Genotypic Diversity and Population Structure of *Vibrio vulnificus* Strains Isolated in Taiwan and Korea as Determined by Multilocus Sequence Typing

Hye-Jin Kim[1], Jae-Chang Cho[1,2]*

1 Institute of Environmental Sciences and Department of Environmental Sciences, Hankuk University of Foreign Studies, Yong-In, Korea, **2** Solana Informatics, San Diego, CA, United States of America

* chojc@hufs.ac.kr

Abstract

The genetic diversity and population structure of *Vibrio vulnificus* isolates from Korea and Taiwan were investigated using PCR-based assays targeting putative virulence-related genes and multilocus sequence typing (MLST). BOX-PCR genomic fingerprinting identified 52 unique genotypes in 84 environmental and clinical *V. vulnificus* isolates. The majority (> 50%) of strains had pathogenic genotypes for all loci tested; moreover, many environmental strains had pathogenic genotypes. Although significant ($p < 0.05$) inter-relationships among the genotypes were observed, the association between genotype and strain source (environmental or clinical) was not significant, indicating that genotypic characteristics alone are not sufficient to predict the isolation source or the virulence of a given *V. vulnificus* strain and vice versa. MLST revealed 23–35 allelic types per locus analyzed, resulting in a total of 44 unique sequence types (STs). Two major monophyletic groups (lineages A and B) corresponding to the two known lineages of *V. vulnificus* were observed; lineage A had six STs that were exclusively environmental, whereas lineage B had STs from both environmental and clinical sources. Pathogenic and nonpathogenic genotypes predominated in MLST lineages B and A, respectively. In addition, *V. vulnificus* was shown to be in linkage disequilibrium ($p < 0.05$), although two different recombination tests (PHI and Sawyer's tests) detected significant evidence of recombination. Tajima's D test also indicated that *V. vulnificus* might be comprised of recently sub-divided lineages. These results suggested that the two lineages revealed by MLST correspond to two distinct ecotypes of *V. vulnificus*.

Introduction

Vibrio vulnificus is a halophilic, Gram-negative, curved, rod-shaped bacterium frequently found in marine environments (e.g., estuarine and coastal waters) as well as in molluscan shellfish (e.g., oysters and clams) as part of the normal microflora [1–3]. *V. vulnificus* is also an important human pathogen, infection by which mainly occurs through the consumption of contaminated seafood and occasionally through open wounds exposed to contaminated

seawater [1, 3, 4]. Although most data regarding the epidemiology of *V. vulnificus* infection are for the United States of America (USA) [2, 5–8], human infections with *V. vulnificus* have also been reported in many other countries [9–16]. The most fatal consequence of *V. vulnificus* infection is primary septicemia, which has a mortality rate of over 50% in immunocompromised individuals [2, 5, 17]. Due to the high mortality rate of *V. vulnificus* infection, *V. vulnificus* contamination is a serious food safety issue in regions where seafood (especially raw oysters) consumption is common. Until recently, studies of *V. vulnificus* focused mainly on the virulence factors that confer its remarkable pathogenicity. The *V. vulnificus* virulence factors identified to date include capsular polysaccharide (CPS) [18, 19], lipopolysaccharide (LPS) [20, 21], hemolysin [22, 23], the RtxA1 toxin [24, 25], and other molecules involved in iron acquisition and the formation of pili and flagella [26–29]. To evaluate and predict the virulence of different *V. vulnificus* strains, several genotyping methods have been developed [30–38]. Although genotyping studies are not a substitute for animal models for evaluating the virulence of *V. vulnificus* strains, such studies provide valuable information regarding the distribution of *V. vulnificus* genotypes.

