) are synonyms of *V. diabollicus*. The expansion of this species is ongoing as three new *V. diabollicus* genomes (FDAARGOS_96, LMG 3418, and FDAARGOS_105) were made publicly available during the review of this manuscript.

AUTHOR CONTRIBUTIONS

JWT, WN, RP, EA, and MS conceived and designed the research project. JWT, JJT, AM, LP, NE, and DNA conducted the genomic analyses and laboratory experiments. All authors contributed to data interpretation and the writing of the manuscript.

FUNDING

This study was funded by the National Oceanic and Atmospheric Administration's (NOAA) Oceans and Human Health Initiative (OHHI), the National Marine Fisheries Service (NMFS), the National Research Council (NRC), the Gordon and Betty Moore Foundation (EVA #537.01), and Texas A&M University-Corpus Christi (TAMU-CC).

ACKNOWLEDGMENTS

This study would not have been possible without the support of our friends and colleagues. In particular, we thank Chris T. Berthiaume, Rhonda Morales, and Thomas Merrick for their assistance with sequencing and data analysis, and we thank Drs. Toshio Kodama and Tetsuya Iida and the Pathogenic Microbes Repository unit of the Research Institute for Microbial Diseases at Osaka University for providing the *V. parahaemolyticus* RIMD2210633 strain. We also thank Dr. Daniele Provenzano for kindly proofreading the manuscript and Dr. Magesh Thiyagarajan for his generous gift of the laboratory supplies used during this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.01893/full#supplementary-material>

FIGURE S1 | Phylogenetic tree of the *V. harveyi* clade. A maximum-likelihood tree representing 49 *V. harveyi* clade genomes was inferred from the concatenated alignment of 1,109 homologous genes. Node labels show the bootstrap support values. Nodes with strong support (>85) were highlighted in green while nodes

with weak support (<85) were highlighted in red. Branch lengths represent the average number of substitutions per site. The tree was rooted to the outgroup comprised of *V. azureus* NBRC104587 and *V. sagamiensis* NBRC 104589.

FIGURE S2 | Phylogenetic tree of β -lactamases. A maximum-likelihood tree showing the relatedness of 18 β -lactamases present in the eight *V. diabollicus* subclade genomes. Node labels show the bootstrap support values greater than 85. Branch lengths represent the average number of substitutions per site. The tree was rooted at the midpoint.

TABLE S1 | List of the 49 Harveyi clade genomes included in this study. Genomes in bold font belong to the new *V. diabollicus* subclade described in this study.

