



# Comparative genomic analysis of clinical and environmental *Vibrio vulnificus* isolates revealed biotype 3 evolutionary relationships

Yael Koton<sup>1,2</sup>, Michal Gordon<sup>3</sup>, Vered Chalifa-Caspi<sup>3</sup> and Naïel Bisharat<sup>1,2\*</sup>

<sup>1</sup> Department of Medicine D, Emek Medical Center, Afula, Israel

<sup>2</sup> Bruce Rappaport Faculty of Medicine, Technion – Israel Institute of Technology, Haifa, Israel

<sup>3</sup> Bioinformatics Core Facility, National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva, Israel

## Edited by:

Radhey S. Gupta, McMaster University, Canada

## Reviewed by:

Nur A. Hasan, University of Maryland College Park, USA  
Carmen Amaro, University of Valencia, Spain

## \*Correspondence:

Naïel Bisharat, Department of Medicine D, Emek Medical Center, 21 Yitzhak Rabin Avenue, Afula 18101, Israel  
e-mail: bisharat\_na@clalit.org.il

In 1996 a common-source outbreak of severe soft tissue and bloodstream infections erupted among Israeli fish farmers and fish consumers due to changes in fish marketing policies. The causative pathogen was a new strain of *Vibrio vulnificus*, named biotype 3, which displayed a unique biochemical and genotypic profile. Initial observations suggested that the pathogen erupted as a result of genetic recombination between two distinct populations. We applied a whole genome shotgun sequencing approach using several *V. vulnificus* strains from Israel in order to study the pan genome of *V. vulnificus* and determine the phylogenetic relationship of biotype 3 with existing populations. The core genome of *V. vulnificus* based on 16 draft and complete genomes consisted of 3068 genes, representing between 59 and 78% of the whole genome of 16 strains. The accessory genome varied in size from 781 to 2044 kbp. Phylogenetic analysis based on whole, core, and accessory genomes displayed similar clustering patterns with two main clusters, clinical (C) and environmental (E), all biotype 3 strains formed a distinct group within the E cluster. Annotation of accessory genomic regions found in biotype 3 strains and absent from the core genome yielded 1732 genes, of which the vast majority encoded hypothetical proteins, phage-related proteins, and mobile element proteins. A total of 1916 proteins (including 713 hypothetical proteins) were present in all human pathogenic strains (both biotype 3 and non-biotype 3) and absent from the environmental strains. Clustering analysis of the non-hypothetical proteins revealed 148 protein clusters shared by all human pathogenic strains; these included transcriptional regulators, arylsulfatases, methyl-accepting chemotaxis proteins, acetyltransferases, GGDEF family proteins, transposases, type IV secretory system (T4SS) proteins, and integrases. Our study showed that *V. vulnificus* biotype 3 evolved from environmental populations and formed a genetically distinct group within the E-cluster. The unique epidemiological circumstances facilitated disease outbreak and brought this genotype to the attention of the scientific community.

**Keywords:** aquaculture, microbial genome, *Vibrio vulnificus*, whole genome shotgun sequences, evolution, core genome, accessory genome

## INTRODUCTION

*Vibrio vulnificus*, like other potentially pathogenic halophilic vibrios, is part of the marine microbiota. It occurs in high numbers in molluscan shellfish and in temperate zones, and especially during the warmer months it reaches sufficient concentrations to cause clinical disease in human (Oliver, 1989). The bacterium is capable of causing primary septicemia following its ingestion, and secondary septicemia through skin lesions in individuals with underlying chronic diseases. People who are most susceptible to *V. vulnificus* infection usually suffer from a chronic liver disease, primarily cirrhosis or alcoholic liver disease, diabetes mellitus, or diseases associated with iron overload such as hemochromatosis and thalassemia major (Oliver, 2006). Worldwide, the vast

majority of human disease has been reported from USA and Southeast Asia due to dietary habits of eating raw or undercooked seafood (Tacket et al., 1984; Klontz et al., 1988; Park et al., 1991; Chuang et al., 1992; Kumamoto and Vukich, 1998; Chiang and Chuang, 2003; Matsumoto et al., 2010). Reports from other parts of the world have been largely sporadic and typically due to wound infection (Bock et al., 1994; Melhus et al., 1995; Dalsgaard et al., 1996; Horre et al., 1998; Torres et al., 2002; Frank et al., 2006).

*V. vulnificus* populations have been divided into subpopulations based on phenotypic (biochemical) and genotypic characteristics. Phenotypically, three biotypes have been described; biotype 1, the most common worldwide (Oliver, 1989), biotype

2 mainly affecting eels (Tison et al., 1982), and biotype 3 that has been reported only in Israel (Bisharat et al., 1999). Several genotypic methods showed that *V. vulnificus* populations resolve into two main clusters, one dominated by strains from environmental sources and shellfish, and the other dominated by strains from human clinical samples (Nilsson et al., 2003; Bisharat et al., 2005; Gonzalez-Escalona et al., 2007). More recently a simple PCR-based assay showed excellent differentiation of *V. vulnificus* into two main genotypes based on their source of isolation, E-genotype (environmental) and C-genotype (clinical) (Rosche et al., 2005). Analysis of whole genome shotgun (WGS) sequence data confirmed this distinction and identified key genes specifically associated with each genotype (Morrison et al., 2012).

The disease outbreak in Israel in 1996 had several unique features. First, it was the first ever reported common source outbreak of *V. vulnificus* occurring among fish farmers and fish consumers handling live tilapia fish cultivated in inland fish farms (Bisharat and Raz, 1996). Second, biochemically the pathogen differed from existing pathogens, and was subsequently named biotype 3 (Bisharat et al., 1999). Finally, genotypically the bacterium displayed a unique pattern later identified as a hybrid clone of existing populations (Bisharat et al., 2005). Furthermore, because all disease cases were caused exclusively by a highly clonal pathogen (Bisharat et al., 1999, 2007a; Colodner et al., 2002, 2004; Miron et al., 2003; Zaidenstein et al., 2008), it was suggested that a recent genetic event have enabled a harmless environmental population to acquire capabilities to cause a deadly disease in human.

Our previous observations highlighted the importance of recombination in generating genetic diversity within *V. vulnificus* and showed that hybrids may be a recurring feature of *V. vulnificus* evolutionary biology (Bisharat et al., 2005, 2007b; Bishrat, 2010). However, the genetic divergence of *V. vulnificus* populations into two distinct genetic clusters was maintained across the genome and systematic over both chromosomes. Against a background of so much potential recombination we suggested two possible scenarios for the evolution of *V. vulnificus* populations, are the strains falling into two populations because of an old lineage split into two clonal complexes, which have accumulated differences through both mutation and horizontal transfer of diversity from a range of unknown sources, perhaps from other *Vibrio* species? Alternatively, is *V. vulnificus* a highly recombining species that structures into two populations for contemporary rather than phylogenetic reasons, implying some unknown ecological barrier that limits recombination between populations compared to within? (Bisharat et al., 2007b). Nevertheless, some inter-cluster hybrids do emerge and persist; the emergence of *V. vulnificus* biotype 3 was considered as such (Bisharat et al., 2005), yet we could not identify the parental lineage from which it had emerged. Recent observations based on WGS sequencing of a single biotype 3 strain implied that a single episode of genome hybridization of two bacterial populations is less likely to be the main event for the emergence of *V. vulnificus* biotype 3 and that it may have evolved by lateral gene transfer from other bacteria, such as *Shewanella* (Efimov et al., 2013). In order to elucidate the evolutionary pathways that led to the emergence of *V. vulnificus* biotype 3, we carried out a comparative genomic approach using three human

pathogenic biotype 3 strains and two environmental strains, all isolated from Israel. We aimed to study the pan genome structure of *V. vulnificus*, identify the set of novel sequences characterizing the human pathogenic biotype 3 and determine its phylogenetic relationship with other human pathogenic and environmental populations.