Since most cases of *V. vulnificus* infection have been reported primarily in the Gulf Coast region of the USA [39–42] and in specific regions of Asia [10, 43–45], environmental factors, geographic factors, and seafood consumption patterns of inhabitants in these regions could all play an important role in *V. vulnificus* epidemiology. Interestingly, *V. vulnificus* infection is most prevalent in temperate coastal regions [3, 46, 47]. Moreover, the majority (>70%) of *V. vulnificus* infections in Korea occur on the southwest coast, whereas *V. vulnificus* infection is very rare on the east coast [48–50]. This is a striking observation, since the dietary preferences of the inhabitants of the southwest coast and the east coast are quite similar. Additionally, in Taiwan, most reported cases (>90%) of *V. vulnificus*-induced illness have occurred in the southern part of Taiwan [10]. The prevalence of virulent strains might vary by geographic region, although the proportion of susceptible individuals might also vary within a region according to socioeconomic status or other factors [6].

Extensive studies of *V. vulnificus* virulence factors have been performed using highly virulent strains (CMCP6, MO6-24/O, and YJ016) isolated in Korea, the United States, and Taiwan [51–53]. However, the natural populations of *V. vulnificus* have been poorly characterized, especially the *V. vulnificus* populations found in Asian countries. The genotypic distribution patterns of *V. vulnificus* have been investigated mainly in the Gulf Coast region [54–57], while other studies have examined the genetic diversity of *V. vulnificus* in the northeastern United States and in European countries [35, 36, 58, 59]. In this study, we investigated the genetic diversity of *V. vulnificus* strains isolated in Taiwan and Korea. These strains were subjected to PCR-based genotypic characterization of putative virulence factors and to multilocus sequence typing (MLST) using seven genetic loci. The association between the genotypic characteristics of *V. vulnificus* strains and their source of origin, as well as the population structure of *V. vulnificus* as inferred from the MLST data, are reported here.

Materials and Methods

Bacterial strains, biochemical tests, and general molecular techniques

A total of 84 strains of *V. vulnificus* (hereafter referred to as VV) originating from the United States, Taiwan, and Korea were included in this study (Table 1). Clinical strains (n = 21) had been previously isolated from hospitalized patients with septicemia [51–53, 60–64], whereas environmental strains (n = 63) were isolated from seawater, seafood, and tidal mudflat samples. Three clinical strains whose genomes had been sequenced (CMCP6, MO6-24/O, and YJ016) [51, 53, 64] were used as reference strains. All VV strains were routinely cultured in Luria-

Table 1. *Vibrio vulnificus* strains used in this study.