REFERENCES

- Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., and Beatson, S. A. (2011). BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. doi: 10.1186/1471-2164-12-402
- Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Austin, B., and Zhang, X. H. (2006). *Vibrio harveyi*: a significant pathogen of marine vertebrates and invertebrates. *Lett. Appl. Microbiol.* 43, 119–124. doi: 10.1111/j.1472-765X.2006.01989.x
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75–95. doi: 10.1186/1471-2164-9-75
- Baker-Austin, C., McArthur, J. V., Lindell, A. H., Wright, M. S., Tuckfield, R. C., Gooch, J., et al. (2009). Multisite analysis reveals widespread antibiotic resistance in the marine pathogen *Vibrio vulnificus*. *Microb. Ecol.* 57, 151–159. doi: 10.1007/s00248-008-9413-8
- Baker-Austin, C., McArthur, J. V., Tuckfield, R. C., Najarro, M., Lindell, A. H., Gooch, J., et al. (2008). Antibiotic resistance in the shellfish pathogen *Vibrio parahaemolyticus* isolated from the coastal water and sediment of Georgia and South Carolina, USA. *J. Food Prot.* 71, 2552–2558. doi: 10.4315/0362-028X-71.12.2552
- Bej, A. K., Patterson, D. P., Brasher, C. W., Vickery, M., Jones, D. D., and Kaysner, C. A. (1999). Detection of total and hemolysin-producing *Vibrio parahaemolyticus* in shellfish using multiplex PCR amplification of *tdh* and *trh*. *J. Microbiol. Methods* 36, 215–225. doi: 10.1016/S0167-7012(99)00037-8
- Bergey, D. H., Buchanan, R. E., and Gibbons, N. E. (1974). *Bergey's manual of determinative bacteriology*, 8th Edn, vol. 1. Baltimore, MD: Williams & Wilkins Co.
- Bryant, D., and Moulton, V. (2004). Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.* 21, 255–265. doi: 10.1093/molbev/msh018
- Cao, Y., Liu, X. F., Zhang, H. L., Chen, Y. J., and Hu, C. J. (2013). Draft genome sequence of the human-pathogenic bacterium *Vibrio alginolyticus* E0666. *Genome Announc.* 1:e686-13. doi: 10.1128/genomeA.00686-13
- Capella-Gutierrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
- Chakraborty, S., Nair, G. B., and Shinoda, S. (1997). Pathogenic *Vibriosis* in the natural aquatic environment. *Rev. Environ. Health* 12, 63–80. doi: 10.1515/REVEH.1997.12.2.63
- Chikhi, R., and Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. *Bioinformatics* 30, 31–37. doi: 10.1093/bioinformatics/btt310
- Clinical and Laboratory Standards Institute [CLSI] (2010). *Methods for Antimicrobial Dilution and Disk Susceptibility Testing of Infrequently Isolated or Fastidious Bacteria, Approved Guideline*, 2nd Edn. Wayne, IL: Clinical Laboratory Standards Institute.
- Contreras-Moreira, B., and Vinuesa, P. (2013). GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl. Environ. Microbiol.* 79, 7696–7701. doi: 10.1128/AEM.02411-13
- Cordero, O. X., and Polz, M. F. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nat. Rev. Microbiol.* 12, 263–273. doi: 10.1038/nrmicro3218
- Devi, R., Surendran, P. K., and Chakraborty, K. (2009). Antibiotic resistance and plasmid profiling of *Vibrio parahaemolyticus* isolated from shrimp farms along the southwest coast of India. *World J. Microbiol. Biotechnol.* 25, 2005–2012. doi: 10.1007/s11274-009-0101-8