## METHODS

### BACTERIAL STRAINS

We used five *V. vulnificus* strains for the study (Table 1): three biotype 3 strains isolated from Israeli patients with invasive infection—VV9-09, VV 4-03, and 491771—and two biotype 1 strains, 101/4 and 2322, isolated from fish and fish pond water in Israel, respectively. All biotype 3 strains belong to sequence type 8 as determined by multi-locus sequence typing (MLST) (Bisharat et al., 2005).

### GENOME AMPLIFICATION, BUILDING AN ILLUMINA LIBRARY, AND SEQUENCING

Genomic DNA was extracted using a commercial kit (Qiagen DNeasy kit). Adaptors were added to each library during preparation according to the TruSeq protocol (Illumina) to produce multiplexed paired-end libraries. Pools of four samples were run on a sequencer (Illumina MiSeq) at the Technion Genome Center, Haifa, Israel, generating 250 base paired-end reads.

### READ MAPPING, GENOME ASSEMBLY, AND ANNOTATION

Sequence read data were analyzed using a pipeline of initial analysis that consisted of mapping reads to reference genomes and variant calling, genome assembly and gene annotation. The reads were mapped to two reference genomes, CMCP6 and YJ016 (biotype 1 human pathogenic *V. vulnificus* strains with complete genome sequence). Reads were mapped to the reference genomes using Burrows-Wheeler Aligner (BWA) 0.6.1 (Li and Durbin, 2009). Raw sequence data were tested for quality control using FastQC (Andrews, 2014). Due to decreased quality, the last 25 bp in read2 were trimmed from all the samples, improving the mapping of these reads against the reference genomes with only a minor reduction in the average coverage (~5%). The trimmed reads were used for the next steps of the analysis. In order to determine how many mismatches to allow per read during mapping to reference genomes, several mapping attempts were made for each sample against the two reference genomes (5, 15, 20, 25 mismatches per read). Allowing 20 mismatches per read increased the percentage of unique mapping to the reference genomes, while 25

**Table 1 | Characteristics of strains used in the study.**

Strain	Biotype	Source	Year	Country	MLST <sup>a</sup>
101/04	1	Fish	1997	Israel	59
2322	1	Water	1997	Israel	10
491771	3	Human	1997	Israel	8
VV9-09	3	Human	1999	Israel	8
VV 4-03	3	Human	2003	Israel	8

<sup>a</sup> Sequence type as determined by multi-locus sequence typing (MLST).

mismatches resulted in only a minor addition of unique mapping. We used Artemis (Rutherford et al., 2000) to visualize alignment of reads to the reference genomes and quantify the number of reads per gene in each sample. Reads per kilobase of exon model per million mapped reads (RPKM) values were calculated for each gene and ORFs in each sample. Identification of genes shared by the human pathogenic strains and absent from the environmental strains was estimated based on read coverage to the human-pathogenic and the environmental strains. Genes were regarded as “present” in a sample if they had more than 30 aligned reads, and “absent” if they had less than 30 such reads. CLC bio’s de novo assembler (v 6.5) (Qiagen) was used for de novo assembly of all the samples, using default parameters. RAST (Rapid Annotation using Subsystem Technology) (Aziz et al., 2008) was used for gene annotation.

### COMPARATIVE GENOMICS AND PHYLOGENETIC ANALYSIS

The assembled genomes generated in the current study were used for comparative genome sequencing analysis together with three reference human-pathogenic strains with complete genome sequence; CMCP6 (RefSeq: NC\_004459 and RefSeq: NC\_004460), YJ016 (RefSeq: NC\_005128; RefSeq: NC\_005139; RefSeq: NC\_005140), M06-24/O (RefSeq: NC\_014965; RefSeq: NC\_014966). Comparisons were also made with sequencing data from WGS sequencing projects available at <http://www.ncbi.nlm.nih.gov/Traces/wgs/>, this included VVyb1 (biotype 3 strain isolated from tilapia fish in Israel) (NZ\_AOCM00000000.1), BAA87 (biotype 3 strain isolated from human wound in Israel) (JDSE00000000.1), ATCC 27562 (biotype 1 strain isolated from human blood in the USA) (AMQV00000000.1), B2 (*V. vulnificus* strain isolated from human blood in China) (NZ\_AMQR00000000.1). There were also three environmental strains isolated from oysters in the USA: JY1305 (AFSW00000000.1), E64MW (AFSX00000000.1), and JY1701 (AFSY00000000.1). In addition, we used WGS sequence data of a biotype 2 strain (ATCC 33147) isolated from an infected eel in Japan 1979, the draft genome of this strain was recently published by our group (Koton et al., 2014) (JRQR01000000). The pan-genome was studied using Panseq (Laing et al., 2010) with application of two main modules; the Core and Accessory Genome Finder (CAGF) and the Novel Region Finder (NRF). For the purposes of the analysis, the CAGF module considers the “pan genome” to be comprised of all sequences selected as input for the analysis. The software uses a sequence file as a seed to which all other sequences are compared using MUMmer (Delcher et al., 1999). If a segment greater than the “minimum sequence size” is found in other than the seed, that segment is added to the pan genome. Next, the software fragments the entire pan-genome into segments of user-defined length (in the current study we used the default measures of 500 bp), and determines the presence or absence of each of these fragments in each of the original sequences based on the percent sequence identity cutoff using the Basic Local Alignment Search Tool (BLAST) algorithm. Fragments above the cutoff (we used a percent nucleotide sequence identity cutoff  $\geq 90\%$ ) found in every original sequence are considered part of the “core” genome, while fragments below the cutoff in at least one strain are considered

part of the “accessory” genome. For the CAGF module, we used the assembled genomes of five biotype 3 strains: VV9-09, VV 4-03, 491771, VVyb1, and BAA87. The NRF module compares an input sequence(s) with a database of sequences and produces an output file of sequences found in the query sequences and absent from the reference strains.

REALPHY (reference sequence alignment based phylogeny builder) (Bertels et al., 2014) was used for the phylogenetic analysis using sequence data from the present study, and sequence data from WGS sequencing projects of eight *V. vulnificus* strains (VVyb1, BAA87, ATCC27562, B2, JY1305, E64MW, JY1701, and ATCC 33147). In addition, we used the complete genome sequence of three reference genomes (CMCP6, YJ016, and M06-24/O). REALPHY infers phylogenetic trees from whole genome sequence data where all provided sequences are mapped to each of the references via bowtie2 (Langmead et al., 2009). From these alignments, phylogenetic trees are inferred via PhyML (phylogenetic estimation using maximum likelihood) (Guindon et al., 2010). The phylogenetic analysis was carried out using the draft whole genomes and then repeated using core and accessory genomes separately. The core and accessory genomes for this analysis were extracted using SPINE and AGEent (Ozer et al., 2014). SPINE identifies a core genome from genomic regions found among all submitted genomes (using default parameters; 100% of all input genomes in which sequence must be present to be considered “core,” and  $\geq 85\%$  identity of nucleotide alignments to be considered homologous). CD-HIT, a program for clustering large datasets of nucleotide or amino acid sequence data, was used for clustering protein sequences sharing sequence similarity among all the human-pathogenic strains and absent from all the environmental strains, using default parameters (amino acid sequence similarity  $\geq 90$ ) (Li and Godzik, 2006; Fu et al., 2012). The clustering algorithm is an incremental clustering algorithm. Briefly, sequences are first sorted in order of decreasing length. The longest sequence becomes the representative of the first cluster. Then, each remaining sequence is compared with the representatives of existing clusters. If the similarity with any representative is above a given threshold (amino acid sequence similarity  $\geq 90$ ), it is grouped into that cluster. Otherwise, a new cluster is defined with that sequence as the representative.

