



Evolutionary Model of Cluster Divergence of the Emergent Marine Pathogen *Vibrio vulnificus*: From Genotype to Ecotype

Mario López-Pérez,^a Jane M. Jayakumar,^a Jose M. Haro-Moreno,^b Asier Zaragoza-Solas,^b Geethika Reddi,^a Francisco Rodriguez-Valera,^b Orr H. Shapiro,^c Munirul Alam,^d  Salvador Almagro-Moreno^{a,e}

^aBurnett School of Biomedical Sciences, College of Medicine, University of Central Florida, Orlando, Florida, USA

^bEvolutionary Genomics Group, División de Microbiología, Universidad Miguel Hernández, Alicante, Spain

^cInstitute for Postharvest and Food Sciences, Volcani Research Center, Rishon LeZion, Israel

^dMolecular Ecology and Metagenomics Laboratory, International Center for Diarrheal Disease Research, Dhaka, Bangladesh

^eNational Center for Integrated Coastal Research, University of Central Florida, Orlando, Florida, USA

ABSTRACT *Vibrio vulnificus*, an opportunistic pathogen, is the causative agent of a life-threatening septicemia and a rising problem for aquaculture worldwide. The genetic factors that differentiate its clinical and environmental strains remain enigmatic. Furthermore, clinical strains have emerged from every clade of *V. vulnificus*. In this work, we investigated the underlying genomic properties and population dynamics of the *V. vulnificus* species from an evolutionary and ecological point of view. Genome comparisons and bioinformatic analyses of 113 *V. vulnificus* isolates indicate that the population of *V. vulnificus* is made up of four different clusters. We found evidence that recombination and gene flow between the two largest clusters (cluster 1 [C1] and C2) have drastically decreased to the point where they are diverging independently. Pangenome and phenotypic analyses showed two markedly different lifestyles for these two clusters, indicating commensal (C2) and bloomer (C1) ecotypes, with differences in carbohydrate utilization, defense systems, and chemotaxis, among other characteristics. Nonetheless, we identified frequent intra- and interspecies exchange of mobile genetic elements (e.g., antibiotic resistance plasmids, novel “chromids,” or two different and concurrent type VI secretion systems) that provide high levels of genetic diversity in the population. Surprisingly, we identified strains from both clusters in the mucosa of aquaculture species, indicating that manmade niches are bringing strains from the two clusters together. We propose an evolutionary model of *V. vulnificus* that could be broadly applicable to other pathogenic vibrios and facultative bacterial pathogens to pursue strategies to prevent their infections and emergence.

IMPORTANCE *Vibrio vulnificus* is an emergent marine pathogen and is the cause of a deadly septicemia. However, the genetic factors that differentiate its clinical and environmental strains and its several biotypes remain mostly enigmatic. In this work, we investigated the underlying genomic properties and population dynamics of the *V. vulnificus* species to elucidate the traits that make these strains emerge as a human pathogen. The acquisition of different ecological determinants could have allowed the development of highly divergent clusters with different lifestyles within the same environment. However, we identified strains from both clusters in the mucosa of aquaculture species, indicating that manmade niches are bringing strains from the two clusters together, posing a potential risk of recombination and of emergence of novel variants. We propose a new evolutionary model that provides a perspective that could be broadly applicable to other pathogenic vibrios and facultative bacterial pathogens to pursue strategies to prevent their infections.

Citation López-Pérez M, Jayakumar JM, Haro-Moreno JM, Zaragoza-Solas A, Reddi G, Rodriguez-Valera F, Shapiro OH, Alam M, Almagro-Moreno S. 2019. Evolutionary model of cluster divergence of the emergent marine pathogen *Vibrio vulnificus*: from genotype to ecotype. *mBio* 10:e02852-18. <https://doi.org/10.1128/mBio.02852-18>.

Editor Michael T. Laub, Massachusetts Institute of Technology

Copyright © 2019 López-Pérez 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 Salvador Almagro-Moreno, samoreno@ucf.edu.

This article is a direct contribution from a Fellow of the American Academy of Microbiology. Solicited external reviewers: Paul Jensen, University of California, San Diego; Laura Gómez-Consarnau, University of Southern California.

Received 20 December 2018

Accepted 9 January 2019

Published 19 February 2019

KEYWORDS *Vibrio vulnificus*, emergence, evolution, genomics, waterborne pathogens

The family *Vibrionaceae* encompasses a ubiquitous group of moderately halophilic bacteria that are natural inhabitants of marine and brackish environments (1). Over the past decades, the number of *Vibrio*-related human infections rose steadily, with a similar increase observed in *Vibrio* infections in aquaculture environments (2–5). This rise in *Vibrio* virulence and pathogenicity is often attributed to the ongoing increase in sea surface temperatures associated with climate change (6, 7). Indeed, the distribution and geographical range of these opportunistic pathogens has been gradually widening, with outbreaks of *Vibrio* infections reported at latitudes as high as the Baltic Sea (4) or Alaska (8), previously considered too cold for *Vibrio* to thrive.

While the majority of the more than 100 described *Vibrio* species are harmless to humans, several species have emerged as opportunistic human pathogens, most notably *Vibrio cholerae*, *V. parahaemolyticus*, and *V. vulnificus* (9, 10). *Vibrio* infections are associated with a wide range of diseases and symptoms ranging from cholera, and other gastrointestinal infections, to necrotizing fasciitis and acute septicemia (1). *Vibrio* infections occur through the consumption of contaminated water or of raw or undercooked seafood or through exposure of open wounds to seawater (11, 12). According to CDC reports, an estimated 80,000 illnesses, 500 hospitalizations, and 100 deaths in the United States occur annually due to *Vibrio* infections (13). The increase in prevalence of *Vibrio* infections in the United States is unique among foodborne pathogens, with all infections associated with other major pathogens such as *Salmonella*, *Escherichia coli*, *Campylobacter*, *Listeria*, or *Shigella* steadily decreasing over the same time period (14).

Among vibrios, *V. vulnificus* has gained particular notoriety as an opportunistic and often fatal human pathogen and as an emergent pathogen of aquaculture species (12, 15). Specifically, *V. vulnificus* infects humans through consumption of raw seafood, causing severe gastroenteritis, or by direct contact of an open wound with seawater, producing wound infections, leading to necrotizing fasciitis or primary septicemia (16, 17). *V. vulnificus* is responsible for up to 94% of noncholera *Vibrio*-related deaths (12). Most deaths occur in patients with preexisting conditions, such as a compromised immune system or elevated serum iron levels (primarily in alcohol-associated liver cirrhosis), where primary septicemia may lead to mortality rates of over 50% (18, 19).

Strains of *V. vulnificus* are currently subdivided into three biotypes based on their biochemical characteristics and phylogeny (20, 21). Biotype 1 is associated with most of the clinical infections. Biotype 2 is primarily considered a pathogen of aquaculture-raised species, particularly eels, but is also found in association with human infections (22). Biotype 3, the smallest and most recently discovered, is thus far limited to Israel, where it caused a serious outbreak of wound infections among fish farmers and consumers of tilapia grown in aquaculture (12). A recent classification based on phylogenetic lineages broadly matches the biotype classification (23).

In contrast to *V. cholerae*, where all strains capable of causing cholera belong to a single clade, a comparison between *V. vulnificus* strains reveals a more complex pattern in the distribution of its clinical strains (24–27). Clinical *V. vulnificus* strains encompass a large range of genomic diversity (12, 28, 29) and lack specific markers that can be used to clearly distinguish isolates with pathogenic potential (16). Despite possessing a wide range of essential virulence factors (e.g., capsular polysaccharide [CPS], iron acquisition, cytotoxicity systems, etc.), the general consensus is that virulence in *V. vulnificus* is mostly host dependent (12, 16, 30). Nonetheless, it remains puzzling why most of the clinical strains isolated to date belong to biotype 1.

In this study, we aimed to gain insights into the emergence of this enigmatic human pathogen by understanding the evolutionary differences at the population, genomic, and phenotypic levels that differentiate strains belonging to the different *V. vulnificus* biotypes. To this end, we performed the most comprehensive computational analysis

of *V. vulnificus* to date. For our analysis, we compared all *V. vulnificus* genomes currently available in public databases, comprising a total of 113 worldwide isolates from various habitats and hosts collected over a period of 40 years. We analyzed these genomes using a diverse suite of bioinformatic tools and performed phenotypic analyses of representative strains in order to infer the mechanistic processes driving their evolution. We identified four major clusters: cluster 1 (C1) to cluster 4 (C4). We show that the two largest and most divergent ones (C1 and C2) are adapted to different lifestyles that may include behavioral barriers leading to speciation. Nonetheless, frequent exchange of mobile genetic elements (MGEs) across family barriers occurs. Surprisingly, we identified strains from both C1 and C2 cohabitating in the mucosa of eels from aquaculture farms, which raises the concern of manmade environments bringing strains of these two clusters together. Overall, our findings shed light on the underlying genomic properties that are required for the emergence of pathogenic *V. vulnificus* strains and determine their host range. Information derived from our results may be applied to develop novel strategies for the prevention of future infections in aquaculture environments and subsequent spread to human hosts.

RESULTS

Phylogenomic and population structure of *V. vulnificus*. In order to investigate the evolutionary changes that led to the divergent expansion of *V. vulnificus*, we compared 113 publicly available genomes in GenBank, complete or draft, using a phylogenomic tree based on both single nucleotide polymorphisms (SNPs) and average nucleotide identity (ANI). A total of 27,366 SNPs were identified among the aligned core genomes of 113 strains recovered from a wide range of geographical and ecological sources (see Table S1 in the supplemental material) to produce a phylogenomic tree. Using both approximations, all strains were clustered into four groups with ANI values of >97% (Fig. 1A; see also Fig. S1 in the supplemental material), here referred to as clusters 1 to 4 (C1 to C4) for simplicity.

The genetic population structure, inferred based on the pattern of SNPs shared among the strains (Fig. 1B), revealed four ancestral populations that closely corresponded with the phylogenomic approach. We did not observe significant mixing between the different clusters, suggesting low connectivity and gene flow. Significantly, our analysis indicated that clusters 1 and 2 are widely divergent lineages (ANI, ca. 95%) and are on the verge of qualifying as different species (Fig. S1). For our subsequent analyses, we focused on clusters C1 and C2, which between them include close to 90% of the strains, including the bulk of the diversity of clinical isolates. Indeed, the low number of representatives in cluster 3 (C3) and cluster 4 (C4), combined with their high clonality, did not provide enough genetic diversity to obtain meaningful results. Nonetheless, future studies should analyze additional strains from C3 and C4, with higher divergence, to obtain a more comprehensive view of the evolution and emergence of additional *V. vulnificus* clusters.

Despite the genomic divergence among clusters, we could not identify a distinct pattern linking strain phylogeny, source of isolation, and virulent capabilities (Fig. 1). Nevertheless, C1 appears to contain a significantly higher proportion of strains isolated from humans, while C2 is dominated by strains derived from multiple marine hosts, including a large proportion of eels, and appears to be closely associated with aquaculture environments (Fig. 1B). An additional distinguishing feature is the relatively high clonality of C2 strains, both clinical and environmental, compared to a much higher divergence among strains belonging to C1 (Fig. 1A). Combined, these observations point toward different evolutionary pathways taken by these clusters that may be partially driven by anthropogenic influences.