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Frans, I., Michiels, C. W., Bossier, P., Willems, K. A., Lievens, B., and Rediers, H. (2011). *Vibrio anguillarum* as a fish pathogen: virulence factors, diagnosis and prevention. *J. Fish Dis.* 34, 643–661. doi: 10.1111/j.1365-2761.2011.01279.x
- Fraser, C., Hanage, W. P., and Spratt, B. G. (2007). Recombination and the nature of bacterial speciation. *Science* 315, 476–480. doi: 10.1126/science.1127573
- Garg, P., Chakraborty, S., Basu, I., Datta, S., Rajendran, K., Bhattacharya, T., et al. (2000). Expanding multiple antibiotic resistance among clinical strains of *Vibrio cholerae* isolated from 1992–1997 in Calcutta, India. *Epidemiol. Infect.* 124, 393–399. doi: 10.1017/S0950268899003957
- Giardine, B., Riemer, C., Hardison, R. C., Burhans, R., Elnitski, L., Shah, P., et al. (2005). Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* 15, 1451–1455. doi: 10.1101/gr.4086505
- González-Escalona, N., Martínez-Urtaza, J., Romero, J., Espejo, R. T., Jaykus, L. A., and DePaola, A. (2008). Determination of molecular phylogenetics of *Vibrio parahaemolyticus* strains by multilocus sequence typing. *J. Bacteriol.* 190, 2831–2840. doi: 10.1128/JB.01808-07
- Goudenège, D., Boursicot, V., Versigny, T., Bonnetot, S., Ratskol, J., Sinquin, C., et al. (2014). Genome sequence of *Vibrio diabollicus* and identification of the exopolysaccharide HE800 biosynthesis locus. *Appl. Microbiol. Biotechnol.* 98, 10165–10176. doi: 10.1007/s00253-014-6086-8
- Hasan, N. A., Grim, C. J., Lipp, E. K., Rivera, I. N., Chun, J., Haley, B. J., et al. (2015). Deep-sea hydrothermal vent bacteria related to human pathogenic *Vibrio* species. *Proc. Natl. Acad. Sci. U.S.A.* 112, E2813–E2819. doi: 10.1073/pnas.1503928112
- Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030
- Johnson, C. N., Bowers, J. C., Griffitt, K. J., Molina, V., Clostio, R. W., Pei, S., et al. (2012). Ecology of *Vibrio parahaemolyticus* and *Vibrio vulnificus* in the coastal and estuarine waters of Louisiana, Maryland, Mississippi, and Washington (United States). *Appl. Environ. Microbiol.* 78, 7249–7257. doi: 10.1128/AEM.01296-12
- Kachlany, S. C., Planet, P. J., DeSalle, R., Fine, D. H., and Figurski, D. H. (2001). Genes for tight adherence of *Actinobacillus actinomycetemcomitans*: from plaque to plague to pond scum. *Trends Microbiol.* 9, 429–437. doi: 10.1016/S0966-842X(01)02161-8
- Kalyanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Karaolis, D. K., Lan, R., and Reeves, P. R. (1995). The sixth and seventh cholera pandemics are due to independent clones separately derived from environmental, nontoxicogenic, non-O1 *Vibrio cholerae*. *J. Bacteriol.* 177, 3191–3198. doi: 10.1128/jb.177.11.3191-3198.1995
- Ke, H. M., Prachumwat, A., Yu, C. P., Yang, Y. T., Promsri, S., Liu, K. F., et al. (2017). Comparative genomics of *Vibrio campbellii* strains and core species of the *Vibrio harveyi* clade. *Sci. Rep.* 7, 41394–41311. doi: 10.1038/srep41394
- Klein, S. L., Gutierrez-West, C. K., Mejia, D. M., and Lovell, C. R. (2014). Genes similar to the *Vibrio parahaemolyticus* virulence-related genes *tdh*, *tlh*, and *vscC2* occur in other *Vibrionaceae* species isolated from a pristine estuary. *Appl. Environ. Microbiol.* 80, 595–602. doi: 10.1128/AEM.02895-13