## RESULTS

### SEQUENCING QUALITY AND READ MAPPING TO REFERENCE GENOMES

The quality of base calling from images and sequences was determined by quality score (Q). Approximately (average of 2 runs) 80% of inserts (paired end segments that were sequenced) passed the quality filter (Q = 30), indicating a 99.9% accuracy of base calling at a particular sequence position. Mapping statistics of all the strains to the reference genomes CMCP6 and YJ016 showed, as rather expected, that the human pathogenic strains showed higher rates of uniquely mapped reads to the reference genomes than the environmental strains.

Read mapping of three human-pathogenic biotype 3 strains (VV9-09, VV 4-03, and 491771) and two environmental biotype 1 strains (101/4 and 2322) to the reference genomes CMCP6 and YJ016 showed that between 64 and 66% of the reads obtained

from biotype 3 strains were uniquely mapped (passed quality filter with up to five mismatches), while for the environmental strains only 55–58% of the reads were uniquely mapped. Based on the uniquely mapped reads we extracted the genes that were common to the human-pathogenic strains and absent from the environmental strains (File S1). Altogether there were 176 genes common to the human pathogenic strains, 71 genes (40.3%) encoded hypothetical proteins, while other major groups included genes encoding outer membrane assembly and transcriptional regulators. We identified 43,021 and 47,468 SNPs present in the human pathogenic strains and absent from the environmental strains which were also found in the reference genomes CMCP6 and YJ016, respectively.

The five Israeli samples—three human pathogenic biotype 3 strains and two environmental biotype 1 strains—were subjected to QUASt (Gurevich et al., 2013), a quality assessment tool for genome assemblies, using CMCP6 as a reference genome. We used the assembled genomes of two other biotype 3 strains, BAA87 and VVyb1, which were published by others for comparison purposes (Danin-Poleg et al., 2013; Phillips et al., 2014). The assembled genomes of biotype 3 strains (VV9-09, 491771, VV 4-03, BAA87, and VVyb1) showed similar characteristics and were different from the environmental strains (2322 and 101/04) (Table 2). This WGS project has been deposited at DDBJ/EMBL/GenBank under the accession IDs: JQDW000000000 (VV9-09), JQDV000000000 (VV4-03), JQDU000000000 (491771), JQDT000000000 (101/4), and JQDS000000000 (2322). The versions described in this paper are versions JQDW01000000, JQDV01000000, JQDU01000000, JQDT01000000, and JQDS01000000.

## CORE AND ACCESSORY GENOMES

The core and accessory genomes were studied using Panseq, SPINE, and AGEnt. We included all publicly available genome assemblies in addition to the strains sequenced in the current study and another two biotype 3 strains that were studied by others (VVyb1 and BAA87). We also included the recently published draft genome of biotype 2 strain ATCC 33147. Altogether the analysis included 16 genomes (three complete genomes of the reference strains belonging to the C-genotype and 13 draft genomes). All input sequences from each assembly were concatenated into a single sequence and multiple sequence alignment was produced. The core genome of *V. vulnificus* based on 16 genomes consisted of 3068 genes totaling 3,435,751 bp in size with a GC content of 47.3% (File S2), representing 67% of the whole genome of reference strain CMCP6 (being lowest for the whole genome of biotype 3 strain – VVyb1, 59%, and highest for biotype 1 strain – ATCC 27562, 78%). A total of 214,879 SNPs were identified among the aligned core genomes. The core genome increased as fewer genomes were analyzed reaching 3638 genes when only three genomes representing the three biotypes were used, CMCP6 (biotype 1), ATCC 33147 (biotype 2), and 491771 (biotype 3) (Figure 1). Coding sequences of core and accessory genomes were assigned to functional categories and subcategories using the Clusters of Orthologous Groups of proteins (COG) database (Tatusov et al., 2001) (Figure 2). Nearly 22 and 56% of the core and accessory genomes encoded poorly characterized proteins (including proteins categorized as general function, unknown function, or not in COG database). The average size of the accessory genome of *V. vulnificus* strain was 1662 kbp (being smallest in the clinical strain ATCC 27562, 781 kbp, and largest in

**Table 2 | Quality assessment of assembled genomes.**

Strain	101/04	2322	VV9-09	491771	VV 4-03	BAA87 <sup>a</sup>	VVyb1 <sup>a</sup>
No. of reads	9,786,036	9,150,870	7,465,222	7,291,670	7,944,338	NA	NA
Average length – contigs (bp)	14,853	12,811	13,463	9,774	11,454	23,531	41,029
No. contigs (≥ 0 bp)	370	413	391	543	460	218	140
No. contigs (≥ 1000 bp)	63	78	177	200	190	187	115
GC (%)	46.32	46.4	46.42	46.43	46.42	46.49	46.73
N50	506,351	324,391	54,238	57,095	57,166	52,210	230,903
N75	143,535	161,992	33,564	34,718	33,212	32,756	109,426
No. of fully unaligned contigs	133	162	101	175	139	68	43
Fully unaligned length (bp)	170,338	152,003	326,751	390,311	369,492	284,904	226,696
No. mismatches per 100 kbp <sup>b</sup>	3717.52	3753.33	3024.61	3013.1	3022.73	3019.71	2875.88
No. indels per 100 kbp <sup>c</sup>	73.28	93.49	77.05	80.98	77.7	80.33	59.75
Genome fraction (%) <sup>d</sup>	63.6	61.4	79.8	78.9	79.4	79.9	79.3
No. predicted genes (≥ 0 bp)	5017	4836	4961	5068	5026	4907	5291
No. predicted genes (≥ 300 bp)	4380	4255	4206	4264	4196	4279	4699
No. predicted genes (≥ 1500 bp)	752	722	719	713	714	680	806
No. predicted genes (≥ 3000 bp)	85	84	73	71	72	63	78

N50 is the length for which the collection of all contigs of that length or longer covers at least half an assembly. N75, defined similarly to N50, with 75% instead of 50%.

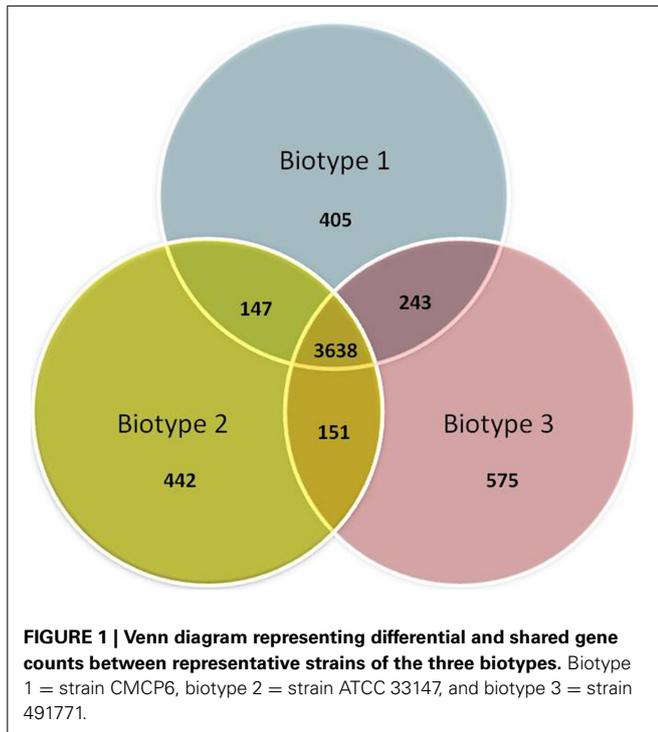
<sup>a</sup>Strains sequenced by others (Danin-Poleg et al., 2013; Phillips et al., 2014).