Evolution and genomic diversity of *V. vulnificus* clusters. In order to further elucidate the evolutionary drivers behind the divergence of the two main clusters, we looked for signatures of genetic drift acting on the genome of each cluster. Assembly of 12 clinical and environmental strains from each cluster showed that the synteny in both chromosomes was remarkably well preserved within and also between clusters,

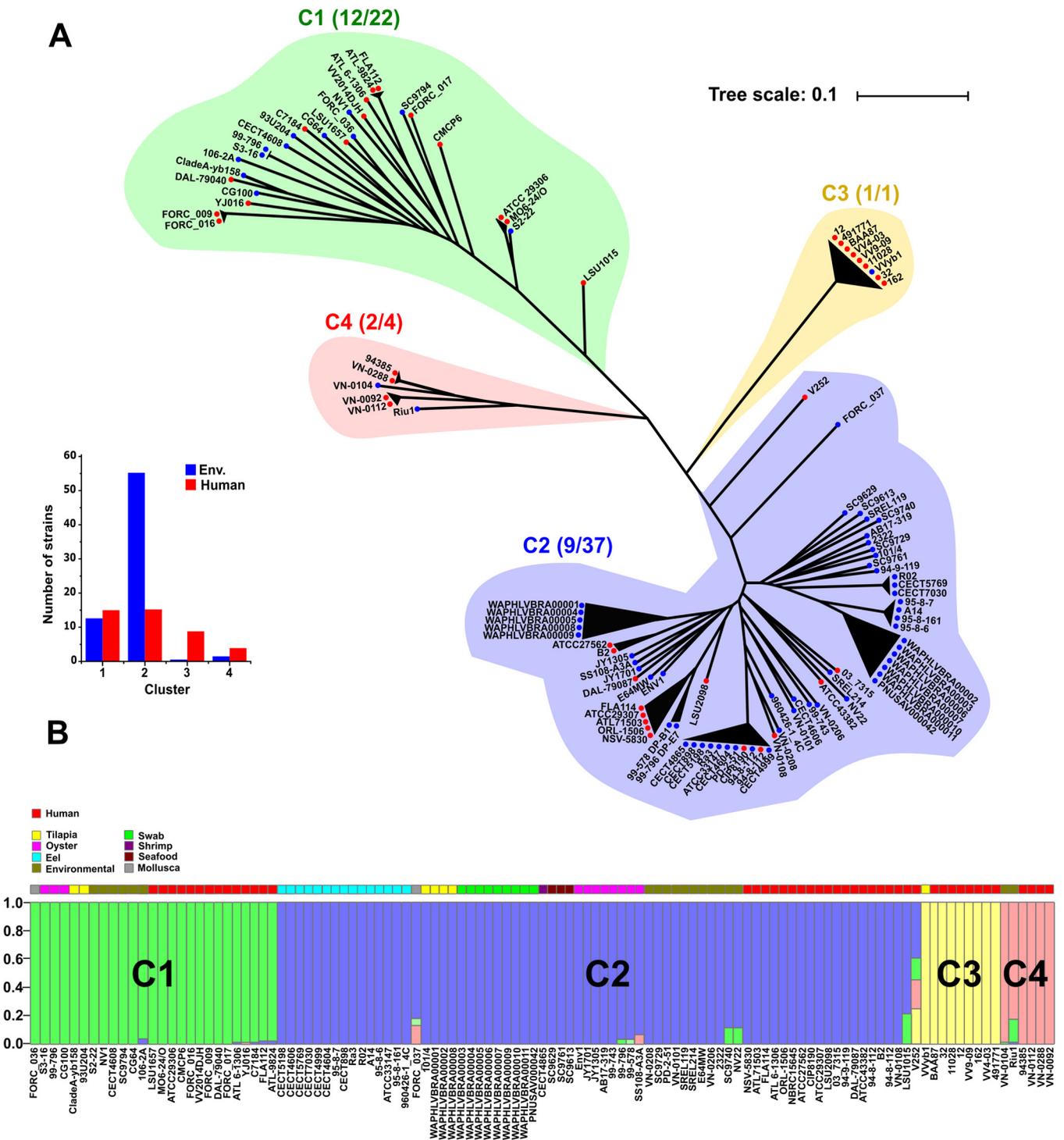


FIG 1 Phylogenomic and population structure of *V. vulnificus*. (A) Maximum likelihood tree reconstructed from single nucleotide polymorphisms of the core genome. Blue and red circles mark strains isolated from environmental and human samples, respectively. Members of the same cluster (C1 to C4) (ANI > 97%) are indicated with the same color. The smaller inset shows the proportions of nonclonal human and environmental isolates in each cluster. (B) STRUCTURE plot showing contribution to each strain from each of four ancestral populations (colored). Each vertical line represents one of the *V. vulnificus* strains. The color chart at the top of the plot indicates the isolated source of the corresponding strains.

including positions of the main features of the flexible genome (e.g., CPS and the superintegron) (Fig. S2). The overall means of the estimates of the averages of non-synonymous (dN) to synonymous (dS) substitution rates for the analyzed genomes were 0.42 ± 0.02 and 0.46 ± 0.01 for C1 and C2, respectively, indicating weak stabilizing

selection on both clusters. Similar dN/dS values were obtained for bacterial pathogens such as *Chlamydia trachomatis* (0.40) (31), *Salmonella enterica* serovar Typhi (0.45) (32) and *Burkholderia mallei* (0.47) (33).

We found substantial differences in the degree of divergence of the genomes within each cluster. Similarly to other *Vibrio* species, *V. vulnificus* possess two chromosomes with different sizes. The size of the large chromosome is ca. 3.2 Mb, whereas the size of the small chromosome is ca. 1.8 Mb. Despite having a lower relative abundance of strains, both chromosomes of C1 were found to be genetically more diverse than those of C2 and to be accumulating greater amounts of SNPs (Fig. S3). Furthermore, pairwise nucleotide diversity was higher in both clusters in chromosome 2 (Chr2) (Fig. S3), showing that genes are evolving faster in that chromosome than in chromosome 1 (Chr1). To further assess the relative effects of recombination and mutation between the two chromosomes of strains belonging to the two main clusters, we estimated the ratio of recombination events to point mutation events (R/θ) (34). The mean R/θ values for all strains from both clusters were 0.38 ± 0.02 for Chr1 and 0.41 ± 0.01 for Chr2. However, for strains within C1, ratios of recombination-associated replacements were higher in Chr I (0.69 ± 0.002) even though the R/θ values were similar for the second chromosome (0.32 ± 0.001). Similar R/θ values (0.85 ± 0.008 for Chr I and 0.30 ± 0.001 for Chr II) were estimated within C2 representatives.

These data indicate low gene flow between the clusters and limited recombination between them, possibly leading toward speciation (Fig. 1). The most plausible scenario that explains our findings is that physical isolation has decreased the probability of encounter and recombination between the two clusters leading to allopatric speciation and the generation of distinct ecotypes. This model of bacterial ecotype evolution and separation restricting recombination has already been observed previously in pathogens such as *Yersinia enterocolitica* (35) and *Campylobacter jejuni* (36) and in populations of the hyperthermophilic archaeon *Sulfolobus* (37).

Virulence factors and capsular polysaccharide diversity. Next, we compared data corresponding to the presence and distribution of known virulence factors in both clusters. In order to investigate this, we contrasted all protein-coding genes from the *V. vulnificus* strains against the Virulence Factors Database (38). We considered putative virulence factors for all genes sharing over 90% amino acid sequence identity with any entry in that database (Fig. S4). All *V. vulnificus* strains possess a wide array of putative virulence factors related to attachment and adhesion, iron acquisition, quorum sensing, secretion, and cytotoxicity systems. Surprisingly, with the exception of the CPS cluster, the virulence factors were similarly distributed in all strains regardless of cluster, source of isolation, or clinical/environmental designation (Fig. S4). The prevalence of antibiotic resistance genes, analyzed using the MegaRES database (39), pointed to an intrinsic resistance to tetracycline, with *tet*(34) and *tet*(35) genes present in the core genome of all the strains.

Since the CPS, which is part of the flexible genome of *V. vulnificus* (Fig. S2), was found to be the most diverse virulence factor, we investigated whether its genetic diversity might yield insights into the divergence of C1 and C2. First, we extracted and generated a heat map based on data representing the similarity of the CPS genomic islands using only one representative of each clonal frame within the *V. vulnificus* species (Fig. 2A). Pairwise comparison showed that the variability of this region is very high, even within members of the same cluster, since we did not find two identical versions (Fig. 2A). Furthermore, genomic comparison revealed that the CPS cluster has two separate hypervariable regions (Fig. 2B). SNP analysis of the common part of the CPS showed a large number of SNPs, mostly synonymous, detected at one side of these variable regions (Fig. 2B). This phenomenon has been suggested to be the result of events of recombination between divergent genomes and might elicit the complete replacement of the gene cluster (40).

Interestingly, we identified several syntenic groups in distantly related genomes. Highly similar CPS clusters (ca. 70% to 80% tBLASTx identity) were found in compari-

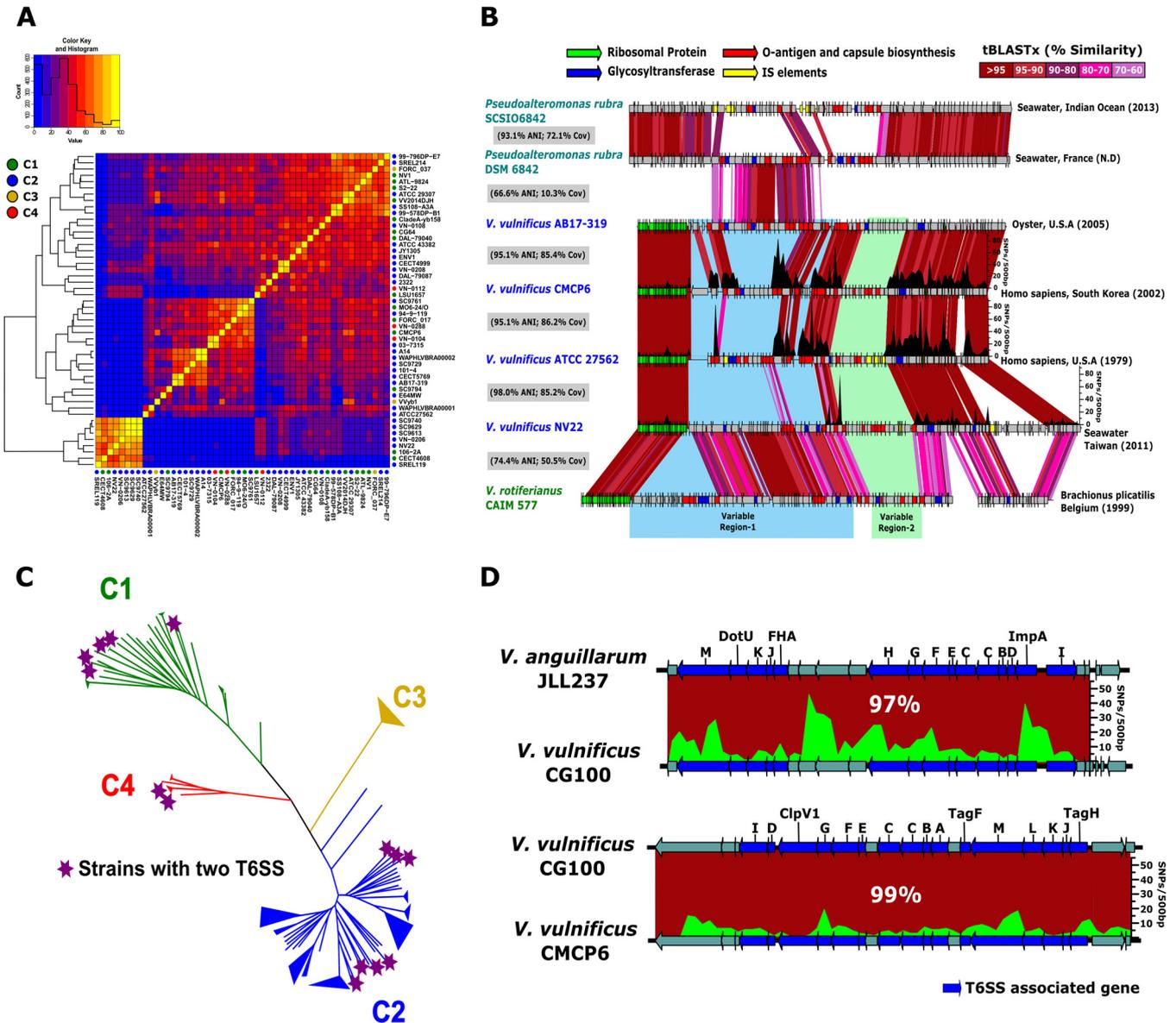


FIG 2 Capsular polysaccharide and type VI secretions system genomic region comparison. (A) Hierarchical clustering of pairwise average nucleotide identity comparison of the variable region of the capsular polysaccharide genomic island using one representative of each clonal frame within the *V. vulnificus* species. The correspondence of each strain with its cluster is shown by a color-coded circle. (B) Schematic representation of the same cassettes that were found to still be syntenic but at lower similarity in members of the genus *Pseudoalteromonas*. Color-coded arrows show locations of important genomic features. The plot above the genomes indicates the number of SNPs in a 500-bp window. Variable regions 1 and 2 are highlighted in blue and green, respectively. (C) The location of the 14 strains with two type VI secretion systems (T6SS) is highlighted with a purple star in the tree reconstructed from single nucleotide polymorphisms of the core genome (Fig. 1). (D) Comparison of the two T6SS clusters found in strain CG100. Genes associated with this cluster are highlighted in blue. The plots above the genome indicate the number of SNPs in a 500-bp window.