Strain	Isolation source[a]	Country of isolation	Reference[b]	Strain	Isolation source[a]	Country of isolation	Reference[b]
ATCC 29307	C	United States	70, ATCC	SC9613	E (crab)	Korea	58, 59
CG108	E	Taiwan	60	SC9624	E (shell)	Korea	58, 59
CG110	E	Taiwan	60	SC9629	E (clam)	Korea	58, 59
CG122	E	Taiwan	60	SC9631	E (clam)	Korea	58, 59
CG21	E	Taiwan	60	SC9641	E (octopus)	Korea	58, 59
CG26	E	Taiwan	60	SC9648	E (clam)	Korea	58, 59
CG27	E	Taiwan	60	SC9649	E (shell)	Korea	58, 59
CG33	E	Taiwan	60	SC97100	E (conch)	Korea	58, 59
CG46	E	Taiwan	60	SC97114	E (tidal mudflat)	Korea	58, 59
CG54	E	Taiwan	60	SC97116	E (crab)	Korea	58, 59
CG55	E	Taiwan	60	SC97118	E (tidal mudflat)	Korea	58, 59
CG57	E	Taiwan	60	SC97126	E (oyster)	Korea	58, 59
CG58	E	Taiwan	60	SC9716	E (seawater)	Korea	58, 59
CG62	E	Taiwan	60	SC9717	E (tidal mudflat)	Korea	58, 59
CG64	E	Taiwan	60	SC9720	E (tidal mudflat)	Korea	58, 59
CNUH94-6	C	Korea	58, 59, CNUH	SC9721	E (seawater)	Korea	58, 59
CMCP6	C	Korea	68	SC9728	E (seawater)	Korea	58, 59
CN7	C	Korea	58, 59, CNUH	SC9729	E (seawater)	Korea	58, 59
CN8	C	Korea	58, 59, CNUH	SC9730	E (tidal mudflat)	Korea	58, 59
CN9	C	Korea	58, 59, CNUH	SC9731	E (seawater)	Korea	58, 59
CNUH94-3	C	Korea	58, 59, CNUH	SC9733	E (fish)	Korea	58, 59
CNUH94-4	C	Korea	58, 59, CNUH	SC9737	E (tidal mudflat)	Korea	58, 59
CS91133	C	Korea	58, 59, CNUH	SC9738	E (oyster)	Korea	58, 59
MO6-24/O	C	United States	67	SC9740	E (seawater)	Korea	58, 59
NV1	E	Taiwan	60	SC9761	E (oyster)	Korea	58, 59
NV101	E	Taiwan	60	SC9763	E (seawater)	Korea	58, 59
NV15	E	Taiwan	60	SC9766	E (clam)	Korea	58, 59
NV18	E	Taiwan	60	SC9771	E (shell)	Korea	58, 59
NV22	E	Taiwan	60	SC9793	E (seawater)	Korea	58, 59
NV24	E	Taiwan	60	SC9794	E (tidal mudflat)	Korea	58, 59
NV28	E	Taiwan	60	SC9795	E (seawater)	Korea	58, 59
NV31	E	Taiwan	60	V-15	C	Korea	58, 59, CNUH
NV33	E	Taiwan	60	V-16	C	Korea	58, 59, CNUH
NV37	E	Taiwan	60	V-19	C	Korea	58, 59, CNUH
NV42	E	Taiwan	60	WK13	C	Korea	58, 59, WKUH
NV43	E	Taiwan	60	WK15	C	Korea	58, 59, WKUH
NV55	E	Taiwan	60	WK16	C	Korea	58, 59, WKUH
NV61	E	Taiwan	60	WK20	C	Korea	58, 59, WKUH
NV63	E	Taiwan	60	WK22	C	Korea	58, 59, WKUH
NV69	E	Taiwan	60	WK3	C	Korea	58, 59, WKUH
NV72	E	Taiwan	60	WK6	C	Korea	58, 59, WKUH
NV78	E	Taiwan	60	YJ016	C	Taiwan	69

[a] C and E denote clinical (human patient with septicemia) and environmental (seawater, seafood, and tidal mudflat) sources, respectively.

[b] Literature or additional information (e.g., strain provider). ATCC, American Type Culture Collection; CNUH, Chonnam National University Hospital, Korea; WKUH, Won Kwang University Hospital, Korea. i, Hollis et al. [61]; ii, Wong et al. [123]; iii, Kim et al. [62] and Lee et al. [63]; iv, Kim et al. [51]; v, Wright et al. [52]; vi, Shao and Hor [60].

doi:10.1371/journal.pone.0142657.t001

Bertani medium, either with or without 2.5% NaCl (LB or LBS), at 37°C. Frozen stocks were stored at -80°C in LBS with 70% (vol/vol) glycerol.

VV strains were biotyped using the following biochemical tests: i) indole production from tryptophan (indole reaction), ii) putrescine production from orinithine (ornithine decarboxylation, ODC reaction), and iii) o-nitrophenol production from o-nitrophenyl-β-D-galactopyranoside (ONPG reaction). Biotypes were assigned according to the methods of Tison et al. [65] and Bisharat et al. [66] (biotype 1, positive for indole and ODC reactions; biotype 2, negative for indole and ODC reactions; biotype 3, ONPG-negative).

Genomic DNA was extracted from bacterial cultures in exponential growth phase using Genomic-Tip kits (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol and stored in sterile water at -20°C until use. Unless specified otherwise, all other general experimental procedures were performed according to a standard laboratory manual [67].

BOX-PCR genome fingerprinting

Repetitive extragenic palindromic PCR (rep-PCR) genomic fingerprinting of the VV strains was carried out with a BOX-A1R primer (BOX-PCR) according to the protocol of Cho and Tiedje [68], with minor modifications. Similarity matrices of whole densitometric curves of the gel tracks were calculated using the pairwise Pearson's product-moment correlation coefficients (r values) [69] as described by Rademaker et al. [70]. Cluster analyses of the similarity matrices were performed using the unweighted pair group method with arithmetic averages (UPGMA).