<sup>b</sup>The average number of mismatches per 100,000 aligned bases. True SNPs and sequencing errors are not distinguished and are counted equally.

<sup>c</sup>The average number of indels per 100,000 aligned bases. Several consecutive single nucleotide indels are counted as one indel.

<sup>d</sup>Percentage of aligned bases in the reference.

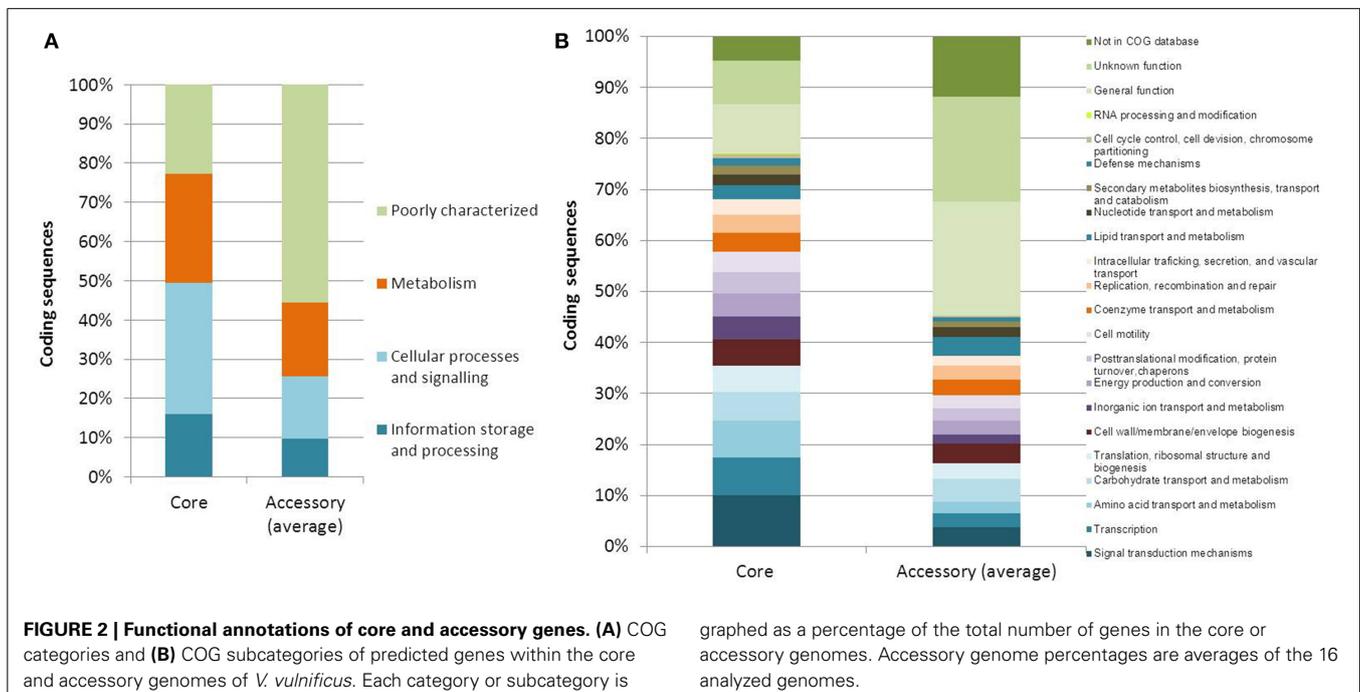
the environmental strain 101/4, 2044 kbp). BLAST alignment of the non-core genomes against the reference clinical strain CMCP6 showed that most of the non-core genomic regions aligned to the large chromosome (data not shown). Accessory element composition was diverse, largely consisting of hypothetical proteins, integrative and conjugative elements, prophages and phage-like elements, transposons, insertion sequences, and integrons.



### CHARACTERISTIC FEATURES OF BIOTYPE 3 GENOMES

Analysis of sequence variance of the *vcg* gene (virulence correlated gene) (Rosche et al., 2005) showed that all biotype 3 strains were classified as E-genotype. In addition, two environmental biotype 1 strains 101/04 and 2322, biotype 2 strain (ATCC 33147), and two clinical biotype 1 strains, ATCC27562 and B2, were all E-genotype. Clustering and phylogenetic analysis of biotype 3 strains showed that strains VV9-09, VV 4-03, VVy1, and BAA87 were genetically more closely related and clustered separately from strain 491771. Applying the NRF module of Panseq to search for sequences present in strain 491771 and absent from the other biotype 3 strains resulted in 114,855 bp sequences that resolved into 127 genes, of which the vast majority encoded hypothetical proteins and phage-related proteins (File S3). Biotype 3 strains shared between 85.4 and 89.5% of their genomes with the genomes of the rest of the strains, being highest for the clinical strains CMCP6, YJ016, MO6-24/O, ATCC 27562, and B2 and lowest for the genomes of the environmental strains 2322 and 101/4.

Next we searched for accessory genomic regions found in biotype 3 strains and absent from the core genome of all the strains. Annotation of these accessory genomic sequences yielded 1732 genes, of which the vast majority encoded hypothetical proteins, phage-related proteins, and mobile element proteins (File S4). Several genes encoded plasmid conjugative transfer proteins that displayed high sequence similarity to parts of the genomes of *V. vulnificus* pR99 plasmid, *Vibrio* sp. 04Ya090 plasmid pAQU2, *V. vulnificus* pC4602-1 plasmid, *Vibrio cholerae* plasmid pVC, and *Vibrio* phage kappa proviral DNA. Significant sequence alignment was found to nearly the entire length of *V. cholerae* O139 class 4 integron genes for hypothetical protein, *Vibrio parahaemolyticus* HTH gene for HTH-domain protein, *V. vulnificus* super-integron, and insertion sequence ISVpa2 KX-V237 found in many *vibrio*



species including *V. parahaemolyticus* O3:K6, *V. vulnificus*, *Vibrio penaeicida*, and *Vibrio splendidus*.

