



# Draft Genome Sequences of 42 Environmental *Vibrio vulnificus* Strains Isolated from the Northern Gulf of Mexico

Megan M. Mullis,<sup>a</sup> I-Shuo Huang,<sup>a</sup> Githzette M. Planas-Costas,<sup>a</sup> Reaveyn Pray,<sup>a</sup> Gregory W. Buck,<sup>a</sup> Arian Nassiri,<sup>b</sup> Courtney Fuentes,<sup>b</sup> Lauren Turner,<sup>b</sup> Gabriel D. Ramirez,<sup>a\*</sup> Joanna B. Mott,<sup>c</sup> Jeffrey W. Turner<sup>a</sup>

<sup>a</sup>Department of Life Sciences, Texas A&M University—Corpus Christi, Corpus Christi, Texas, USA

<sup>b</sup>Virginia Division of Consolidated Laboratory Services, Department of General Services, Richmond, Virginia, USA

<sup>c</sup>College of Natural Sciences and Mathematics, California State University, Sacramento, California, USA

**ABSTRACT** *Vibrio vulnificus* is a Gram-negative bacterium and an opportunistic pathogen that can cause septicemia or necrotizing fasciitis. Here, we report the draft genome sequences of 42 environmental *V. vulnificus* strains collected from the northern Gulf of Mexico. These data will allow for more robust comparisons between clinical and environmental strains.

*Vibrio vulnificus* is a Gram-negative marine bacterium with a single polar flagellum (1). This bacterium is capable of producing necrotizing fasciitis if it enters the human body through broken skin or may cause sepsis if ingested, usually through the consumption of raw oysters (2). Previous studies have identified clinical and environmental ecotypes based on sequence polymorphism of the 16S rRNA gene (3) and the virulence-correlated gene (*vcg*) (4). Whole-genome sequencing has confirmed the existence of distinct clinical and environmental ecotypes (5), and the comparative analysis of clinical and environmental genomes has aided in the identification of genes likely required for virulence (6, 7). However, virulence remains poorly understood, and the lack of a large public collection of nonpathogenic genomes has been cited as a limitation (8). To address this data gap, we report the draft genome sequences of 42 environmental strains isolated from the northern Gulf of Mexico (NGOM).

The 42 environmental *V. vulnificus* strains were isolated previously from seven locations in the Texas segment of the NGOM between August 2006 and July 2007 (9). Isolates were cultured at 37°C overnight on tryptic soy blood agar (Remel, San Diego, CA) and confirmed as *V. vulnificus* with matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry (MALDI Biotyper; Bruker, Billerica, MA), as described previously (10). Genomic DNA was extracted with a QIAamp DNA minikit (Qiagen, Hilden, Germany) and quantified with a Qubit double-stranded DNA (dsDNA) broad-range (BR) assay kit (Fisher Scientific, Hampton, NH). Multiplexed paired-end libraries were prepared using a Nextera XT index kit (Illumina, San Diego, CA), per the manufacturer's instructions, with the modification that libraries were normalized by the estimated genome size prior to pooling. Sequencing was completed using the Illumina MiSeq 500-cycle kit (v2) following standard FastQ-only generation protocols to produce 250-bp paired-end reads. Overlapping paired reads were merged with FLASH version 1.2.11 (option -M 300) (11). Adapter sequences and low-quality bases were trimmed using Trim Galore! version 0.4.0 (options -paired -retain\_unpaired) (12). Trimmed reads for the 42 genomes were assembled using SPAdes version 3.9.0 (option -k 127) (13). The 42 genomes were also assembled using MaSuRCA version 3.2.8 (default options) (14) using the insert size distribution estimated with BWA (default options) (15). For this purpose, paired reads were aligned to a draft genome (assembled with SPAdes), and the mean insert size and standard deviation were estimated by parsing the final

**Citation** Mullis MM, Huang I-S, Planas-Costas GM, Pray R, Buck GW, Nassiri A, Fuentes C, Turner L, Ramirez GD, Mott JB, Turner JW. 2019. Draft genome sequences of 42 environmental *Vibrio vulnificus* strains isolated from the northern Gulf of Mexico. Microbiol Resour Announc 8:e00200-19. <https://doi.org/10.1128/MRA.00200-19>.