Genotypic characteristics

Representative strains that were selected based on the UPGMA clustering of their BOX-PCR fingerprinting patterns were subjected to genotype analyses. Six genetic loci, *pilF* (pilus-type IV assembly gene), *vcg* (ORF no. VV0401, virulence-correlated gene of strain YJ016), *viuB* (vulnibactin gene), *vuuA* (ferric vulnibactin receptor), *vvhA* (VV hemolysin gene), and CPS (capsular polysaccharide) alleles were analyzed according to the methods of Roig et al. [30], Rosche et al. [71], Panicker et al. [72], Kim et al. [73], Kaysner and DePaola [74], and Han et al. [75], respectively. Associations between genotypic characteristics and strain origin (environmental or clinical), as well as associations among the genotypic characteristics themselves, were evaluated with the chi-square (χ^2) test and Fisher's exact test ($\alpha = 0.05$).

Multilocus sequence typing (MLST)

Representative strains of BOX-PCR types (Table 2) were subjected to MLST of seven genetic loci. These loci included six housekeeping genes that encoded glutamine synthetase (*glnA*), glucose-6-phosphate isomerase (*glp*), DNA gyrase subunit B (*gyrB*), malate-lactate dehydrogenase (*mdh*), dihydroorotase (*pyrC*), and recombinase A (*recA*), in addition to a virulence-associated gene encoding VV hemolysin (*vvhA*) (details described in S1 Table). *vvhA* was included to supplement the phylogenetic information collected from the housekeeping genes with information from a VV-specific gene. Partial sequences of the MLST loci were obtained using specific primer pairs and the amplification conditions described in the PubMLST database [76] for *glp*, *mdh*, and *pyrC*; by the conditions described by Gutacker et al. [77] for *glnA* and *recA*; and by the conditions described by Kotetishvili et al. [78] for *gyrB*. PCR amplicons were sequenced using an ABI3700 DNA analyzer (Applied Biosystems, Foster City, CA, USA). DNA sequences of each gene were aligned with CLUSTAL W [79]. No gaps (indels) were introduced into any of the alignments, and no manual editing was performed.

Table 2. Genotypic characteristics of representative *Vibrio vulnificus* strains as assessed by BOX-PCR genomic fingerprinting.

BOX-PCR profile no.[a]	Representative strain[b]	No. of strains[c]	Soruce of isolation[d]	16S rRNA type[e]	CPS type[e]	*vcg* type[e]	*pilF*[e]	*viuB*[e]	*vuuA*[e]
1	SC9629	2	E	A	NA[f]	E	-	-	-
2	NV22	2	E	A	NA	C	-	-	-
3	SC9729	1	E	A	2	E	-	-	-
4	SC9740	1	E	A	NA	E	-	-	-
5	SC9613	1	E	A	NA	E	+ [g]	-	-
6	CNUH94-4	1	C	B	1	C	+	-	+
7	V-16	1	C	B	2	C	+	-	+
8	CN8	2	C	B	1	C	+	+	+
9	CN7	1	C	B	1	C	+	+	+
10	YJ016	2	C	B	2	C	+	+	+
11	SC9733	1	E	B	1	C	+	-	+
12	CG122	1	E	B	NA	C	+	+	-
13	CG55	2	E	B	2	C	+	+	+
14	CG108	1	E	B	1	C	+	+	+
15	SC9648	1	E	B	2	C	+	-	+
16	SC9720	1	E	B	2	C	+	-	+
17	SC9761	1	E	A	2	E	+	-	+
18	NV63	2	E	B	2	C	+	+	+
19	SC9737	1	E	B	1	C	+	+	+
20	NV72	2	E	B	2	C	+	+	+
21	CG27	4	E	B	1	C	+	-	+
22	CG26	2	E	B	1	C	+	+	+
23	SC97118	4	E	B	2	C	+	+	+
24	WK22	2	C	B	1	C	+	+	+
25	NV101	1	E	B	NA	C	-	-	-
26	WK20	4	C	B	1	C	+	+	+
27	V-19	1	C	B	1	C	+	+	+
28	SC9717	1	E	B	1	C	+	+	-
29	CG33	1	E	B	1	C	+	-	+
30	CG62	2	E	B	2	C	+	+	+
31	NV42	1	E	B	1	C	+	+	+
32	CG21	1	E	B	1	C	+	+	+
33	SC9794	1	E	B	1	C	+	-	+
34	NV37	3	E	B	1	C	-	+	-
35	SC9730	1	E	B	2	C	+	+	-
36	SC9649	2	E	B	2	C	+	+	+
37	CNUH94-3	1	C	B	1	C	+	-	+
38	SC9721	1	E	B	2	C	+	+	-
39	NV43	1	E	B	1	C	+	+	-
40	NV28	3	E	B	NA	C	+	+	-
41	NV31	2	E	B	NA	C	+	-	-
42	NV18	1	E	B	NA	C	+	+	-
43	SC9766	3	E	B	NA	C	+	-	-
44	WK15	2	C	B	1	C	+	+	+
45	WK6	1	C	B	1	C	+	+	+
46	CG64	1	E	B	1	C	+	+	+