### PHYLOGENETIC ANALYSIS

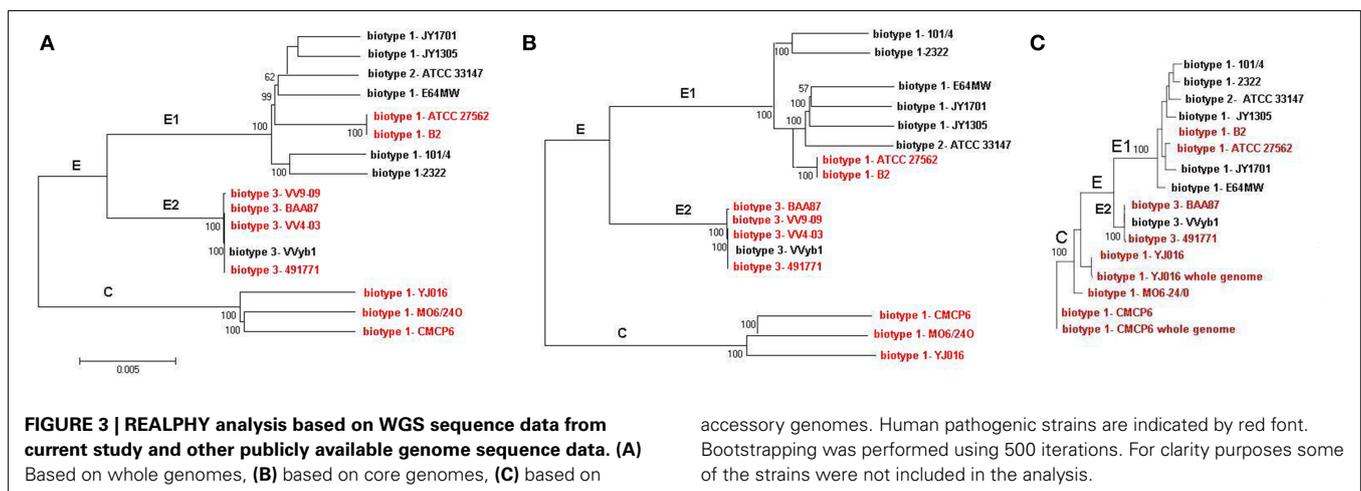
Phylogenetic analysis was carried out based on whole genomes and then repeated using core and accessory genomes separately. Both analyses showed similar phylogenetic relationships with two distinct clusters consistent with the division into C and E genotype paradigm where the clinical reference genomes clustered separately from the rest of the strains (**Figure 3**). Two groups were identified within the E cluster, the first, designated E1, included five biotype 1 environmental strains, two biotype 1 clinical strains (ATCC 27562 and B2), and an eel-pathogenic biotype 2 strain (ATCC 33147). All biotype 3 strains clustered separately within the E-cluster and formed another group designated E2 (**Figure 3A**). Repeating the analysis using core genomes (**Figure 3B**) and accessory genomes (**Figure 3C**) resulted in similar clustering pattern with minor differences in grouping the strains within the clusters. BLAST alignment of the core genomes of groups E1, E2, and cluster C against the complete genome sequence of clinical reference strain CMCP6 showed that the core genome of biotype 3 strains (E2) was the largest and shared more regions with the core genome of cluster C than with group E1 (**Figure 4**). The genomic regions found in group E2 and absent from group E1 consisted of 259 contigs totaling 1,407,644 bp in size, from which 1273 genes were identified. The largest group consisted of genes encoding hypothetical proteins (32%), while other groups were made up of phage-related proteins, mobile elements proteins, genes involved in DNA metabolism, and genes encoding membrane transport proteins.

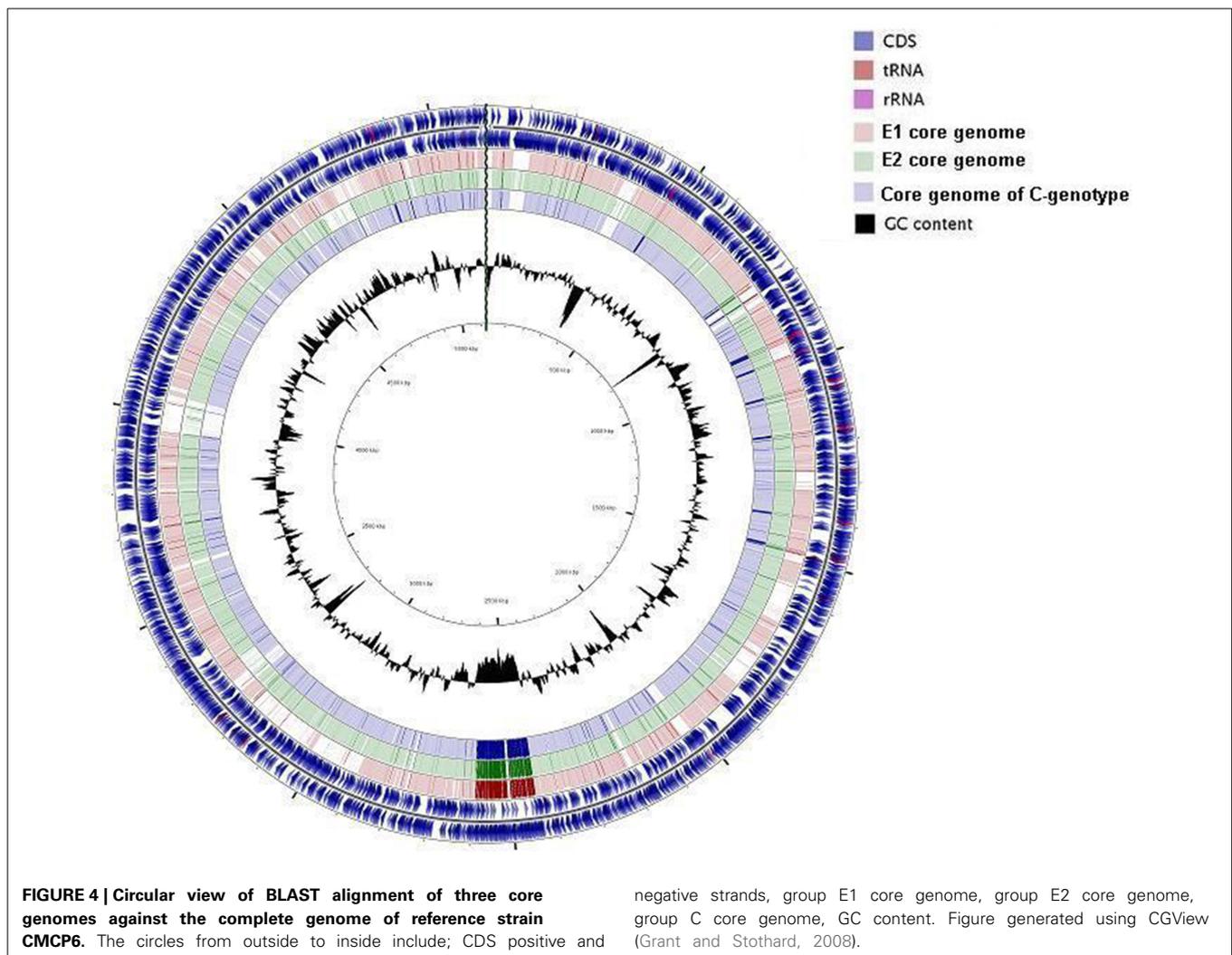
In order to characterize the set of sequences shared by the human-pathogenic strains and absent from environmental strains we compared each and every human-pathogenic strain (VV9-09, VV 4-03, 491771 from the present study, complete genomes of three reference human pathogenic strains; CMCP6, YJ016, MO6-24/O, and WGS sequence data from strains BAA87, ATCC275962, and B2) against all five environmental strains (101/04, 2322, E64MW, JY1701, and JY1305 – E1

group in **Figure 3**). A total of 1916 sequences/proteins (including 713 hypothetical proteins) were present in the human pathogenic strains and absent from the environmental strains. Clustering analysis of the non-hypothetical proteins ( $n = 1203$ ) revealed 148 protein clusters shared by the human-pathogenic strains (File S5). The proteins shared among the pathogenic strains included transcriptional regulators, arylsulfatases, methyl-accepting chemotaxis proteins, acetyltransferases, GGDEF family proteins, transposases, type IV secretory system (T4SS) proteins, and integrases. The 713 hypothetical proteins resolved into 319 clusters consisting largely of proteins shared by the same phylogenetic group: that is clusters unique to the reference genomes, clusters unique to biotype 3 strains, and clusters unique to strains ATCC 27562 and B2 present within E-genotype population, probably indicative of their common ancestral origin. Very few hypothetical protein clusters were shared by all the human pathogenic strains.