sons between *V. vulnificus* NV22 and *Vibrio rotiferianus* CAIM 577 (74.4% ANI, 50.5% coverage [Cov]) (Fig. 2B), with preserved synteny and location in the genome. Similarly, the CPS cluster of *V. vulnificus* AB17-319 shares a syntenic cluster with *Pseudoalteromonas rubra* DSM6842 (66.6% ANI, 10.3% Cov), a strain isolated from seawater in Nice, France (Fig. 2B). This region contains genes that encode glycosyltransferases, amino-transferases, and polysaccharide biosynthesis proteins, among other products.

With the exception of the CPS cluster, our analyses showed virulence factors to be both widely distributed and highly conserved among all analyzed *V. vulnificus* strains. Here we provide evidence of two hypervariable CPS regions that are frequently exchanged by homologous recombination within and between species, as there is

evidence for their import from distant taxonomic units. It is well documented that CPS is an essential virulence factor for *V. vulnificus* and other bacterial pathogens (19, 41–43). These highly variable regions are involved in the production of different sugar skeletons that form or decorate the extracellular structure and allow the bacterium to avoid predation by protozoa and other grazers in the natural environment (44, 45). The high variability in the CPS region of closely related microbes indicates that, even though it is essential for colonization and survival in the human host, it cannot be used as a factor to differentiate between clinical and environmental strains or between C1 and C2 strains.

Presence of two type VI secretion systems in *V. vulnificus*. The type VI secretion system (T6SS) functions as an antibacterial mechanism facilitating elimination of competing bacteria during host colonization (46–48). We hypothesized that the distribution of this key virulence factor within *Vibrio* species may thus shed light on the evolution of virulence in *V. vulnificus* (49). Our bioinformatic analysis revealed T6SS-associated genes in all *V. vulnificus* strains in our census. Moreover, these genes were invariably located in a conserved region on Chr2. Interestingly, 14 of the isolates, belonging to different clusters, harbored a second T6SS homolog (T6SS-2) on Chr2 (Fig. 2C). A phylogenetic analysis of the concatenation genes encoding the TssB and TssC proteins (Fig. S5), previously suggested as markers for evolutionary relationships between distantly related T6SS systems (50), revealed the core T6SS (T6SS-1) to be highly congruent with the phylogeny of the whole genome, suggesting vertical transmission (Fig. S5). In contrast, T6SS-2 clustered into four different groups, with similarity to other species within the genus *Vibrio* but no apparent correlation with overall phylogeny (Fig. S5). Fig. 2D shows a schematic of the two T6SS homologs from strain CG100, which encodes both systems. While T6SS-1 from this strain shared 99% similarity with the T6SS cluster from reference strain *V. vulnificus* CMCP6, the T6SS-2 (absent in CMCP6) showed 97% similarity to a T6SS cluster in *V. anguillarum* JLL237, isolated from fish tissue. To our knowledge, this is the first report of a human pathogen encoding two T6SS. It remains to be determined whether T6SS-2 contributes to the virulence of *V. vulnificus*. Redundancy in T6SS gene clusters has been previously described in *Vibrio fluvialis* (strain 85003) and *Vibrio proteolyticus* (51, 52). It is possible that this novel T6SS might have specificity for some bacterial or protozoal species, thus increasing the fitness of *V. vulnificus* in its natural environment and increasing the ability of the bacterium to outcompete the intestinal microbiota or to overcome the host's immune system.

***V. vulnificus* panmobilome.** Mobile genomic regions enable the rapid recombination of genetic elements, which may facilitate the expansion of physioecological range of a microbe, including the dissemination of antimicrobial resistance and virulence factors within a population (53–55). We analyzed the panmobilome of different *V. vulnificus* strains to identify unique mobile genetic elements (MGEs) associated with the different clusters.

(i) Plasmids. Several virulence plasmids were described in *V. vulnificus*, primarily in strains associated with C2 (23), that were suggested to provide resistance to the innate immunity of eels (56, 57). Here we describe a novel megaplasmid (404 kb) in C1 strain *V. vulnificus* CECT4608, originally isolated from the water of an eel tank (Spain, 1990), that is highly similar to and syntenic with (ANI, 98.7%; Cov, 92.9%) a megaplasmid found in *V. coralliilyticus* strain RE98 (p380; 380 kb) isolated from a shellfish hatchery (United States, 2000) (Fig. S6A). We also identified a 290-kb conjugative plasmid in strain CECT898 (Japan, 1979) which shows significant gene and organizational similarities to plasmids found in three *Vibrio* species (*Vibrio harveyi*, *V. cholerae*, and *V. parahaemolyticus*) isolated over a period of 35 years from different hosts and locations in Southeastern Asia (Fig. 3A; see also Fig. S6B). An important feature of this plasmid was the presence of a region with high GC content (ca. 66% compared with ca. 43% in the rest of the plasmid) flanked by a class 1 integron-integrase gene (Fig. 3B). In close proximity to this region, we identified a complete *mer* operon conferring mercury

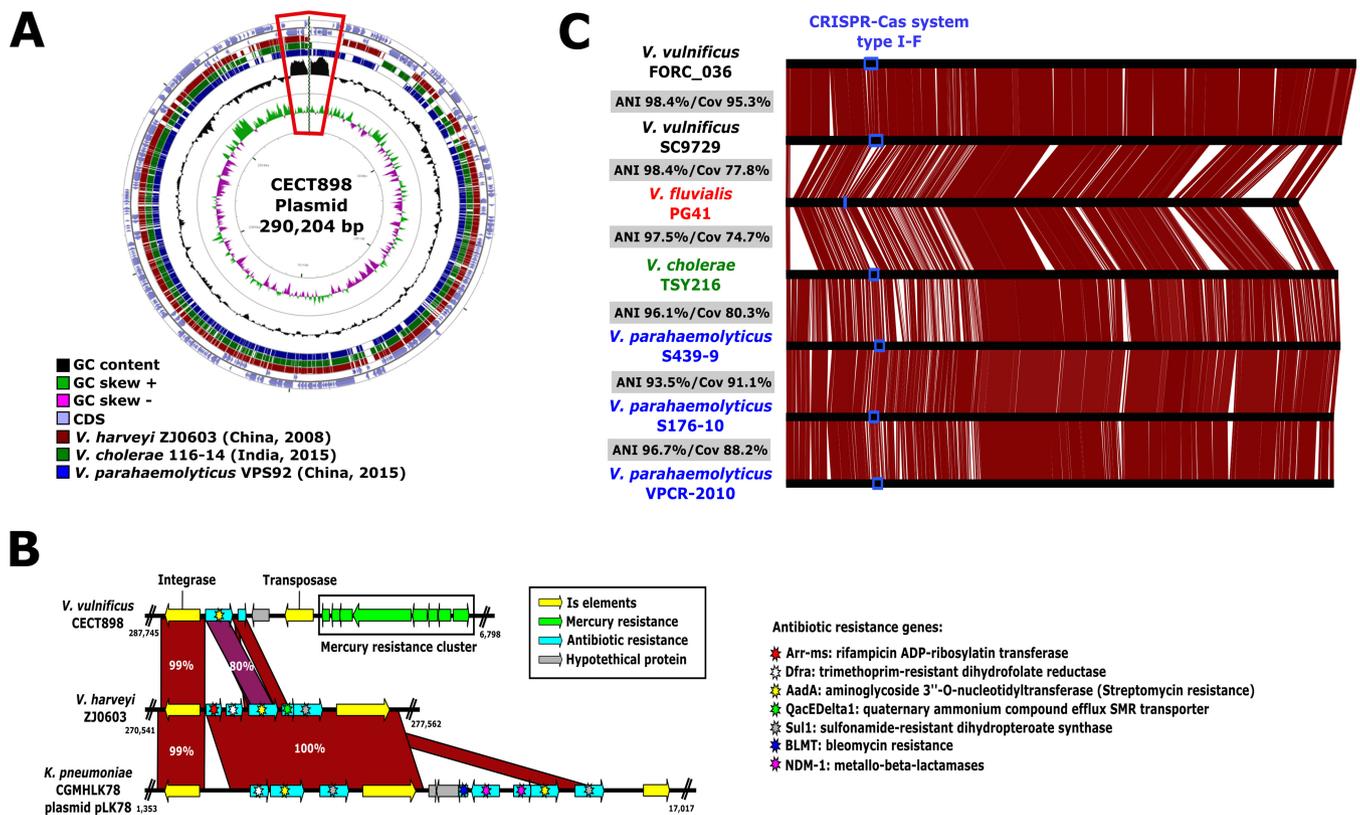


FIG 3 Panmobilome of *V. vulnificus*. (A) Circular representation of the CECT898 plasmid (290 kb). The rings are defined from outside to inside as follows: circles 1 and 2, coding DNA sequence (CDS) in the positive strand and negative strand, respectively; circles 3, 4, and 5, BLAST against *Vibrio harveyi* ZJ0603, *Vibrio cholerae* 116-14, and *Vibrio parahaemolyticus* VPS92; circle 6, GC content; circle 7, GC skew. The region bounded by red shows the enlarged region in section B. (B) Regions homologous to CECT898 plasmid with high GC content in *Vibrio harveyi* ZJ0603 and *Klebsiella pneumoniae* CGMHLK78 plasmid pLK78. (C) Schematic representation comparing the first two chromids described in *V. vulnificus* with other, similar examples found in several *Vibrio* species. CRISPR-Cas system, clustered regularly interspaced short palindromic repeat-Cas system.

resistance and showing high similarity to genes in distantly related microbes such as *Acinetobacter baumannii*, as well as to two antibiotic resistance genes, *aadA5* and *qacEdelta1*, conferring resistance to streptomycin and quaternary ammonium compounds, respectively (Fig. 3B). The presence of the integron-integrase gene, along with the 3'-conserved region containing *qacEdelta1* and *sul1* genes, clearly classifies this segment as a class 1 resistance integron, widely distributed among clinical strains involved in the capture and spread of antibiotic resistance genes (58).