**Editor** Julia A. Maresca, University of Delaware

**Copyright** © 2019 Mullis et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Jeffrey W. Turner, [jeffrey.turner@tamucc.edu](mailto:jeffrey.turner@tamucc.edu).

\* Present address: Gabriel D. Ramirez, Water Department, City of Corpus Christi, Corpus Christi, Texas, USA.

**Received** 4 March 2019

**Accepted** 17 June 2019

**Published** 11 July 2019

**TABLE 1** Summary of the 42 *Vibrio vulnificus* draft genome assemblies from the northern Gulf of Mexico

Isolate	GenBank accession no.	Genome size (bp)	No. of contigs	$N_{50}$ value (bp)	G+C content (%)	Sequence coverage (×)	No. of coding genes	Provenance <sup>a</sup>
18022	RCGD000000000	4,845,321	65	270,380	46.75	123.8	4,544	Aransas Bay, 2007
18023	RCGC000000000	4,984,587	83	256,006	46.62	97.1	4,664	Bird Island, 2007
18024	RCGB000000000	4,941,116	111	117,910	46.68	96.34	4,540	Bayside, 2007
18025	RCGA000000000	5,240,241	176	90,860	46.47	82.96	4,999	Bayside, 2007
18026	RCFZ000000000	4,844,420	80	233,461	46.64	75.32	4,564	Bayside, 2007
18027	RCFY000000000	4,843,483	84	195,670	46.70	68.94	4,590	Bayside, 2007
18028	RCFX000000000	4,966,197	79	704,086	46.66	63	4,686	Bayside, 2007
18029	RCFW000000000	4,793,233	95	136,022	46.64	58.78	4,566	Bayside, 2007
18030	RCFV000000000	4,979,258	119	141,772	46.63	40.66	4,653	Bayside, 2007
18031	RCFU000000000	4,850,485	92	130,871	46.70	95.98	4,593	Bayside, 2007
18032	RCFT000000000	4,801,036	113	141,861	46.70	107	4,496	Bayside, 2007
18033	RBZB000000000	5,055,915	162	80,566	46.69	53.08	4,849	Bayside, 2006
18034	RBZD000000000	4,823,242	63	344,960	46.68	102.24	4,471	Bayside, 2006
18035	RBZH000000000	5,240,024	158	94,332	46.49	51.7	4,953	Bayside, 2006
18036	RBZG000000000	4,873,384	71	185,180	46.69	135.3	4,526	Bayside, 2006
18037	RBZF000000000	5,094,022	170	74,007	46.43	119.22	4,720	Copano Bay, 2007
18038	RBZE000000000	5,195,336	252	59,128	46.40	112.54	4,850	Copano Bay, 2007
18039	RBZJ000000000	4,988,278	99	208,583	46.61	58.08	4,678	Copano Bay, 2007
18040	RBZK000000000	4,840,145	79	145,069	46.69	139.02	4,450	Copano Bay, 2007
18041	RBZL000000000	5,412,903	82	235,674	46.51	86.9	5,035	Copano Bay, 2007
18042	RBZC000000000	4,933,479	57	393,935	46.82	104.6	4,492	Copano Bay, 2007
18043	RBZI000000000	4,905,570	90	202,561	46.66	60.38	4,614	Copano Bay, 2007
18044	RPGM000000000	4,935,504	82	207,050	46.63	59.44	4,706	Copano Bay, 2007
18045	RHHD000000000	4,800,809	82	147,435	46.77	42.42	4,469	Copano Bay, 2007
18047	RHHE000000000	4,820,655	103	125,851	46.78	41.88	4,429	Copano Bay, 2007
18048	RHHF000000000	5,082,584	94	209,253	46.72	144.84	4,816	Copano Bay, 2007
18049	RHHG000000000	4,907,792	104	224,385	46.66	99.46	4,612	Copano Bay, 2007
18050	RHHH000000000	4,957,444	99	210,035	46.69	89.38	4,663	Copano Bay, 2007
18051	RHHI000000000	4,927,428	75	187,128	46.79	52.24	4,496	Copano Bay, 2007
18052	RHHJ000000000	5,130,949	88	202,763	46.70	121.64	4,867	Copano Bay, 2007
18053	RPGN000000000	4,884,811	73	204,930	46.88	47.54	4,479	Copano Bay, 2007
18054	RHHK000000000	4,778,453	93	166,931	46.72	68.9	4,489	Copano Bay, 2006
18055	RHHL000000000	4,842,045	79	198,581	46.69	127.62	4,519	Copano Bay, 2006
18056	RBWJ000000000	4,929,131	44	304,520	46.86	105.94	4,490	Copano Bay, 2007
18057	RBWK000000000	4,931,307	79	213,852	46.71	132.62	4,699	Cole Park, 2007
18058	RBWL000000000	4,911,551	97	158,422	46.64	72.88	4,576	Cole Park, 2006
18059	RBWM000000000	4,799,654	65	294,393	46.76	144.34	4,479	Cole Park, 2006
18060	RBWN000000000	4,990,343	46	375,098	46.75	98.92	4,617	Nueces Bay, 2007
18061	RBWQ000000000	5,081,056	36	450,849	46.56	119.28	4,721	Nueces Bay, 2007
18062	RBWP000000000	4,923,184	51	390,616	46.83	124.3	4,484	Nueces Bay, 2007
18063	RBWQ000000000	4,929,780	77	188,289	46.71	166.44	4,680	Redfish Bay, 2007
18064	RBWR000000000	5,057,894	117	130,275	46.70	90.8	4,816	Redfish Bay, 2006