### DISCUSSION

We used sequence data from 16 genomes (draft and complete) to study the pan genome of *V. vulnificus* and investigate the evolutionary relationship of biotype 3 with existing populations. The core genome of *V. vulnificus* consisted of 3068 genes comprising nearly two-thirds of the set of genes of the complete genome sequence of three reference genomes. All biotype 3 strains resolved into the E-genotype cluster yet formed a distinct group from the rest of the strains. The core genome of biotype 3 strains (group E2) was the largest among the dataset, consistent with its recent evolution and highly clonal nature as previously described (Bisharat et al., 2007a), while the core genome of group E1 was the smallest consistent with the genetic diversity of environmental populations. Annotation of the genomic regions found in the core genome of biotype 3 strains and absent from the core genome of group E1 strains showed that nearly a third of the genes encoded hypothetical proteins in addition to many phage related and mobile element proteins. The core genome of biotype 3 strains (E2) shared more regions with the core genome of the clinical reference strains (cluster C) than with the core genome of the other environmental group (E1) (**Figure 4**). This





is likely due to the relatively small number of strains included in the analysis (three reference clinical strains representing cluster C and five strains representing group E2, while group E1 was represented by eight strains). Using an equal and rather large number of strains from each group, would have probably resulted in minor changes in the size of the core genome of each group and largely representing the core genome of the species.

Phylogenetic analysis divided the dataset into two distinct populations, in agreement with previous studies (Gutacker et al., 2003; Rosche et al., 2005; Bisharat et al., 2007b). This structuring pattern of *V. vulnificus* populations into two distinct clusters has been observed across housekeeping genes (Bisharat et al., 2005), genes encoding outer membrane proteins (Bisharat et al., 2007a), 16s rRNA genes (Aznar et al., 1994), and across whole genomes (current study). In addition, these observations were confirmed in various geographical regions, USA (Rosche et al., 2005; Warner and Oliver, 2008; Reynaud et al., 2013), Baltic Sea region (Bier et al., 2013), South and Southeast Asia (Mahmud et al., 2008, 2010), and the Middle East (Bisharat et al., 2007b). Furthermore, inferred phylogeny from current study based on

negative strands, group E1 core genome, group E2 core genome, group C core genome, GC content. Figure generated using CGView (Grant and Stothard, 2008).

whole, core, or accessory genomes resulted in nearly identical clustering pattern. Overall these findings strongly imply that *V. vulnificus* populations may have diverged into two main clusters in ancient times. Nevertheless, the emergence of a new genotype with unique phenotypic profile in Israel, suggested that a new genetically distinct population may have existed and affected public health due to human behavior (Bisharat and Raz, 1996). Subsequent molecular analysis suggested that this genotype may have emerged due to recombination between the two distinct populations (Bisharat et al., 2005). Data from current study suggests that biotype 3 is a clone that diverged from the parental population, cluster E. This clonal lineage may have emerged and succeeded due to the acquisition of a strong selective advantage, allowing it to rise in frequency in the population and eventually creating a distinct lineage (Smith et al., 1993). The genetic distinction of group E2 (biotype 3) from the parent population and from the C-genotype cluster was not driven by the accessory genome as core-genome and accessory-genome based analysis exhibited similar phylogenetic relationships. We have previously emphasized the role of recombination and mutation in generating genetic diversity within *V. vulnificus* and suggested that it

may have impacted the evolution of biotype 3 strains (Bisharat et al., 2007b). Nevertheless, it seems that the accessory genome has likely evolved together with the core genome in view of the fixed clustering pattern observed over different genomic levels of analysis.

A yet unanswered question is why human disease in Israel is almost entirely caused by biotype 3 (Bisharat et al., 2005; Zaidenstein et al., 2008). Previous studies by our group and others have identified strains, exhibiting non-biotype 3 genotype, belonging to both clusters in the environment (Bisharat et al., 2007b; Broza et al., 2009). Is it because biotype 3 has better fitness or greater virulent potential than other genotypes, or is it because of higher frequencies within the water or fish? Previous environmental surveys conducted in Israel in fish farms during 2004–2006 have shown that biotype 3 frequencies ranged from 2 to 21% of environmental populations which largely consisted of biotype 1 strains (Broza et al., 2009). We speculate that biotype 3 genotype is a subtype of the E-cluster that is of greater virulent potential than other E or C genotypes circulating in fish farms. Luckily, *V. vulnificus* disease burden in Israel have decreased dramatically in the past 10 years and it's restricted now to fish farm associated activities (farming, fishing, and maintenance). In addition, the population at risk has decreased dramatically due to preventive measures aimed to increase the awareness of the population to this dreadful pathogen and decrease risk of infection by changing fish marketing policies (Bisharat, 2002).

The list of protein clusters shared by the human pathogenic strains and absent from the environmental strains is similar to reports published by others in recent years (Gulig et al., 2010; Morrison et al., 2012). Nearly all the genes/proteins shared by the human pathogenic biotype 3 strains were found in all the strains tested, while not all the listed genes were found in all the clinical reference genomes, suggesting that some genes are not entirely involved in virulence and may have other functional roles.

The epidemiology of *V. vulnificus* disease in Israel was unique and unexpected, and the only common source outbreak reported to date. It erupted due to changes in fish marketing policies where tilapia fish were sold live in freshwater instead of dead and packed in ice (Bisharat and Raz, 1996). Had this not occurred, the emergence of biotype 3 may have gone undetected, causing only sporadic cases as occurred in some European countries (Melhus et al., 1995; Dalsgaard et al., 1996; Garcia Cuevas et al., 1998; Horre et al., 1998; Torres et al., 2002; Mitra, 2004; Frank et al., 2006).

Our study showed that *V. vulnificus* biotype 3 is a distinct clone that have descended from the parental environmental population and may have acquired pathogenic potential by lateral gene transfer from other *vibrios* thus enabling a harmless environmental species to cause disease in humans. The unique epidemiological circumstances facilitated disease outbreak and brought this genotype to the attention of the scientific community. These novel observations reveal yet another way by which epidemic organisms arise.

## AUTHOR CONTRIBUTIONS

Yael Koton wrote the initial draft, Michal Gordon and Vered Chalifa-Caspi carried out part of the bioinformatics analysis.

Naiel Bisharat planned, coordinated, analyzed data, and wrote the manuscript.