(ii) *V. vulnificus* chromid. Two megaplasmids (ca. 900 kb), containing all the hallmarks of a "chromid" (59), were found in *V. vulnificus* strains FORC_36 and SC9729, belonging to C1 and C2, respectively. Despite substantial divergence between the hosting strains (ANI, 95.7%), the chromids share markedly high similarity (ANI, 98.49%; Cov, 95.3%) (Fig. 3C). Furthermore, the GC content of these chromids is 37.2%, substantially lower than the 46.5% average GC content determined for *V. vulnificus* strains. Interestingly, we found five nearly identical chromids hosted by three additional *Vibrio* species (*V. parahaemolyticus*, *V. fluvialis*, and *V. cholerae*) (Fig. 3C). To our knowledge, this is the first evidence of the widespread distribution of this kind of conjugative element and also of its presence across the species barrier of any known microbe.

(iii) Prophages. We have found 77 prophage-related elements in the *V. vulnificus* genomes associated with all clusters, located on both chromosomes and ranging in size from 3.7 to 59 kb. The most prevalent element, found in 29 *V. vulnificus* genomes, was a small (ca. 12-kb) prophage encoding two toxins, similarly to *V. cholerae* KSF-1 Φ phage (60, 61) (Fig. S7A). The phylogenomic analysis of both proteins clustered them separately into two well distinguishable branches (Fig. S7B and C). The largest protein was

annotated as zonula occludens toxin (Zot), required for phage morphogenesis, which was shown in *V. cholerae* to increase the permeability of the intestinal epithelium (62, 63). These two proteins are also found in the lysogenic *V. cholerae* CTX Φ phage, although the similarities among them were lower than 30%. Furthermore, the RtsA and RtsB proteins, which are also encoded in CTX Φ and KSF-1 Φ , were conserved in all the *V. vulnificus* sequences. In CTX Φ , RstA is required for DNA replication, whereas the RtsB gene facilitates the integration of the prophage into the *V. cholerae* genome (60).

Overall, our findings indicate that despite divergence between C1 and C2, exchanges of MGEs appear to happen frequently and indiscriminately. Nonetheless, this phenomenon also occurs between different species, as shown by the detection of identical elements in distant relatives (Fig. 3; see also Fig. S6). Our results further highlight the importance of plasmids for the dispersion of harmful genetic determinants among groups of strains even beyond the family barriers. Furthermore, they support the idea of the role of aquatic ecosystems as antibiotic resistance reservoirs. This poses a serious problem with respect to treatment of emergent *V. vulnificus* strains as they have the clear potential of becoming a multidrug-resistant pathogen.

***V. vulnificus* pangenome.** Given the marked divergence between the two main clusters (ANI, ca. 95%), we decided to investigate the evolutionary differences in their overall gene content through pangenome analysis. Despite the nearly 3-fold-larger number of C2 genomes, the C2 pangenome was only 12.7% larger than that of C1, possibly due to the high clonality within C2. Nevertheless, analysis of the core genome of each of the clusters indicated that C1 has a core genome that is nearly 40% larger than that of C2 (2,263 versus 1,641 genes). It seems that the core genome has already reached a plateau in both clusters (Fig. S8). The pangenome curve in both clusters has not saturated, indicating an open pangenome and high genetic diversity in *V. vulnificus* (Fig. S8). This interpretation is further supported by the high prevalence of “cloud genes” (e.g., genes found in up to 15% of the strains), which corresponds to approximately 70% of both pangenomes.

(i) Origins of pangenome divergence. Next, we aimed to elucidate the origins of the *V. vulnificus* divergence by identifying a potential common ancestor node among the four clusters. Using *V. cholerae* as an outgroup, we determined that *V. vulnificus* clusters started branching out at a node close to C3 and C4 (Fig. 4A). We unified the pangenome of both C3 and C4 to generate what we termed the *V. vulnificus* common ancestor (VVCA). We reasoned that subtraction of VVCA from the pangenome of the highly divergent C1 and C2 clusters would provide us with a reference point and give us an unbiased and specific in-depth evolutionary history of each group at the functional level. We compared the functional classifications of the gene coding sequences from the pangenomes of C1 and C2 after subtraction of VVCA. After this process, while the core of C1 was reduced by 40% to 650 genes, the core of C2 was reduced by up to 54%, leaving the number of genes possessed in common at 1,238. Comparison against the SEED Subsystems database (64) revealed differences between the clusters in the relative abundances of genes analyzed using a cutoff criterion of fold change greater than 1.5 in comparison to their prevalence in the VVCA (Fig. 4B).

In C1, we identified an increased proportion of genes related to the functional classification “carbohydrates” (monosaccharide and aminosugars), as well as to “membrane transport,” mainly, a greater proportion of type II secretion systems, which represent the vehicle for the secretion of the degradative enzymes (e.g., proteases, nucleases, phospholipases, and chitinases) that support bacterial persistence in different environments. Interestingly, this system plays a major role in the colonization and virulence of *V. cholerae* through degradation of intestinal mucin and cholera toxin secretion (65, 66). We also found more genes involved in chemotaxis, within the “motility and chemotaxis” classification, the process by which motile cells modulate flagellar rotation in response to the surrounding environment (67). Finally, genes involved in “dormancy and sporulation” were significantly enriched in C1 (Fig. 4B). These genes appear to regulate entry into and resuscitation from a persister-like state

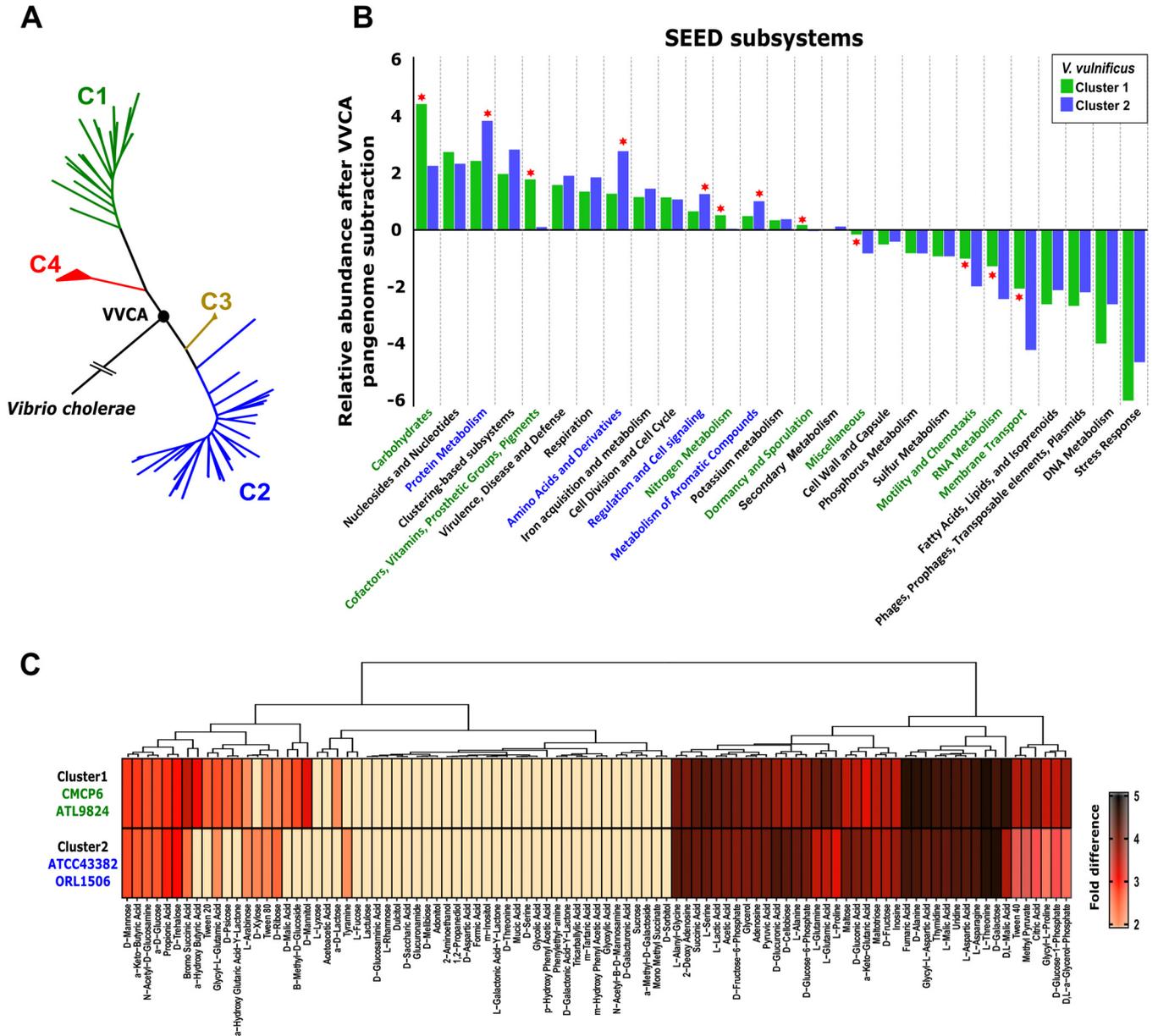


FIG 4 Pangenome analysis and carbon utilization of *V. vulnificus* strains. (A) Maximum likelihood tree reconstructed from single nucleotide polymorphisms of the core genome using *V. cholerae* as an outgroup. The black circle shows the root of the tree. Members of the same cluster are indicated with the same color. VVCA, *Vibrio vulnificus* common ancestor. (B) Functional characterization of the pangenome using the SEED Subsystems database and the difference between the two clusters and VVCA pangenome. The red stars indicate the results that differed using a cutoff criterion of fold change greater than 1.5 between the clusters. (C) Biolog phenotypic microarrays measuring bacterial ability to metabolize a variety of carbon sources by the use of PM1 phenotypic microarray plates. The heat map shows the average levels of carbon utilization of C1 representatives (CMCP6 and ATL9824) and C2 representatives (ATCC 43382 and ORL1506) in comparison to the negative control. Analyses were carried out in duplicate.

called viable but nonculturable, which is essential for both pathogenesis and survival in the environment in *V. cholerae* and other pathogens (68, 69). “Cofactors, vitamins, prosthetic groups, and pigments” and “RNA metabolism” are two additional categories that are overrepresented in C1. Overall, differential functional characterization of C1 suggested an opportunistic (*r*-strategist or bloomer) lifestyle, typical of microbes that grow rapidly, taking advantage of the sporadic inputs of organic matter that appear in the environment.

On the other hand, genes involved in protein biosynthesis (“protein metabolism”) and synthesis of amino acids (“amino acids and derivatives”) were more abundant in C2, suggesting possible adaptation to long-term colonization of nutrient-rich environ-

ments. The skin mucus of fish and eel is rich in proteins and carbohydrates and supports diverse commensal microbial populations (70). In fact, only *V. vulnificus* isolates of C2 have been recovered from diseased eels cultured in brackish and freshwater farms (23). We found that systems conferring resistance against phage predation such as restriction-modification and toxin-antitoxin (“regulation and cell signaling”) were enriched in the C2 pangenome. It has been demonstrated that bacteriophages are present in higher concentrations in the mucus layers than in the surrounding environment as a defense mechanism that ultimately protects the underlying epithelium from bacterial infections (56–58). Preferential colonization of mucosal surfaces is one possible explanation for the lower level of divergence and smaller core genome of C2, as strains from this cluster might not have to encounter conditions that are as variable as those encountered by the strains that typically inhabit the water column. Finally, C2 contained an enrichment of genes involved in aromatic carbon catabolism (“metabolism of aromatic compounds”).