<sup>a</sup> Provenance provided as location and year of collection. All locations are bodies of water in Texas.

alignment (in SAM format) with an awk script (16). The draft genome assemblies were filtered by length (500-bp cutoff), and assembly metrics (e.g., total genome size [bp], number of contigs,  $N_{50}$  value, and G+C content [%]) were calculated by QUAST version 4.1 (default options) (17) to determine which assembler (SPAdes or MaSuRCA) produced a higher-quality genome. For the genomes of two strains, 18057 and 18063, MaSuRCA produced higher-quality assemblies. Draft genomes were annotated using the National Center for Biotechnology Information's (NCBI's) Prokaryotic Genome Annotation Pipeline (PGAP) (18).

Table 1 shows the assembly metrics for the 42 *V. vulnificus* genomes, including the overall genome size (bp), number of contigs,  $N_{50}$  value, and G+C content (%). The availability of these genomes will augment future comparative genomic analyses focused on the differentiation of clinical and environmental strains. Additionally, the availability of these genomes will advance the understanding of the environmental *V. vulnificus* reservoir.

**Data availability.** These whole-genome shotgun projects have been deposited at GenBank under the accession numbers listed in Table 1. The raw sequence reads were deposited in the Sequence Read Archive under BioProject accession number [PRJNA475262](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA475262).

## ACKNOWLEDGMENTS

We thank the Virginia Division of Consolidated Laboratory Services for providing MALDI-TOF testing and Illumina sequencing of the isolate collection. We also thank Lee J. Pinnell for his assistance with genome assembly.

Food and Drug Administration Cooperative Agreement funding for the Maintenance and Enhancement of ISO/IEC 17025 Accreditation and Whole Genome Sequencing for State Food Testing Laboratories supported the sequencing of this sample collection (grant 5U18FD006188-02). This cooperative agreement aims to build and diversify public pathogen genome databases.