## ACKNOWLEDGMENTS

This work was supported by a grant from the Legacy Heritage Foundation. Grant No. 3-2010.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2014.00803/abstract>

## REFERENCES

- Andrews, S. (2014). *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (Accessed December 11, 2014).
- 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. doi: 10.1186/1471-2164-9-75
- Aznar, R., Ludwig, W., Amann, R. I., and Schleifer, K. H. (1994). Sequence determination of rRNA genes of *Vibrio* species and whole-cell identification of *Vibrio vulnificus* with rRNA-targeted oligonucleotide probes. *Int. J. Syst. Bacteriol.* 44, 330–337. doi: 10.1099/00207173-44-2-330
- Bertels, F., Silander, O. K., Pachkov, M., Rainey, P. B., and van Nimwegen, E. (2014). Automated reconstruction of whole-genome phylogenies from short-sequence reads. *Mol. Biol. Evol.* 31, 1077–1088. doi: 10.1093/molbev/msu088
- Bier, N., Bechlers, S., Diescher, S., Klein, F., Hauk, G., Duty, O., et al. (2013). Genotypic diversity and virulence characteristics of clinical and environmental *Vibrio vulnificus* isolates from the Baltic Sea region. *Appl. Environ. Microbiol.* 79, 3570–3581. doi: 10.1128/AEM.00477-13
- Bisharat, N. (2002). *Vibrio* infections can be avoided. *IMAJ* 4, 631–633.
- Bishrat, N. (2010). "Population genetics of vibrios," in *Bacterial Population Genetics in Infectious Diseases*, eds A. Robinson, D. Falush, and E. Feil (Hoboken, NJ: John Wiley & Sons), 378–401.
- Bisharat, N., Agmon, V., Finkelstein, R., Raz, R., Ben-Dror, G., Lerner, L., et al. (1999). Clinical, epidemiological, and microbiological features of *Vibrio vulnificus* biogroup 3 causing outbreaks of wound infection and bacteraemia in Israel. Israel *Vibrio* Study Group. *Lancet* 354, 1421–1424. doi: 10.1016/S0140-6736(99)02471-X
- Bisharat, N., Amaro, C., Fouz, B., Llorens, A., and Cohen, D. I. (2007a). Serological and molecular characteristics of *Vibrio vulnificus* biotype 3: evidence for high clonality. *Microbiology* 153, 847–856. doi: 10.1099/mic.0.2006/003723-0
- Bisharat, N., Cohen, D. I., Harding, R. M., Falush, D., Crook, D. W., Peto, T., et al. (2005). Hybrid *Vibrio vulnificus*. *Emerg. Infect. Dis.* 11, 30–35. doi: 10.3201/eid1101.040440
- Bisharat, N., Cohen, D. I., Maiden, M. C., Crook, D. W., Peto, T., and Harding, R. M. (2007b). The evolution of genetic structure in the marine pathogen, *Vibrio vulnificus*. *Infect. Genet. Evol.* 7, 685–693. doi: 10.1016/j.meegid.2007.07.007
- Bisharat, N., and Raz, R. (1996). *Vibrio* infection in Israel due to changes in fish marketing. *Lancet* 348, 1585–1586. doi: 10.1016/S0140-6736(05)66199-5
- Bock, T., Christensen, N., Eriksen, N. H., Winter, S., Rygaard, H., and Jorgensen, F. (1994). The first fatal case of *Vibrio vulnificus* infection in Denmark. *APMIS* 102, 874–876. doi: 10.1111/j.1699-0463.1994.tb05247.x
- Broza, Y. Y., Danin-Poleg, Y., Lerner, L., Valinsky, L., Broza, M., and Kashi, Y. (2009). Epidemiologic study of *Vibrio vulnificus* infections by using variable number tandem repeats. *Emerg. Infect. Dis.* 15, 1282–1285. doi: 10.3201/eid1508.080839
- Chiang, S. R., and Chuang, Y. C. (2003). *Vibrio vulnificus* infection: clinical manifestations, pathogenesis, and antimicrobial therapy. *J. Microbiol. Immunol. Infect.* 36, 81–88.
- Chuang, Y. C., Yuan, C. Y., Liu, C. Y., Lan, C. K., and Huang, A. H. (1992). *Vibrio vulnificus* infection in Taiwan: report of 28 cases and review of clinical manifestations and treatment. *Clin. Infect. Dis.* 15, 271–276. doi: 10.1093/clinids/15.2.271
- Colodner, R., Chazan, B., Kopelowitz, J., Keness, Y., and Raz, R. (2002). Unusual portal of entry of *Vibrio vulnificus*: evidence of its prolonged survival on the skin. *Clin. Infect. Dis.* 34, 714–715. doi: 10.1086/338871