(Continued)

Table 2. (*Continued*)

BOX-PCR profile no.[a]	Representative strain[b]	No. of strains[c]	Source of isolation[d]	16S rRNA type[e]	CPS type[e]	*vcg* type[e]	*pilF*[e]	*viuB*[e]	*vuuA*[e]
47	NV1	3	E	B	1	C	+	+	+
48	SC9793	1	E	B	1	C	+	+	+
49	MO6-24/O	2	C	B	1	C	+	+	+
50	ATCC29307	1	C	B	2	C	+	-	+
51	CMCP6	1	C	B	1	C	+	-	+
52	WK13	2	C	B	1	C	+	-	+

[a] As shown in Fig 1.

[b] Selected randomly, with the exception of the reference clinical strains CMCP6, MO6-24/O, and YJ016.

[c] Number of strains belonging to each BOX-PCR profile.

[d] C, clinical source; E, environmental source.

[e] Determined according to the methods of Nilsson et al. [37], Han et al. [75], Roche et al. [71], Roig et al. [30], Panicker et al. [72], and Kim et al. [73].

[f] Not amplified.

[g] PCR amplicon with the expected size.

doi:10.1371/journal.pone.0142657.t002

The seven multiple sequence alignments were concatenated, and phylogenetic reconstruction was performed from the concatenated sequence of the MLST loci. A phylogenetic tree was inferred from the concatenated sequence using the neighbor-joining (NJ) algorithm implemented in MEGA software (ver. 5) [80]. The evolutionary distances between the sequences were calculated according to the Jukes-Cantor (JC69) substitution model [81]. The tree topology was statistically evaluated by 1,000 bootstrap resamplings and was further confirmed using the maximum likelihood (ML) algorithm (general time reversal [GTR] + gamma [Γ] distribution model) implemented in the high performance computing version of RAxML [82]. Phylogenetic analyses were also performed for individual genes, as described above, to determine the alleles of each gene (S1 Fig).

An allelic type (AT) was defined as a unique combination of polymorphisms within a gene. Each AT was assigned an arbitrary number. A sequence type (ST) was defined as a unique combination of ATs within the concatenated sequences (seven genes combined) according to START (ver. 2) [83]. Each ST was also assigned an arbitrary number.