- Klimke, W., Agarwala, R., Badretin, A., Chetverin, S., Ciufu, S., Fedorov, B., et al. (2009). The national center for biotechnology information's protein clusters database. *Nucleic Acids Res.* 37, D216–D223. doi: 10.1093/nar/gkn734
- Laing, C., Buchanan, C., Taboada, E. N., Zhang, Y., Kropinski, A., Villegas, A., et al. (2010). Pan-genome sequence analysis using Panseq: an online tool for the rapid analysis of core and accessory genomic regions. *BMC Bioinformatics* 11:461. doi: 10.1186/1471-2105-11-461
- Letchumanan, V., Pusparajah, P., Tan, L. T.-H., Yin, W.-F., Lee, L.-H., and Chan, K.-G. (2015). Occurrence and antibiotic resistance of *Vibrio parahaemolyticus* from shellfish in Selangor, Malaysia. *Front. Microbiol.* 6:1417. doi: 10.3389/fmicb.2015.01417
- Li, L., Stoeckert, C. J., and Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178–2189. doi: 10.1101/gr.1224503
- Lin, B., Wang, Z., Malanoski, A. P., O'Grady, E. A., Wimpee, C. F., Vuddhakul, V., et al. (2010). Comparative genomic analyses identify the *Vibrio harveyi* genome sequenced strains BAA-1116 and HY01 as *Vibrio campbellii*. *Environ. Microbiol. Rep.* 2, 81–89. doi: 10.1111/j.1758-2229.2009.00100.x
- Liu, B., and Pop, M. (2009). ARDB-antibiotic resistance genes database. *Nucleic Acids Res.* 37, D443–D447. doi: 10.1093/nar/gkn656
- Liu, X. F., Zhang, H., Liu, X., Gong, Y., Chen, Y., Cao, Y., et al. (2013). Pathogenic analysis of *Vibrio alginolyticus* infection in a mouse model. *Folia Microbiol.* 59, 167–171. doi: 10.1007/s12223-013-0279-x
- Magoc, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507
- Markert, S., Arndt, C., Felbeck, H., Becher, D., Sievert, S. M., Hügler, M., et al. (2007). Physiological proteomics of the uncultured endosymbiont of *Riftia pachyptila*. *Science* 315, 247–250. doi: 10.1126/science.1132913
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12. doi: 10.14806/ej.17.1.200
- Matsumoto, C., Okuda, J., Ishibashi, M., Iwanaga, M., Garg, P., Rammamurthy, T., et al. (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.
- Minh, B. Q., Nguyen, M. A. T., Haeseler, and von, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* 30, 1188–1195. doi: 10.1093/molbev/mst024
- Nasu, H., Iida, T., Sugahara, T., Yamaichi, Y., Park, K.-S., Yokoyama, T., et al. (2000). A filamentous phage associated with recent pandemic *Vibrio parahaemolyticus* O3:K6 strains. *J. Clin. Microbiol.* 38, 2156–2161.
- Nguyen, L. T., Schmidt, H. A., Haeseler, von, A., and Minh, B. Q. (2014). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226
- Pacini, F. (1854). Osservazioni microscopiche e deduzioni patologiche sul cholera asiatico. *Gaz. Med. Ital. Toscana* 6, 397–405.
- Paranjpye, R., Hamel, O. S., Stojanovski, A., and Liermann, M. (2012). Genetic diversity of clinical and environmental *Vibrio parahaemolyticus* strains from the Pacific Northwest. *Appl. Environ. Microbiol.* 78, 8631–8638. doi: 10.1128/AEM.01531-12
- Polz, M. F., Hunt, D. E., Preheim, S. P., and Weinreich, D. M. (2006). Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 2009–2021. doi: 10.1098/rstb.2006.1928
- Raguénès, G., Christen, R., Guezennac, J., Pignet, P., and Barbier, G. (1997). *Vibrio diabollicus* sp. nov., a new polysaccharide-secreting organism isolated from a deep-sea hydrothermal vent polychaete annelid, *Alvinella pompejana*. *Int. J. Syst. Bacteriol.* 47, 989–995. doi: 10.1099/00207713-47-4-989
- Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19126–19131. doi: 10.1073/pnas.0906412106
- Richter, M., Rosselló-Móra, R., Oliver Glöckner, F., and Peplies, J. (2016). JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32, 929–931. doi: 10.1093/bioinformatics/btv681
- Rosenberg, E., Koren, O., Reshef, L., Efrony, R., and Zilber-Rosenberg, I. (2007). The role of microorganisms in coral health, disease and evolution. *Nat. Rev. Microbiol.* 5, 355–362. doi: 10.1038/nrmicro1635
- Sawabe, T., Kita-Tsukamoto, K., and Thompson, F. L. (2007). Inferring the evolutionary history of *Vibrios* by means of multilocus sequence analysis. *J. Bacteriol.* 189, 7932–7936. doi: 10.1128/JB.00693-07
- Sawabe, T., Ogura, Y., Matsumura, Y., Feng, G., Mino, S., Nakagawa, S., et al. (2013). Updating the *Vibrio* clades defined by multilocus sequence phylogeny: proposal of eight new clades, and the description of *Vibrio tritonius* sp. nov. *Front. Microbiol.* 4:414. doi: 10.3389/fmicb.2013.00414