- Colodner, R., Raz, R., Meir, I., Lazarovich, T., Lerner, L., Kopelowitz, J., et al. (2004). Identification of the emerging pathogen *Vibrio vulnificus* biotype 3 by commercially available phenotypic methods. *J. Clin. Microbiol.* 42, 4137–4140. doi: 10.1128/JCM.42.9.4137-4140.2004
- Dalsgaard, A., Frimodt-Moller, N., Bruun, B., Hoi, L., and Larsen, J. L. (1996). Clinical manifestations and molecular epidemiology of *Vibrio vulnificus* infections in Denmark. *Eur. J. Clin. Microbiol. Infect. Dis.* 15, 227–232. doi: 10.1007/BF01591359
- Danin-Poleg, Y., Elgavish, S., Raz, N., Efimov, V., and Kashi, Y. (2013). Genome sequence of the pathogenic bacterium *Vibrio vulnificus* biotype 3. *Genome Announc.* 1:e0013613. doi: 10.1128/genomeA.00136-13
- Delcher, A. L., Kasif, S., Fleischmann, R. D., Peterson, J., White, O., and Salzberg, S. L. (1999). Alignment of whole genomes. *Nucleic. Acids Res.* 27, 2369–2376. doi: 10.1093/nar/27.11.2369
- Efimov, V., Danin-Poleg, Y., Raz, N., Elgavish, S., Linetsky, A., and Kashi, Y. (2013). Insight into the evolution of *Vibrio vulnificus* biotype 3's genome. *Front. Microbiol.* 4:393. doi: 10.3389/fmicb.2013.00393
- Frank, C., Littman, M., Alpers, K., and Hallauer, J. (2006). *Vibrio vulnificus* wound infections after contact with the Baltic Sea, Germany. *Euro. Surveill.* 11:E060817.1.
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W., (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565
- Garcia Cuevas, M., Collazos Gonzalez, J., Martinez Gutierrez, E., and Mayo Suarez, J. (1998). *Vibrio vulnificus* septicemia in Spain. *An. Med. Interna.* 15, 485–486.
- Gonzalez-Escalona, N., Jaykus, L. A., and Depaola, A. (2007). Typing of *Vibrio vulnificus* strains by variability in their 16S-23S rRNA intergenic spacer regions. *Foodborne Pathog. Dis.* 4, 327–337. doi: 10.1089/fpd.2007.0005
- Grant, J. R., and Stothard, P. (2008). The CGView server: a comparative genomics tool for circular genomes. *Nucleic. Acids Res.* 36, W181–W184. doi: 10.1093/nar/gkn179
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Gulig, P. A., de Crecy-Lagard, V., Wright, A. C., Walts, B., Telonis-Scott, M., and McIntyre, L. M. (2010). SOLiD sequencing of four *Vibrio vulnificus* genomes enables comparative genomic analysis and identification of candidate clade-specific virulence genes. *BMC Genomics* 11:512. doi: 10.1186/1471-2164-11-512
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086
- Gutacker, M., Conza, C., Benagli, A., Pedrolì, M., Bernasconi, V., Permin, L., et al. (2003). Population genetics of *Vibrio vulnificus*: identification of two divisions and a distinct eel-pathogenic clone. *Appl. Environ. Microbiol.* 69, 3203–3212. doi: 10.1128/AEM.69.6.3203-3212.2003
- Horre, R., Becker, S., Marklein, G., Shimada, T., Stephan, R., Steuer, K., et al. (1998). Necrotizing fasciitis caused by *Vibrio vulnificus*: first published infection acquired in Turkey is the second time a strain is isolated in Germany. *Infection* 26, 399–401. doi: 10.1007/BF02770844
- Klontz, K. C., Lieb, S., Schreiber, M., Janowski, H. T., Baldy, L. M., and Gunn, R. A. (1988). Syndromes of *Vibrio vulnificus* infections. Clinical and epidemiologic features in Florida cases. *Ann. Intern. Med.* 109, 318–323. doi: 10.7326/0003-4819-109-4-318
- Koton, Y., Eghbaria, S., Gordon, M., Chalifa-Caspi, V., and Bisharat, N. (2014). Draft genome sequence of fish pathogenic *Vibrio vulnificus* biotype 2. *Genome Announc.* 2:e01224-14. doi: 10.1128/genomeA.01224-14
- Kumamoto, K. S., and Vukich, D. J. (1998). Clinical infections of *Vibrio vulnificus*: a case report and review of the literature. *J. Emerg. Med.* 16, 61–66. doi: 10.1016/S0736-4679(97)00230-8
- 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
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158
- Mahmud, Z., Neogi, S., Kassu, A., Mai Huong, B., Jahid, I., Islam, M., et al. (2008). Occurrence, seasonality and genetic diversity of *Vibrio vulnificus* in coastal seaweeds and water along the Kii channel, Japan. *FEMS Microbiol. Ecol.* 64, 209–218. doi: 10.1111/j.1574-6941.2008.00460.x
- Mahmud, Z., Wright, A., Mandal, S., Dai, J., Jones, M., Hasan, M., et al. (2010). Genetic characterization of *Vibrio vulnificus* strains from tilapia aquaculture in Bangladesh. *Appl. Environ. Microbiol.* 76, 4890–4895. doi: 10.1128/AEM.00636-10
- Matsumoto, K., Ohshige, K., Fujita, N., Tomita, Y., Mitsumizo, S., Nakashima, M., et al. (2010). Clinical features of *Vibrio vulnificus* infections in the coastal areas of the Ariake Sea, Japan. *J. Infect. Chemother.* 16, 272–279. doi: 10.1007/s10156-010-0050-Z
- Melhus, A., Holmdahl, T., and Tjernberg, I. (1995). First documented case of bacteremia with *Vibrio vulnificus* in Sweden. *Scand. J. Infect. Dis.* 27, 81–82. doi: 10.3109/00365549509018980
- Miron, D., Lev, A., Colodner, R., and Merzel, Y. (2003). *Vibrio vulnificus* necrotizing fasciitis of the calf presenting with compartment syndrome. *Pediatr. Infect. Dis. J.* 22, 666–668. doi: 10.1097/01.inf.0000076546.27263.15
- Mitra, A. K. (2004). *Vibrio vulnificus* infection: epidemiology, clinical presentations, and prevention. *South Med. J.* 97, 118–119. doi: 10.1097/01.SMJ.0000092520.47509.C2
- Morrison, S. S., Williams, T., Cain, A., Froelich, B., Taylor, C., Baker-Austin, C., et al. (2012). Pyrosequencing-based comparative genome analysis of *Vibrio vulnificus* environmental isolates. *PLoS ONE* 7:e37553. doi: 10.1371/journal.pone.0037553
- Nilsson, W. B., Paranjypte, R. N., Depaola, A., and Strom, M. S. (2003). Sequence polymorphism of the 16S rRNA gene of *Vibrio vulnificus* is a possible indicator of strain virulence. *J. Clin. Microbiol.* 41, 442–446. doi: 10.1128/JCM.41.1.442-446.2003
- Oliver, J. D. (1989). “*Vibrio vulnificus*,” in *Food-Borne Bacterial Pathogens*, ed M. P. Doyle (New York, NY: Marcel Dekker, Inc.), 569–600.
- Oliver, J. D. (2006). “*Vibrio vulnificus*,” in *The Biology of Vibrios*, eds F. L. Thompson, B. Austin, and J. G. Swings (Washington, DC: ASM Press), 349–366. doi: 10.1128/9781555815714.ch25
- Ozer, E. A., Allen, J., and Hauser, A. (2014). Characterization of the core and accessory genomes of *Pseudomonas aeruginosa* using bioinformatic tools Spine and AGEnt. *BMC Genomics.* 15:737 doi: 10.1186/1471-2164-15-737
- Park, S. D., Shon, H. S., and Joh, N. J. (1991). *Vibrio vulnificus* septicemia in Korea: clinical and epidemiologic findings in seventy patients. *J. Am. Acad. Dermatol.* 24, 397–403. doi: 10.1016/0190-9622(91)70059-B
- Phillips, K. E., Schipma, M. J., and Satchell, K. J. (2014). Draft genome sequence of Israeli outbreak-associated *Vibrio vulnificus* biotype 3 clinical isolate BAA87. *Genome Announc.* 2:e00032-14. doi: 10.1128/genomeA.00032-14
- Reynaud, Y., Pitchford, S., de Decker, S., Wilkfors, G. H., and Brown, C. L. (2013). Molecular typing of environmental and clinical strains of *Vibrio vulnificus* isolated in the northeastern USA. *PLoS ONE* 8:e83357. doi: 10.1371/journal.pone.0083357
- Rosche, T. M., Yano, Y., and Oliver, J. D. (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. doi: 10.1111/j.1348-0421.2005.tb03731.x
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., et al. (2000). Artemis: sequence visualization and annotation. *Bioinformatics* 16, 944–945. doi: 10.1093/bioinformatics/16.10.944
- Smith, J. M., Smith, N. H., O'Rourke, M., and Spratt, B. G. (1993). How clonal are bacteria? *Proc. Natl. Acad. Sci. U.S.A.* 90, 4384–4388 doi: 10.1073/pnas.90.10.4384
- Tacket, C. O., Brenner, F., and Blake, P. A. (1984). Clinical features and an epidemiological study of *Vibrio vulnificus* infections. *J. Infect. Dis.* 149, 558–561. doi: 10.1093/infdis/149.4.558

- Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., et al. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 29, 22–28. doi: 10.1093/nar/29.1.22
- Tison, D. L., Nishibuchi, M., Greenwood, J. D., and Seidler, R. (1982). *Vibrio vulnificus* biogroup 2: new biogroup pathogenic for eels. *Appl. Environ. Microbiol.* 44, 640–646.
- Torres, L., Escobar, S., Lopez, A. I., Marco, M. L., and Pobo, V. (2002). Wound infection due to *Vibrio vulnificus* in Spain. *Eur. J. Clin. Microbiol. Infect. Dis.* 21, 537–538. doi: 10.1007/s10096-002-0767-4
- Warner, E., and Oliver, J. D. (2008). Population structures of two genotypes of *Vibrio vulnificus* in oysters (*Crassostrea virginica*) and seawater. *Appl. Environ. Microbiol.* 74, 80–85. doi: 10.1128/AEM.01434-07
- Zaidenstein, R., Sadik, C., Lerner, L., Valinsky, L., Kopelowitz, J., Yishai, R., et al. (2008). Clinical characteristics and molecular subtyping of *Vibrio vulnificus* illnesses, Israel. *Emerg. Infect. Dis.* 14, 1875–1882. doi: 10.3201/eid1412.080499
- 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.
- Received: 12 October 2014; accepted: 29 December 2014; published online: 15 January 2015.
- Citation: Koton Y, Gordon M, Chalifa-Caspi V and Bisharat N (2015) Comparative genomic analysis of clinical and environmental *Vibrio vulnificus* isolates revealed biotype 3 evolutionary relationships. *Front. Microbiol.* 5:803. doi: 10.3389/fmicb.2014.00803
- This article was submitted to *Evolutionary and Genomic Microbiology*, a section of the journal *Frontiers in Microbiology*.
- Copyright © 2015 Koton, Gordon, Chalifa-Caspi and Bisharat. 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) or licensor 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.