(ii) Carbohydrate utilization. The most abundant functional classification in both clusters involved genes associated with carbohydrate metabolism, with a significant overrepresentation in strains from C1 (Fig. 4B). The ability to utilize a diverse set of carbon sources has been shown to be crucial for the pathogenicity and emergence of other pathogenic vibrios (71). We further analyzed the genomic potential for carbohydrate metabolism using the Carbohydrate-Active enZYmes (CAZy) database (72). We found a significantly higher percentage of glycoside hydrolases (GHs) and polysaccharide lyases in the C1 pangenome, suggesting a greater carbohydrate degradation capacity in this cluster (Fig. S9A). With the exception of two α -mannosidase families (GH38 and GH92), all GH families were found in equal or greater numbers in the C1 pangenome, including 9 families that were found only in that cluster (Fig. S9A). Mannose is among the main monosaccharides that constitute mucus glycoproteins, together with N-acetyl- α -galactosamine, N-acetyl- β -galactosamine, N-acetylglucosamine, fucose, and neuraminic acid, the latter two of which are found in the terminal residues of mucin glycoproteins (73).

The highest differences in abundances between the clusters corresponded to GH13 (α -amylase), GH23 (lysozyme with activity for several polysaccharides, including chitin), and GH109 (α -N-acetylgalactosaminidase) that might degrade the peptidoglycan of cell walls and mucus (74) (Fig. S9A). In order to experimentally determine the potential differences between C1 and C2 in carbohydrate utilization, we analyzed the growth of a couple of strains from each cluster (CMCP6 and ATL9824 [C1] and ATCC 43382 and ORL1506 [C2]) in a diverse range of carbon sources. We tested their ability to metabolize a variety of carbohydrates and other carbon sources using Biolog phenotypic microarray PM1 (Fig. 4C). Our analysis indicated that C1 is capable of utilizing a much larger range of carbon sources, in agreement with the genomic data of the pangenome (Fig. 4C; see also Fig. S9B). While members of the C1 grew better than the C2 representatives in 70 carbon sources, C2 outgrew C1 in only 14 (Fig. 4C). Specifically, C1 had at least twice the final optical density (OD) of C2 in D-mannitol, B-methyl-D-glucoside, α -D-lactose, and D-malic acid (Fig. 4C). Interestingly, the C2 ATCC 43382 strain was capable of utilizing L-fucose, a major component of the terminal glycans found in mucins, as a sole carbon source, whereas C1 representatives did not grow in the presence of L-fucose as a sole carbon source (Fig. S9B). Genome analysis results confirmed the phenotypic assay results, since ATCC 43382 contains genes encoding L-fucose permease in its genome, while CMCP6 does not. It has been demonstrated that uptake and utilization of L-fucose by *Campylobacter jejuni*, a prevalent gastrointestinal pathogen in humans, provided a distinct competitive advantage in its pathogenesis (75).

Adaptations to eel mucosa colonization. Our analyses (phenotype and genomic evolution) and the distribution of the isolates indicate that the two main clusters of *V. vulnificus* have different lifestyles. Thus, we speculated that these phylogenetically distant clusters occupy different niches, possibly differing in their natural hosts or

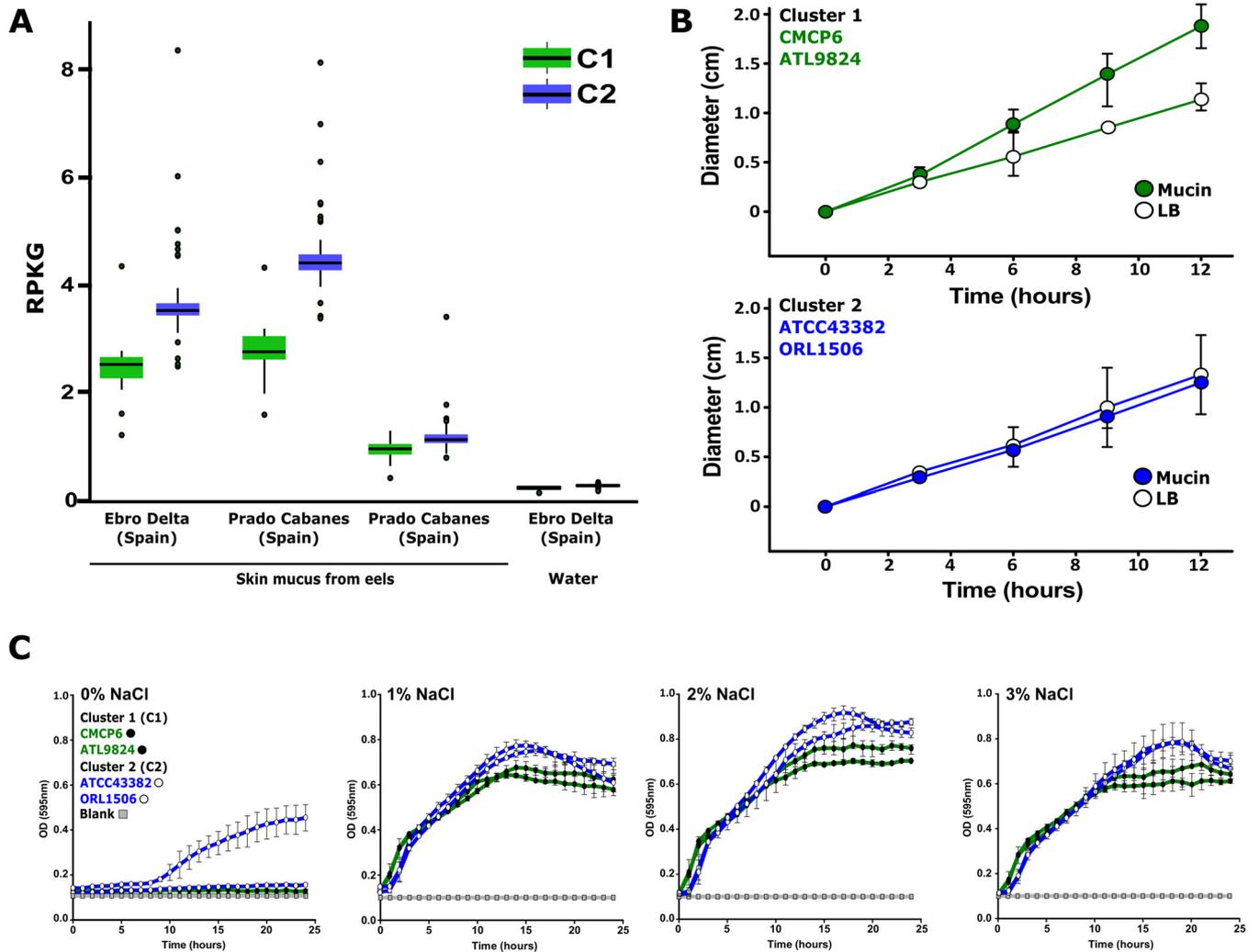


FIG 5 Adaptations of *V. vulnificus* strains to colonization of eel mucosa. (A) Relative abundances of members of the C1 and C2 based on metagenomic recruitment in eel skin mucus and seawater metagenome samples. Data are expressed as RPKG (reads recruited per kilobase of genome per gigabase of metagenome). (B) Motility assays of representatives of C1 (CMCP6 and ATL9824; green) and C2 (ATCC 43382 and ORL1506; blue) minimal media supplemented with mucin and LB at RT. Graphics represent the average diameter of motility zone of three replicates and error bars the standard deviation at 0 h, 3 h, 6 h, 9 h, and 12 h. (C) Growth of representatives of C1 (CMCP6 and ATL9824; green lines and white dots) and C2 (ATCC 43382 and ORL1506; blue lines and black dots) at different salinities (0% to 3% NaCl). Error bars represent standard deviations of results from three replicates.

habitat. This physical isolation would result in distinct evolutionary pressures that would decrease the probability of encounter and recombination, ultimately leading to genetic isolation and the generation of distinct ecotypes. To test our hypothesis, we analyzed the abundance of both clusters through recruitment in the metagenome of the eel skin mucus and seawater (76). Even though all the sequenced strains isolated from healthy and diseased eels from worldwide locations belonged to C2 (Fig. 1), metagenomic recruitment showed that both clusters were present in the eel mucus. However, C2 was consistently found at greater frequencies than C1 (Fig. 5A) further supporting a scenario where it preferentially lives as a commensal of marine organisms. None of the clusters were detected in seawater metagenomes (Fig. 5A). *V. vulnificus* genomes do not generate a large number of reads from marine metagenomes, and most analysis of the possible host microbiota derived from 16S rRNA has shown that detection at the cluster level is impossible. It would be revealing to investigate the presence of members of the different clusters in other *V. vulnificus* hosts in future studies.

The consistent enrichment of C2 in the eel mucus metagenome prompted us to investigate the physiological adaptations that might lead to these differences. First, we

investigated the motility response of C1 (strains CMCP6 and ATL9824) and C2 (strains ATCC 43382 and ORL1506) in the presence of mucin. We compared their motility characteristics in soft-agar plates containing M9 minimal media supplemented with 0.1% mucin or Luria-Bertani (LB) (Fig. 5B). We measured their motility zones at different time points for a period of 12 h. Strains from C1 exhibited greater motility in mucin than in LB media on average, while no significant differences between the two conditions were found for C2 (Fig. 5B). Therefore, C1 appears to react to some mucin components, increasing its motility in their presence. This has been previously identified in pathogenic strains of *V. cholerae* and other vibrios (77, 78). Our results are consistent with the differences in the pangenome for “motility and chemotaxis” in the SEED classification between the clusters and support the idea of a bloomer lifestyle for C1, chemotactic with respect to novel carbon sources, and of a commensal lifestyle for C2, adapted to live in a rich environment. The physicochemical parameters of the environment within the eel mucus differ from those of the surrounding environment. There are pH fluctuations, and the mean osmolarity of the eel mucus (~1% NaCl) is lower than that of seawater (76). Thus, we examined the response of the two clusters to different pH-related conditions by testing their growth on Biolog phenotypic microarray PM10 (Fig. S9B). We found no major differences among the four strains from the two clusters that we analyzed under these conditions (Fig. S9C). We also tested growth under conditions of increased salinity (0% to 3% NaCl) (Fig. 5C). Interestingly, there were differences in growth on LB over the entire range of salinities (including the average mucus salinity level [1%]), where the C2 strains showed better growth than the representatives from C1 (Fig. 5C), which could be another important factor that explains the predominance of C2 in the metagenomes of the eel mucus.

DISCUSSION

In order to emerge as a human pathogen, a bacterium must acquire numerous novel properties such as resistance to antimicrobials, avoidance of host immune defenses, or the ability to effectively colonize specific host tissues (66, 79). The acquisition and evolution of some of these pathogenic determinants are the results of the interaction of the bacterium with its natural habitat. These interactions prompt the selection of certain traits that increase its fitness in that ecological setting and also play a role in the context of the human host (26, 27). In this study, we investigated the population structure and genomic evolution of the marine pathogen *V. vulnificus* in order to understand the drivers that led to its emergence and cluster divergence.

The combined results of the different analyses in this study suggest that the population of *V. vulnificus* is made up of four distinct clusters. Although the ANI values within the different clusters were >97%, the divergence between the two largest clusters, C1 and C2, indicates that they are widely divergent lineages that are on the borderline of qualifying as different species. We speculate that the acquisition of different ecological determinants allowed the development of diverse lifestyles within the same environment, which has led to higher divergence. Interestingly, despite the genomic and ecological divergence of C1 and C2, the exchange of MGEs appears to happen frequently and indiscriminately, even between different species, as shown by the detection of identical elements in distant relatives (80).