- Shade, A., Jones, S. E., Caporaso, J. G., Handelsman, J., Knight, R., Fierer, N., et al. (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mBio* 5:e1371-14. doi: 10.1128/mBio.01371-14
- Shaw, K. S., Rosenberg Goldstein, R. E., He, X., Jacobs, J. M., Crump, B. C., and Sapkota, A. R. (2014). Antimicrobial susceptibility of *Vibrio vulnificus* and *Vibrio parahaemolyticus* recovered from recreational and commercial areas of Chesapeake Bay and Maryland coastal bays. *PLoS One* 9:e89616-11. doi: 10.1371/journal.pone.0089616
- Strom, M. S., and Paranjpye, R. N. (2000). Epidemiology and pathogenesis of *Vibrio vulnificus*. *Microbes Infect.* 2, 177–188. doi: 10.1016/S1286-4579(00)00270-7
- Takemura, A. F., Chien, D. M., and Polz, M. F. (2014). Associations and dynamics of *Vibrionaceae* in the environment, from the genus to the population level. *Front. Microbiol.* 5:38. doi: 10.3389/fmicb.2014.00038
- Thompson, C. C., Vicente, A., Souza, R. C., Vasconcelos, A., Vesth, T., Alves, N., et al. (2009). Genomic taxonomy of *Vibrios*. *BMC Evol. Biol.* 9:258–273. doi: 10.1186/1471-2148-9-258
- Thompson, F. L., Austin, B., and Swings, J. E. (2006). “Taxonomy of the vibrios,” in *The Biology of Vibrios*, eds F. L. Thompson, B. Austin, and J. Swings (Washington, DC: ASM Press), 29–43. doi: 10.1128/9781555815714.ch3
- Thompson, F. L., Gevers, D., Thompson, C. C., Dawyndt, P., Naser, S., Hoste, B., et al. (2005). Phylogeny and molecular identification of *Vibrios* on the basis of multilocus sequence analysis. *Appl. Environ. Microbiol.* 71, 5107–5115. doi: 10.1128/AEM.71.9.5107-5115.2005
- Thompson, F. L., Gomez-Gil, B., Vasconcelos, A. T., and Sawabe, T. (2007). Multilocus sequence analysis reveals that *Vibrio harveyi* and *V. campbellii* are distinct species. *Appl. Environ. Microbiol.* 73, 4279–4285. doi: 10.1128/AEM.00020-07
- Thompson, F. L., Iida, T., and Swings, J. (2004). Biodiversity of *Vibrios*. *Microbiol. Mol. Biol. Rev.* 68, 403–431. doi: 10.1128/MMBR.68.3.403-431.2004
- Thompson, J. R., and Polz, M. F. (2006). “Dynamics of *Vibrio* populations and their role in environmental nutrient cycling,” in *The Biology of Vibrios*, eds F. L. Thompson, B. Austin, and J. Swings (Washington, DC: ASM Press), 190–203.
- Turner, J. W., Berthiaume, C. T., Morales, R., Armbrust, E. V., and Strom, M. S. (2016). Genomic evidence of adaptive evolution in emergent *Vibrio parahaemolyticus* ecotypes. *Elem. Sci. Anth* 4:117. doi: 10.12952/journal.elementa.000117
- Turner, J. W., Paranjpye, R. N., Landis, E. D., Biryukov, S. V., Gonzalez-Escalona, N., Nilsson, W. B., et al. (2013). Population structure of clinical and environmental *Vibrio parahaemolyticus* from the Pacific Northwest coast of the United States. *PLoS One* 8:e55726. doi: 10.1371/journal.pone.0055726
- Uppal, B., Mehra, B., Panda, P. S., and Kumar, S. K. (2017). Changing epidemiology and antimicrobial resistance pattern of *Vibrio cholerae* isolates at a tertiary care health laboratory in North India (2011–2015). *Trop. J. Med. Res.* 20, 132–137. doi: 10.4103/tjmr.tjmr_17_17
- Urbanczyk, H., Ogura, Y., and Hayashi, T. (2013). Taxonomic revision of *Harveyi* clade bacteria (family *Vibrionaceae*) based on analysis of whole genome sequences. *Int. J. Syst. Evol. Microbiol.* 63, 2742–2751. doi: 10.1099/ijso.0.051110-0
- Urbanczyk, H., Ogura, Y., and Hayashi, T. (2014). Contrasting inter- and intraspecies recombination patterns in the “*Harveyi* Clade” *Vibrio* collected over large spatial and temporal scales. *Genome Biol. Evol.* 7, 71–80. doi: 10.1093/gbe/evu269
- Urbanczyk, Y., Ogura, Y., Hayashi, T., and Urbanczyk, H. (2016). Genomic evidence that *Vibrio inhibens* is a heterotypic synonym of *Vibrio jasicida*. *Int. J. Syst. Evol. Microbiol.* 66, 3214–3218. doi: 10.1099/ijsem.0.001173

- Varghese, N. J., Mukherjee, S., Ivanova, N., Konstantinidis, K. T., Mavrommatis, K., Kyrpides, N. C., et al. (2015). Microbial species delineation using whole genome sequences. *Nucleic Acids Res.* 43, 6761–6771. doi: 10.1093/nar/gkv657
- Yáñez, R., Bastías, R., Higuera, G., Salgado, O., Katharios, P., Romero, J., et al. (2015). Amplification of *tlh* gene in other *Vibrionaceae* species by species-specific multiplex PCR of *Vibrio parahaemolyticus*. *Electron. J. Biotechnol.* 18, 459–463. doi: 10.1016/j.ejbt.2015.09.007
- Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Turner, Tallman, Macias, Pinnell, Elledge, Nasr Azadani, Nilsson, Paranipye, Armbrust and Strom. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.