Tests of neutrality, linkage disequilibrium, and recombination

From the ATs, a number of descriptive properties were determined for each gene. The number of polymorphic sites, the number of parsimonious sites, the nucleotide diversity (π, average pairwise nucleotide difference per site) [84], the average number of nonsynonymous substitutions per nonsynonymous site (dN), the average number of synonymous substitutions per synonymous site (dS), and Tajima's D statistic [85] for each gene were calculated using DNASP (ver. 5.1) [86]. To investigate whether positive (or negative) selection had occurred at the protein level, the ratio of dN to dS was calculated (dN/dS < 1, negative selection; dN/dS = 1, no selection; dN/dS > 1, positive selection). Tajima's D statistic was used to test the neutrality of the observed DNA polymorphisms under the assumption that the Tajima D values exhibit a β distribution (D < 0, high level of low frequency polymorphisms compared with the expected level in a neutral model [possibly due to population size expansion after a bottleneck or selective sweep]; D > 0, low level of polymorphisms compared with the expected level [possibly due to population size contraction or balancing selection]; D = 0, observed level of polymorphisms similar to the expected level [possibly due to genetic drift-mutation equilibrium]) [85, 87].

Multilocus linkage disequilibrium between alleles (nonrandom association of alleles at multiple loci) [88] was examined by determining the index of association ($I_A \neq 0$, disequilibrium; $I_A = [V_O/V_E] - 1$, where V_O = observed variance and V_E = expected variance) using START (ver. 2) [83]. Since I_A depends on sample size (i.e., the number of loci used), the standardized version of I_A ($I^S_A = [V_0/V_E—1]/[L-1]$, where L = the number of loci) [89] was also determined. Statistical significance levels of I_A and I^S_A were measured using a randomization test with 1,000 iterations.

Recombination analyses for the genes analyzed by MLST were performed according to the method described by Martin et al. [90]. The overall evidence of recombination at each locus was assessed by determining the Φ statistic (pairwise homoplasy index, PHI) [91]. The PHI test detects recombination by comparing the frequency of phylogenetically incompatible site-pairs with the frequency of such site pairs expected in the absence of recombination [90]. The PHI test was performed with SPLITSTREE (ver. 4.13) [92]. Window size (W) was set to 100 bases ($\Phi_{W = 100}$), and the statistical significance of Φ_W values were assessed using a permutation test with 1,000 iterations (H_0 [no recombination, $\Phi_W \neq 0$] rejected at $\alpha = 0.05$ in favor of H_1 [recombination]), as recommended by Bruen et al. [91]. Individual recombination events for each locus were inferred from evidence of gene conversions (fragments of DNA sequence that are copied onto another fragment of DNA sequence) [93] as detected by Sawyer's runs test [94]. This test identifies the DNA fragment shared by two sequences via ancestral gene conversion and was performed using GENECONV (www.math.wustl.edu/~sawyer). The statistical significance of detecting the shared fragment (intragenic conversion) was measured using a permutation test with 10,000 iterations. A single recombination event was defined as a group of fragments linked to the same 5' and/or 3' breakpoints, as described by Nightingales et al. [95] and den Bakker et al. [96]. After statistical assessment of recombination, all ST sequences were used to generate a Neighbor-Net network [97] tree depicting the evolutionary relationships between VV strains (ST sequences) with ancestral recombination events [90]. The Neighbor-Net network based on weighted splits in multiple sequence alignment was constructed using SPLITSTREE (ver. 4.13) [92]. The JC69 substitution model [81] used in the bifurcating phylogenetic tree was used to calculate the evolutionary distances in the Neighbor-Net network, and the network topology was evaluated by 1,000 bootstrap resamplings. All other calculations were performed according to the SPLITSTREE manual.

Nucleotide sequence accession numbers

The nucleotide sequences of the 16S RNA gene, *glnA*, *glp*, *gyrB*, *mdh*, *pyrC*, *recA*, and *vvhA* have been deposited at NCBI GenBank under the accession numbers KP223857-KP223908, KP223909-KP223960, KP223961-KP224012, KP224013-KP224064, KP224065-KP224116, KP224117-KP224168, KP224169-KP224220, and KP224221-KP224272, respectively.