It appears that C2 members have a competitive advantage for colonization and growth in different hosts following a commensal lifestyle, which might be a reason explaining why the greatest number of isolates has been obtained from this cluster (Fig. 6B). This specialization model could explain the smaller core genome and lower divergence of the members of C2 and their prevalence in nutrient-rich environments such as mucous surfaces. Our results indicate that C1 is a bloomer that grows when the conditions are favorable due to the high potential to degrade carbohydrates, a greater proportion of secretion systems, or a higher abundance of genes related to “dormancy and sporulation,” which support a “feast to famine” lifestyle allowing bacterial cells to endure long periods of unfavorable environmental conditions (68, 69) (Fig. 6B). Their ability to use a greater pool of nutrients and to tolerate a larger range of stressors in

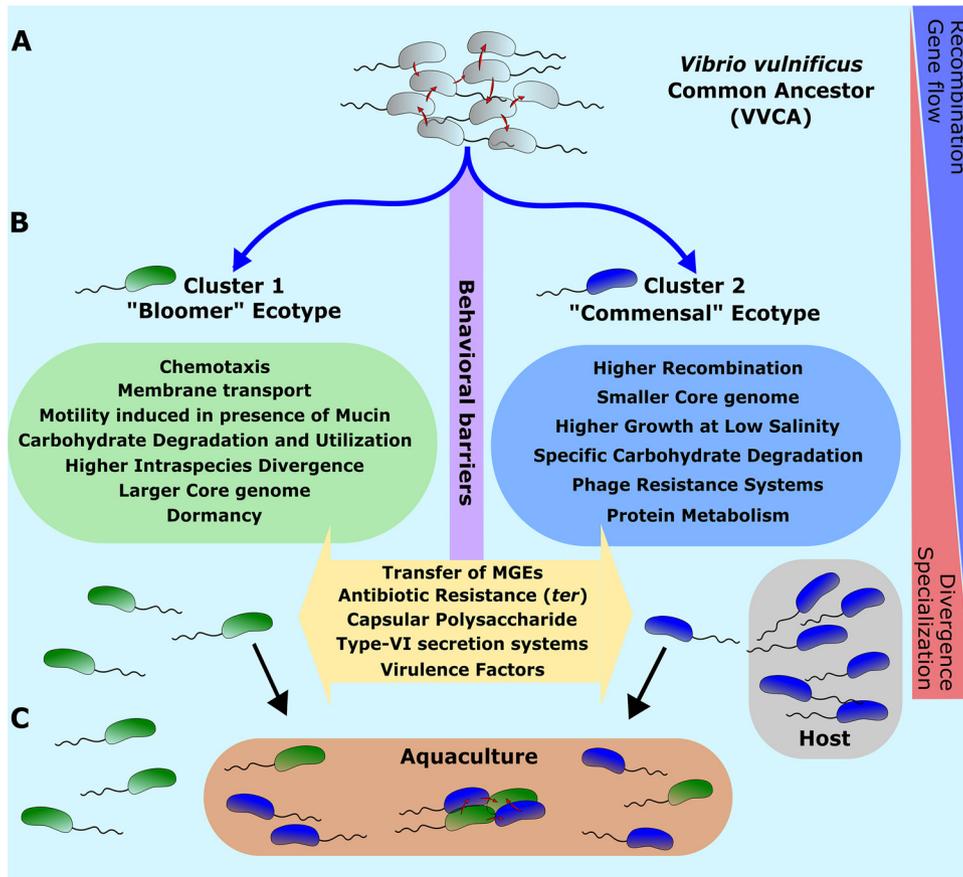


FIG 6 Evolutionary model of cluster divergence in *V. vulnificus*. (A) VVCA. Clonal lineages start diverging from the *V. vulnificus* common ancestor (VVCA). (B) Divergence. The acquisition of different ecological determinants allowed the development of diverse lifestyles within the same environment, which has led to a higher divergence. This divergence led to a recombination and gene flow decrease, although frequent exchange of mobile genetic elements is found within the species and with other species. (C) Convergence. With the advent of aquaculture, we have created an artificial environment that has led to colocalization of strains from the two major clusters.

the environment likely provides an advantage to *V. vulnificus* C1 in coping with the rapid and drastic ecological transition under the unfavorable conditions of the oligotrophic aquatic environments.

Our scenario proposes that strains from the two clusters occupy different niches that lead over time to a greater divergence of the two ecotypes. We contend that these continuing divergences would likely eventually lead to speciation of the two clusters (Fig. 6B). Interestingly, it appears that with the advent of aquaculture we have created an artificial environment where the two clusters can be isolated in sympatry; while C2 is commensal of the eel in aquaculture, the confinement of the fish, together with the organic matter that is added and their depositions, might produce the ideal environment for a continuous bloom of members of C1 (Fig. 6C). Furthermore, this newly created artificial environment increases the possibility of contact between the clusters, thus maximizing the probability of transfer of genetic material and of recombination. This could entail a risk of emergence of novel clusters with potentially devastating consequences for both aquaculture and human health.

Overall, our results shed light on some of the underlying genomic properties associated with the emergence of pathogenic *V. vulnificus*. We consider that our findings could provide information relevant to the pursuit of strategies to prevent and foresee the potential emergence of strains with importance to both human health and aquaculture. Finally, our evolutionary model and genomic approaches are broadly applicable to other pathogenic vibrios and facultative bacterial pathogens.

MATERIALS AND METHODS

Bacterial isolates and population structure. We downloaded all the genomes present in the NCBI belonging to the *V. vulnificus* species. Genomic features, cluster affiliations, and origins of the 113 *V. vulnificus* strains used are listed in Table S1 in the supplemental material. Reciprocal BLASTN and TBLASTX searches were carried out between the genomes, leading to the identification of regions of similarity, insertions, and rearrangements. The values representing ANI between strains were calculated using JSpecies software package v1.2.1 and default parameters (81).

SNPs, population structure, and recombination analyses. The population structure of *V. vulnificus* was reconstructed using STRUCTURE (82). The number of hypothetical ancestral populations (K) was estimated to be equal to $K = 4$. The optimum K value was evaluated by the ΔK method (83) using independent runs for a number of populations K ranging from 2 to 10. The Harvest Suite, a software package which includes tools such as Parsnp and Gingr, was used to perform the core alignment and to obtain the SNPs between strains (84). Indels and SNPs between small regions of the genome such as genomic islands were identified using the nucmer program in the MUMmer3+ package (85). Clonal-FrameML (34) was also used with default parameters to take into account recombination events and to calculate the R/θ ratios (relative rates of recombination and mutation). GenBank files from all the strains downloaded from the NCBI were converted to GFF files. These files were used to estimate the pangenome using the Roary pipeline with a 70% identity cutoff value (86).

Phylogenomic reconstructions. The core genome SNP analyses for all the strains were performed using the KSNP v3.0 program (87) with the optimum kmer size of 19, which was determined by Kchooser. Maximum likelihood trees for the two chromosomes were generated individually using RAxML (version 7.2.6) (88), and the core alignment was obtained with Parsnp software. Then, the file was edited using iTool v3 software (89).

Evolutionary rate. To calculate the nonsynonymous (dN) and synonymous (dS) substitutions for an ortholog in a pair of *V. vulnificus* strains, we used the ortholog package (90). Briefly, this package identifies orthologous gene pairs by choosing the best reciprocal hit using BLASTp and performs codon alignments of the orthologous gene pairs using PAL2NAL (91). Finally, GESTIMATOR (92) computes the dN/dS values of the codon alignments. A low ratio ($dN/dS < 1$) indicates purifying selection, whereas a high ratio ($dN/dS > 1$) is a clear signal of diversifying selection.

Functional classification. Putative functionality, the presence of virulence factors, and the presence of antibiotic resistance factors encoded in the genomes were inferred by comparing all the proteins against the SEED subsystem database (64), virulence factor database (VFDB) (38), and MEGARes database (39), respectively. Proteins were compared using the different databases and DIAMOND (93) (blastp option, top hit, $\geq 50\%$ identity, $\geq 50\%$ alignment length, E value of $< 10^{-5}$). Using dbSCAN (94), we analyzed the presence of glycoside hydrolases, comparing all the proteins against the Carbohydrate-Active enZymes (CAZy) database (72).

Strains and culture conditions. The experimental analyses utilized *Vibrio vulnificus* isolates CMCP6 and ATL9824 as representatives of C1 (laboratory collection) and ATCC 43382 (American Type Culture Collection, Rockville, MD) and ORL1506 (Paul Gulig) as representatives of C2. All strains were routinely grown in Luria-Bertani (LB) broth or agar plates containing 2% (wt/vol) NaCl for 16 h aerobically at 37°C, unless otherwise specified.

Biolog phenotypic microarrays. Differences in carbon utilization and pH tolerance were assessed using the Biolog microbial identification system (Biolog, Hayward, CA). Phenotypic MicroArray 1 (PM1) and PM10 analyses were carried out in duplicate following the manufacturer's instructions. Briefly, colonies from agar plates were suspended in $1 \times$ IF-0a (PM1) or IF-10 (PM10) inoculation media supplemented with NaCl for a final concentration of 1% (wt/vol), as recommended by the manufacturer, and a 1:5 dilution (PM1) or 1:200 dilution (PM10) of this suspension was prepared to obtain an absorbance of 0.07 at 600 nm. Aliquots (100 μ l) of the final cell suspension were added to each well. The plates were incubated under aerobic conditions at 37°C for 48 h with shaking. The optical density (OD) was measured at 595 nm every hour for 48 h using a Tecan Sunrise microplate reader, and the results were evaluated using Magellan plate reader software. Growth curves were plotted using GraphPad Prism V7, and area under the curve (AuC) values were calculated. Data were normalized by taking the ratio of the AuC of the respective carbon sources or pH conditions to that of the negative control. Normalized data were used to plot a heat map to compare the strains.

Metagenomic read recruitments. C1 and C2 members were used to recruit reads from eel mucus and seawater metagenomics data sets (76) using BLASTN and a cutoff value of 99% nucleotide identity over a minimum alignment length of 50 nucleotides. Hits obtained were used to compute the RPKG (reads recruited per kilobase of genome per gigabase of metagenome) values that provide normalized numbers that are comparable across various metagenomes.

Growth curves at different salinities. For the growth curve analysis, the overnight cultures of *V. vulnificus* strains were centrifuged to obtain a pellet, washed with LB, and resuspended in LB media containing no salt. Dilutions (1:100) of the cell suspensions were prepared in LB media containing final concentrations of 0%, 1%, 2%, and 3% (wt/vol) NaCl. Aliquots (200 μ l) of each suspension were added to a 96-well microtiter plate. The experiment was performed in triplicate, with three independent biological replicates. Growth curves were plotted using GraphPad Prism V7.

Motility assays. Assessment of motility was performed using soft-agar motility plates containing 0.3% (wt/vol) agar and either LB or M9 minimal media supplemented with 0.1% (wt/vol) mucin from porcine stomach (Sigma). Single colonies were stabbed in the center of the ~ 10 -cm-diameter soft-agar plates using a sterile inoculating loop. Plates were incubated at room temperature (RT). The diameter of the motility zone was measured at the time points indicated in the x axis. Experiments were conducted

in triplicate. Graphs were plotted using GraphPad Prism software, V7. The motility characteristics of the two strains on the different media were compared using Student's *t* test (*, $P < 0.05$; **, $P < 0.005$).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mBio.02852-18>.

FIG S1, PDF file, 0.1 MB.

FIG S2, PDF file, 0.1 MB.

FIG S3, PDF file, 0.02 MB.

FIG S4, PDF file, 3.9 MB.

FIG S5, PDF file, 0.05 MB.

FIG S6, PDF file, 0.7 MB.

FIG S7, PDF file, 0.1 MB.

FIG S8, PDF file, 0.04 MB.

FIG S9, PDF file, 0.2 MB.