## REFERENCES

- Farmer J. 1979. *Vibrio* ("Benecke") *vulnificus*, the bacterium associated with sepsis, septicaemia, and the sea. *Lancet* 314:903. [https://doi.org/10.1016/S0140-6736\(79\)92715-6](https://doi.org/10.1016/S0140-6736(79)92715-6).
- Strom MS, Paranjpye RN. 2000. Epidemiology and pathogenesis of *Vibrio vulnificus*. *Microbes Infect* 2:177–188. [https://doi.org/10.1016/S1286-4579\(00\)00270-7](https://doi.org/10.1016/S1286-4579(00)00270-7).
- Nilsson WB, Paranjpye RN, DePaola A, Strom MS. 2003. Sequence polymorphism of the 16S rRNA gene of *Vibrio vulnificus* is a possible indicator of strain virulence. *J Clin Microbiol* 41:442–446. <https://doi.org/10.1128/jcm.41.1.442-446.2003>.
- Rosche TM, Yano Y, Oliver JD. 2005. A rapid and simple PCR analysis indicates there are two subgroups of *Vibrio vulnificus* which correlate with clinical or environmental isolation. *Microbiol Immunol* 49:381–389. <https://doi.org/10.1111/j.1348-0421.2005.tb03731.x>.
- Rosche TM, Binder EA, Oliver JD. 2010. *Vibrio vulnificus* genome suggests two distinct ecotypes. *Environ Microbiol Rep* 2:128–132. <https://doi.org/10.1111/j.1758-2229.2009.00119.x>.
- Gulig PA, de Crécy-Lagard V, Wright AC, Walts B, Telonis-Scott M, McIntyre LM. 2010. SOLiD sequencing of four *Vibrio vulnificus* genomes enables comparative genomic analysis and identification of candidate clade-specific virulence genes. *BMC Genomics* 11:512. <https://doi.org/10.1186/1471-2164-11-512>.
- Morrison SS, Williams T, Cain A, Froelich B, Taylor C, Baker-Austin C, Verner-Jeffreys D, Hartnell R, Oliver JD, Gibas CJ. 2012. Pyrosequencing-based comparative genomic analysis of *Vibrio vulnificus* environmental isolates. *PLoS One* 7:e37553. <https://doi.org/10.1371/journal.pone.0037553>.
- Baker-Austin C, Oliver JD. 2018. *Vibrio vulnificus*: new insights into a deadly opportunistic pathogen. *Environ Microbiol* 20:423–430. <https://doi.org/10.1111/1462-2920.13955>.
- Ramirez G, Buck G, Smith A, Gordon K, Mott J. 2009. Incidence of *Vibrio vulnificus* in estuarine waters of the south Texas Coastal Bend region. *J Appl Microbiol* 107:2047–2053. <https://doi.org/10.1111/j.1365-2672.2009.04390.x>.
- Branda JA, Fritsche TR, Burnham CA, Butler-Wu S, Doern C, Fedorko D, Frana TS, Gawoski JM, Ginocchio C, Hamula C, Hazen K, Kiss K, Kus J, Mingle L, Perry M, Pincus DH, Shea Y, Smoot L, Wang WF, Weaver M, Wengenack NL, Ziegelmann-Fjeld KI. 2017. Methods for the identification of cultured microorganisms using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, 1st ed. CLSI guideline M58. Clinical and Laboratory Standards Institute, Wayne, PA.
- Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27:2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>.
- Krueger F. 2015. Trim Galore! A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. [http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. *Bioinformatics* 29:2669–2677. <https://doi.org/10.1093/bioinformatics/btt476>.
- Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595. <https://doi.org/10.1093/bioinformatics/btp698>.
- Elledge NC, Eytan RI, Pinnell LJ, Pray R, Joyner JL, Wares JP, Sutherland KP, Lipp EK, Turner JW. 2019. Draft genome sequences of 38 *Serratia marcescens* isolates associated with acroporid serratiosis. *Microbiol Resour Announc* 8:e00194-19. <https://doi.org/10.1128/MRA.00194-19>.
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.