TABLE S1, XLSX file, 0.03 MB.

ACKNOWLEDGMENTS

We thank the reviewers for their thoughtful comments.

This work was supported by startup funds from the Burnett School of Biomedical Sciences to S.A.-M., funds from the Binational Science Foundation (BSF 2016319) to S.A.-M. and O.H.S., funds from "VIREVO" CGL2016-76273-P [AEI/FEDER, EU] (cofunded with FEDER funds) to F.R.-V., and a Postdoctoral fellowship (IJCI-2017-34002) to M.L.-P. from the Spanish Ministerio de Economía, Industria y Competitividad. J.M.H.-M. and A.Z.-S. were supported by Ph.D. fellowships from the Spanish Ministerio de Economía y Competitividad (BES-2014-067828 and BES-2017-079993, respectively).

REFERENCES

1. Reen FJ, Almagro-Moreno S, Ussery D, Boyd EF. 2006. The genomic code: inferring Vibrionaceae niche specialization. *Nat Rev Microbiol* 4:697–704. <https://doi.org/10.1038/nrmicro1476>.
2. Frank C, Littman M, Alpers K, Hallauer J. 2006. *Vibrio vulnificus* wound infections after contact with the Baltic Sea, Germany. *Euro Surveill* 11:E060817.1.
3. Baker-Austin C, Trinanés J, González-Escalona N, Martínez-Urtaza J. 2017. Non-cholera vibrios: the microbial barometer of climate change. *Trends Microbiol* 25:76–84. <https://doi.org/10.1016/j.tim.2016.09.008>.
4. Baker-Austin C, Trinanés JA, Taylor NGH, Hartnell R, Siitonen A, Martínez-Urtaza J. 2013. Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nat Clim Chang* 3:73–77. <https://doi.org/10.1038/nclimate1628>.
5. Vezzulli L, Grande C, Reid PC, Hélaouët P, Edwards M, Höfle MG, Brettar I, Colwell RR, Pruzzo C. 2016. Climate influence on *Vibrio* and associated human diseases during the past half-century in the coastal North Atlantic. *Proc Natl Acad Sci U S A* 113:E5062–E5071. <https://doi.org/10.1073/pnas.1609157113>.
6. Azam F, Malfatti F. 2007. Microbial structuring of marine ecosystems. *Nat Rev Microbiol* 5:782–791. <https://doi.org/10.1038/nrmicro1747>.
7. Kordas RL, Harley CDG, O'Connor MI. 2011. Community ecology in a warming world: the influence of temperature on interspecific interactions in marine systems. *J Exp Mar Bio Ecol* 400:218–226. <https://doi.org/10.1016/j.jembe.2011.02.029>.
8. Martínez-Urtaza J, Bowers JC, Trinanés J, DePaola A. 2010. Climate anomalies and the increasing risk of *Vibrio parahaemolyticus* and *Vibrio vulnificus* illnesses. *Food Res Int* 43:1780–1790. <https://doi.org/10.1016/j.foodres.2010.04.001>.
9. Austin B. 2010. Vibrios as causal agents of zoonoses. *Vet Microbiol* 140:310–317. <https://doi.org/10.1016/j.vetmic.2009.03.015>.
10. Cui Y, Yang X, Didelot X, Guo C, Li D, Yan Y, Zhang Y, Yuan Y, Yang H, Wang J, Wang J, Song Y, Zhou D, Falush D, Yang R. 2015. Epidemic clones, oceanic gene pools, and eco-LD in the free living marine pathogen *Vibrio parahaemolyticus*. *Mol Biol Evol* 32:1396–1410. <https://doi.org/10.1093/molbev/msv009>.
11. Dechet AM, Yu PA, Koram N, Painter J. 2008. Nonfoodborne *Vibrio* infections: an important cause of morbidity and mortality in the United States, 1997–2006. *Clin Infect Dis* 46:970–976. <https://doi.org/10.1086/529148>.
12. Jones MK, Oliver JD. 2009. *Vibrio vulnificus*: disease and pathogenesis. *Infect Immun* 77:1723–1733. <https://doi.org/10.1128/IAI.01046-08>.
13. Newton A, Kendall M, Vugia DJ, Henao OL, Mahon BE. 2012. Increasing rates of vibriosis in the United States, 1996–2010: Review of surveillance data from 2 systems. *Clin Infect Dis* 54(Suppl 5):S391–S395. <https://doi.org/10.1093/cid/cis243>.
14. Centers for Disease Control and Prevention (CDC). 2011. Vital signs: incidence and trends of infection with pathogens transmitted commonly through food—foodborne diseases active surveillance network, 10 U.S. sites, 1996–2010. *MMWR Morb Mortal Wkly Rep* 60:749–755.
15. Oliver JD. 2013. *Vibrio vulnificus*: death on the half shell. A personal journey with the pathogen and its ecology. *Microb Ecol* 65:793–799. <https://doi.org/10.1007/s00248-012-0140-9>.
16. Phillips KE, Satchell KJF. 2017. *Vibrio vulnificus*: from oyster colonist to human pathogen. *PLoS Pathog* 13:e1006053. <https://doi.org/10.1371/journal.ppat.1006053>.
17. Warnock EW, MacMath TL. 1993. Primary *Vibrio vulnificus* septicemia. *J Emerg Med* 11:153–156.
18. Oliver JD. 2005. Wound infections caused by *Vibrio vulnificus* and other marine bacteria. *Epidemiol Infect* 133:383–391.
19. Gulig PA, Bourdage KL, Starks AM. 2005. Molecular pathogenesis of *Vibrio vulnificus*. *J Microbiol* 43(Spec No):118–131.
20. Bisharat N, Cohen DI, Harding RM, Falush D, Crook DW, Peto T, Maiden MC. 2005. Hybrid *Vibrio vulnificus*. *Emerg Infect Dis* 11:30–35. <https://doi.org/10.3201/eid1101.040440>.
21. Tison DL, Nishibuchi M, Greenwood JD, Seidler RJ. 1982. *Vibrio vulnificus* biogroup 2: new biogroup pathogenic for eels. *Appl Environ Microbiol* 44:640–646.
22. Sanjuán E, González-Candelas F, Amaro C. 2011. Polyphyletic origin of *Vibrio vulnificus* biotype 2 as revealed by sequence-based analysis. *Appl Environ Microbiol* 77:688–695. <https://doi.org/10.1128/AEM.01263-10>.
23. Roig FJ, González-Candelas F, Sanjuán E, Fouz B, Feil EJ, Llorens C, Baker-Austin C, Oliver JD, Danin-Poleg Y, Gibas CJ, Kashi Y, Gulig PA, Morrison SS, Amaro C. 5 January 2018. Phylogeny of *Vibrio vulnificus*

- from the analysis of the core-genome: implications for intra-species taxonomy. *Front Microbiol* <https://doi.org/10.3389/fmicb.2017.02613>.
24. Chun J, Grim CJ, Hasan NA, Lee JH, Choi SY, Haley BJ, Taviani E, Jeon Y-S, Kim DW, Lee J-H, Brettin TS, Bruce DC, Challacombe JF, Dettler JC, Han CS, Munk AC, Chertkov O, Meincke L, Saunders E, Walters RA, Huq A, Nair GB, Colwell RR. 2009. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proc Natl Acad Sci U S A* 106:15442–15447. <https://doi.org/10.1073/pnas.0907787106>.
 25. Boucher Y. 2016. Sustained local diversity of *Vibrio cholerae* O1 biotypes in a previously cholera-free country. *mBio* 3:e00570-16. <https://doi.org/10.1128/mBio.00570-16>.
 26. Shapiro BJ, Levade I, Kovacicova G, Taylor RK, Almagro-Moreno S. 2016. Origins of pandemic *Vibrio cholerae* from environmental gene pools. *Nat Microbiol* 2:16240. <https://doi.org/10.1038/nmicrobiol.2016.240>.
 27. Sakib SN, Reddi G, Almagro-Moreno S. 26 March 2018. Environmental role of pathogenic traits in *Vibrio cholerae*. *J Bacteriol* <https://doi.org/10.1128/JB.00795-17>.
 28. Cohen ALV, Oliver JD, DePaola A, Feil EJ, Fidelma Boyd E. 2007. Emergence of a virulent clade of *Vibrio vulnificus* and correlation with the presence of a 33-kilobase genomic island. *Appl Environ Microbiol* 73:5553–5565. <https://doi.org/10.1128/AEM.00635-07>.
 29. Chatzidakis-Livanis M, Hubbard MA, Gordon K, Harwood VJ, Wright AC. 2006. Genetic distinctions among clinical and environmental strains of *Vibrio vulnificus*. *Appl Environ Microbiol* 72:6136–6141. <https://doi.org/10.1128/AEM.00341-06>.
 30. Horseman MA, Surani S. 2011. A comprehensive review of *Vibrio vulnificus*: an important cause of severe sepsis and skin and soft-tissue infection. *Int J Infect Dis* 15:e157. <https://doi.org/10.1016/j.ijid.2010.11.003>.
 31. Joseph SJ, Didelot X, Rothschild J, De Vries HJC, Morré SA, Read TD, Dean D. 2012. Population genomics of *Chlamydia trachomatis*: insights on drift, selection, recombination, and population structure. *Mol Biol Evol* 29:3933–3946. <https://doi.org/10.1093/molbev/mss198>.
 32. Holt KE, Parkhill J, Mazzoni CJ, Roumagnac P, Weill FX, Goodhead I, Rance R, Baker S, Maskell DJ, Wain J, Dolecek C, Achtman M, Dougan G. 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. *Nat Genet* 40:987–993. <https://doi.org/10.1038/ng.195>.
 33. Hershberg R, Petrov DA. 2010. Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genet* 6:e1001115. <https://doi.org/10.1371/journal.pgen.1001115>.
 34. Didelot X, Wilson DJ. 2015. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput Biol* 11:e1004041. <https://doi.org/10.1371/journal.pcbi.1004041>.
 35. Thomson NR, Reuter S, de Been M, Hall M, Harris S, McNally A, Cheng L, Corander J. 29 September 2015. Directional gene flow and ecological separation in *Yersinia enterocolitica*. *Microb Genomics* <https://doi.org/10.1099/mgen.0.000030>.
 36. Wilson DJ, Gabriel E, Leatherbarrow AJH, Cheesbrough J, Gee S, Bolton E, Fox A, Hart CA, Diggle PJ, Fearnhead P. 2009. Rapid evolution and the importance of recombination to the gastroenteric pathogen *Campylobacter jejuni*. *Mol Biol Evol* 26:385–397. <https://doi.org/10.1093/molbev/msn264>.
 37. Whitaker RJ, Grogan DW, Taylor JW. 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* 301:976–978. <https://doi.org/10.1126/science.1086909>.
 38. Chen L, Zheng D, Liu B, Yang J, Jin Q. 2016. VFDB 2016: hierarchical and refined dataset for big data analysis - 10 years on. *Nucleic Acids Res* 44:D694–D697. <https://doi.org/10.1093/nar/gkv1239>.
 39. Lakin SM, Dean C, Noyes NR, Dettenwanger A, Ross AS, Doster E, Rovira P, Abdo Z, Jones KL, Ruiz J, Belk KE, Morley PS, Boucher C. 2017. MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res* 45:D574–D580. <https://doi.org/10.1093/nar/gkw1009>.
 40. López-Pérez M, Martín-Cuadrado A-B, Rodríguez-Valera F. 22 May 2014. Homologous recombination is involved in the diversity of replacement flexible genomic Islands in aquatic prokaryotes. *Front Genet* <https://doi.org/10.3389/fgene.2014.00147>.
 41. Park NY, Lee JH, Kim MW, Jeong HG, Lee BC, Kim TS, Choi SH. 2006. Identification of the *Vibrio vulnificus* wbpP gene and evaluation of its role in virulence. *Infect Immun* 74:721–728. <https://doi.org/10.1128/IAI.74.1.721-728.2006>.
 42. Lubin JB, Lewis WG, Gilbert NM, Weimer CM, Almagro-Moreno S, Fidelma Boyd E, Lewis AL. 2015. Host-like carbohydrates promote bloodstream survival of *Vibrio vulnificus* in vivo. *Infect Immun* 83:3126–3136. <https://doi.org/10.1128/IAI.00345-15>.
 43. Moxon ER, Kroll JS. 1988. Type b capsular polysaccharide as a virulence factor of *Haemophilus influenzae*. *Vaccine* 6:113–115.
 44. Rodríguez-Valera F, Martín-Cuadrado A-B, Rodríguez-Brito B, Pasić L, Thingstad TF, Rohwer F, Mira A. 2009. Explaining microbial population genomics through phage predation. *Nat Rev Microbiol* 7:828–836. <https://doi.org/10.1038/nrmicro2235>.
 45. Avrani S, Schwartz DA, Lindell D. 2012. Virus-host swinging party in the oceans. *Mob Genet Elements* 2:88–95. <https://doi.org/10.4161/mge.20031>.
 46. Cianfanelli FR, Monlezun L, Coulthurst SJ. 2016. Aim, load, fire: the type VI secretion system, a bacterial nanoweapon. *Trends Microbiol* 24:51–62. <https://doi.org/10.1016/j.tim.2015.10.005>.
 47. MacIntyre DL, Miyata ST, Kitaoka M, Pukatzki S. 2010. The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proc Natl Acad Sci U S A* 107:19520–19524. <https://doi.org/10.1073/pnas.1012931107>.
 48. Fu Y, Waldor MK, Mekalanos JJ. 2013. Tn-seq analysis of *vibrio cholerae* intestinal colonization reveals a role for T6SS-mediated antibacterial activity in the host. *Cell Host Microbe* 14:652–663. <https://doi.org/10.1016/j.chom.2013.11.001>.
 49. Ho BT, Dong TG, Mekalanos JJ. 2014. A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe* 15:9–21. <https://doi.org/10.1016/j.chom.2013.11.008>.
 50. Schwarz S, West TE, Boyer F, Chiang W-C, Carl MA, Hood RD, Rohmer L, Tolker-Nielsen T, Skerrett SJ, Mougous JD. 2010. Burkholderia type VI secretion systems have distinct roles in eukaryotic and bacterial cell interactions. *PLoS Pathog* 6:e1001068. <https://doi.org/10.1371/journal.ppat.1001068>.
 51. Huang Y, Du P, Zhao M, Liu W, Du Y, Diao B, Li J, Kan B, Liang W. 30 March 2017. Functional characterization and conditional regulation of the type VI secretion system in *Vibrio fluvialis*. *Front Microbiol* <https://doi.org/10.3389/fmicb.2017.00528>.
 52. Ray A, Schwartz N, de Souza Santos M, Zhang J, Orth K, Salomon D. 2017. Type VI secretion system MIX-effectors carry both antibacterial and anti-eukaryotic activities. *EMBO Rep* 18:1978–1990. <https://doi.org/10.15252/embr.201744226>.
 53. Schroeder M, Brooks BD, Brooks AE. 2017. The complex relationship between virulence and antibiotic resistance. *Genes (Basel)* 8:39. <https://doi.org/10.3390/genes8010039>.
 54. von Wintersdorff CJH, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, Savelkoul PHM, Wolfs PFG. 2016. Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Front Microbiol* 7:173.
 55. Soucy SM, Huang J, Gogarten JP. 2015. Horizontal gene transfer: building the web of life. *Nat Rev Genet* 16:472–482. <https://doi.org/10.1038/nrg3962>.
 56. Lee CT, Pajuelo D, Llorens A, Chen YH, Leiro JM, Padrós F, Hor LI, Amaro C. 2013. MARTX of *Vibrio vulnificus* biotype 2 is a virulence and survival factor. *Environ Microbiol* 15:419–432. <https://doi.org/10.1111/j.1462-2920.2012.02854.x>.
 57. Lee C, Te maro C, Wu KM, Valiente E, Chang YF, Tsai SF, Chang CH, Hor LI. 2008. A common virulence plasmid in biotype 2 *Vibrio vulnificus* and its dissemination aided by a conjugal plasmid. *J Bacteriol* 190:1638–1648. <https://doi.org/10.1128/JB.01484-07>.
 58. Gillings MR. 2014. Integrons: past, present, and future. *Microbiol Mol Biol Rev* 78:257–277. <https://doi.org/10.1128/MMBR.00056-13>.
 59. Harrison PW, Lower RPJ, Kim NKD, Young JPW. 2010. Introducing the bacterial “chromid”: not a chromosome, not a plasmid. *Trends Microbiol* 18:141–148. <https://doi.org/10.1016/j.tim.2009.12.010>.
 60. Faruque SM, Mekalanos JJ. 2012. Phage-bacterial interactions in the evolution of toxigenic *Vibrio cholerae*. *Virulence* 3:556–565. <https://doi.org/10.4161/viru.22351>.
 61. Faruque SM, Naser IB, Fujihara K, Diraphat P, Chowdhury N, Kamruzzaman M, Qadri F, Yamasaki S, Ghosh AN, Mekalanos JJ. 2005. Genomic sequence and receptor for the *Vibrio cholerae* phage KSF-1Φ: evolutionary divergence among filamentous vibriophages mediating lateral gene transfer. *J Bacteriol* 187:4095–4103. <https://doi.org/10.1128/JB.187.12.4095-4103.2005>.
 62. Uzzau S, Cappuccinelli P, Fasano A. 1999. Expression of *Vibrio cholerae* zonula occludens toxin and analysis of its subcellular localization. *Microb Pathog* 27:377–385. <https://doi.org/10.1006/mpat.1999.0312>.

63. Waldor MK, Mekalanos JJ. 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272:1910–1914. <https://doi.org/10.1126/science.272.5270.1910>.
64. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, de Crécy-Lagard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata RD, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Rückert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 33:5691–5702.
65. Cianciotto NP, White RC. 21 April 2017. Expanding role of type II secretion in bacterial pathogenesis and beyond. *Infect Immun* <https://doi.org/10.1128/IAI.00014-17>.
66. Almagro-Moreno S, Pruss K, Taylor RK. 2015. Intestinal colonization dynamics of *Vibrio cholerae*. *PLoS Pathog* 11:e1004787. <https://doi.org/10.1371/journal.ppat.1004787>.
67. Colin R, Sourjik V. 2017. Emergent properties of bacterial chemotaxis pathway. *Curr Opin Microbiol* 39:24–33. <https://doi.org/10.1016/j.mib.2017.07.004>.
68. Alam M, Sultana M, Nair GB, Siddique AK, Hasan NA, Sack RB, Sack DA, Ahmed KU, Sadique A, Watanabe H, Grim CJ, Huq A, Colwell RR. 2007. Viable but nonculturable *Vibrio cholerae* O1 in biofilms in the aquatic environment and their role in cholera transmission. *Proc Natl Acad Sci U S A* 104:17801–17806. <https://doi.org/10.1073/pnas.0705599104>.
69. Oliver JD. 2010. Recent findings on the viable but nonculturable state in pathogenic bacteria. *FEMS Microbiol Rev* 34:415–425. <https://doi.org/10.1111/j.1574-6976.2009.00200.x>.
70. Benhamed S, Guardiola FA, Mars M, Esteban MÁ. 2014. Pathogen bacteria adhesion to skin mucus of fishes. *Vet Microbiol* 171:1–12. <https://doi.org/10.1016/j.vetmic.2014.03.008>.
71. Almagro-Moreno S, Boyd EF. 2009. Sialic acid catabolism confers a competitive advantage to pathogenic *Vibrio cholerae* in the mouse intestine. *Infect Immun* 77:3807–3816. <https://doi.org/10.1128/IAI.00279-09>.
72. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 21 November 2013. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* <https://doi.org/10.1093/nar/gkt1178>.
73. Alexander JB, Ingram GA. 1992. Noncellular nonspecific defence mechanisms of fish. *Annu Rev Fish Dis* 2:249–279. [https://doi.org/10.1016/0959-8030\(92\)90066-7](https://doi.org/10.1016/0959-8030(92)90066-7).
74. Kamke J, Sczyrba A, Ivanova N, Schwientek P, Rinke C, Mavromatis K, Woyke T, Hentschel U. 2013. Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges. *ISME J* 7:2287–2300. <https://doi.org/10.1038/ismej.2013.111>.
75. Stahl M, Friis LM, Nothhaft H, Liu X, Li J, Szymanski CM, Stintzi A. 2011. L-Fucose utilization provides *Campylobacter jejuni* with a competitive advantage. *Proc Natl Acad Sci U S A* 108:7194–7199. <https://doi.org/10.1073/pnas.1014125108>.
76. Carda-Diéguez M, Ghai R, Rodríguez-Valera F, Amaro C. 2017. Wild eel microbiome reveals that skin mucus of fish could be a natural niche for aquatic mucosal pathogen evolution. *Microbiome* 5:162. <https://doi.org/10.1186/s40168-017-0376-1>.
77. Liu Z, Wang Y, Liu S, Sheng Y, Rueggeberg KG, Wang H, Li J, Gu FX, Zhong Z, Kan B, Zhu J. 2015. *Vibrio cholerae* represses polysaccharide synthesis to promote motility in mucosa. *Infect Immun* 83:1114–1121. <https://doi.org/10.1128/IAI.02841-14>.
78. Reddi G, Pruss K, Cottingham KL, Taylor RK, Almagro-Moreno S. 26 July 2018. Catabolism of mucus components influences motility of *Vibrio cholerae* in the presence of environmental reservoirs. *PLoS One* <https://doi.org/10.1371/journal.pone.0201383>.
79. Vouga M, Greub G. 2016. Emerging bacterial pathogens: the past and beyond. *Clin Microbiol Infect* 22:12–21. <https://doi.org/10.1016/j.cmi.2015.10.010>.
80. Shapiro BJ, Friedman J, Cordero OX, Preheim SP, Timberlake SC, Szabó G, Polz MF, Alm EJ. 2012. Population genomics of early events in the ecological differentiation of bacteria. *Science* 335:48–51. <https://doi.org/10.1126/science.1218198>.
81. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 106:19126–19131. <https://doi.org/10.1073/pnas.0906412106>.
82. Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
83. Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.
84. Treangen TJ, Ondov BD, Koren S, Phillippy AM. 19 November 2014. The harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* <https://doi.org/10.1186/s13059-014-0524-x>.
85. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol* 5:R12. <https://doi.org/10.1186/gb-2004-5-2-r12>.
86. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>.
87. Gardner SN, Slezak T, Hall BG. 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 31:2877–2878. <https://doi.org/10.1093/bioinformatics/btv271>.
88. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
89. Letunic I, Bork P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23:127–128. <https://doi.org/10.1093/bioinformatics/btl529>.
90. Drost HG, Gabel A, Grosse I, Quint M. 2015. Evidence for active maintenance of phylotranscriptomic hourglass patterns in animal and plant embryogenesis. *Mol Biol Evol* 32:1221–1231. <https://doi.org/10.1093/molbev/msv012>.
91. Suyama M, Torrents D, Bork P. 1 July 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* <https://doi.org/10.1093/nar/gkl315>.
92. Thornton K. 2003. libsequence: a C++ class library for evolutionary genetic analysis. *Bioinformatics* 19:2325–2327.
93. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.
94. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. 29 May 2012. DbCAN: a Web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* <https://doi.org/10.1093/nar/gks479>.