Results and Discussion

Collection of biotype 1 strains

Consistent with the finding that biotype 1 strains have been primarily identified as responsible for human infection [1, 2], all clinical strains in our collection belonged to biotype 1. However, all of our environmental strains also belonged to biotype 1, regardless of the isolation source (e.g., seawater, seafood, or tidal mudflat). Unlike the biotype 1 strains, which have been found worldwide in the marine environment (including brackish water) as well as in patients with septicemia [2], VV strains belonging to biotypes 2 and 3 have been reported to occupy specialized ecological niches. Biotype 2 strains were primarily found as eel (genus *Anguilla*) pathogens [65, 77, 98]. Although biotype 3 strains have been reported to cause wound infection in

humans, to date these infections have been limited to persons handling fish (genus *Tilapia*) in Israel at freshwater fish farming sites [66, 99]. Thus, we may not have obtained any biotype 2 or 3 strains because none of our environmental strains were obtained from such specialized environments.

However, the use of 'biotyping' in the taxonomy of VV is somewhat controversial. Recently, Broza [100] challenged the classification of VV strains by the biotype scheme because of the ambiguity of this scheme as observed by Biosca et al. [101] for indole-positive, eel-infecting VV strains. Although indole, ODC, and ONPG findings are certainly useful to some extent and are still widely used for taxonomic tests to differentiate sub-species groups of VV, we agree that classification systems of VV based on only three enzymatic reactions might not properly reflect the exquisite taxonomic patterns of VV. Nonetheless, the scope of our study was somewhat conservative in that it was confined to VV biotype 1. Thus, we recognize that our results might not adequately reflect the entire span of genetic diversity or population structure of VV.

BOX-PCR

Prior to in-depth analyses, the VV strains were subjected to BOX-PCR genomic fingerprinting to identify clonal strains (VV strains with identical BOX-PCR fingerprinting patterns) in our collection (Fig 1). To determine a cut-off value to define a unique BOX-PCR fingerprinting pattern (BOX-PCR type), two strains (CMCP6 and ATCC 29307) were subjected to multiple rounds of BOX-PCR. A comparison of the resulting fingerprinting patterns resolved on independent gels yielded an average similarity coefficient (*r* value) of 0.96 (range, 0.92–0.98), which is consistent with other studies [68, 70, 102]. Thus, a similarity value of 0.90 or more was chosen to indicate strains of the same BOX-PCR type. The cut-off level used in the UPGMA cluster analysis of BOX-PCR genomic fingerprinting patterns corresponded to a genomic DNA-DNA similarity of > 95% [103], suggesting that strains belonging to the same BOX-PCR type in this study can be regarded as nearly identical strains at the genomic DNA level.

Cluster analysis identified a total of 52 BOX-PCR types (Fig 1 and Table 2); the majority (58%) of these were singletons (one member strain per BOX-PCR type), indicating that the intra-species diversity of VV as assessed by BOX-PCR fingerprinting could be higher than that observed in this study (Fig 1 and Table 2). Rarefaction analysis of the BOX-PCR results using the Chao-1 estimator [104, 105] suggested that there could be around 30 additional BOX-PCR types (Chao-1 estimate, 84.7). Based on the numbers of identified BOX-PCR types, the environmental VV strains (41 BOX-PCR types) seemed to be more diverse than the clinical strains (16 BOX-PCR types). However, this simple comparison might not be sufficient to draw an adequate conclusion regarding the relationship between diversity and strain source, since three times more environmental strains than clinical strains were included in our collection. Moreover, the evenness indices (J, normalized form of diversity index [0–1 scale] based on Shannon's entropy H) [106] for the environmental strains and clinical strains were 0.868 and 0.870, respectively (for all strains, J = 0.864), indicating that the diversity of the clinical strains as revealed by BOX-PCR fingerprinting was comparable to that of the environmental strains.

Associations between isolation source and genotypic characteristics

To determine the genotypes of our VV strains, 52 representative strains of BOX-PCR types were subjected to PCR-based genotyping assays. These assays detected polymorphisms at seven genetic loci (the 16S rRNA gene and the *pilF*, *vcg*, *viuB*, *vuuA*, *vvhA*, and CPS alleles) (Table 2) suggested to be predictive of pathogenicity (with the exception of *vvhA*). All VV strains examined in this study were positive for *vvhA*. Since *vvhA* has been shown to be carried by all VV strains isolated to date and could be a hallmark of the VV species itself [107], this
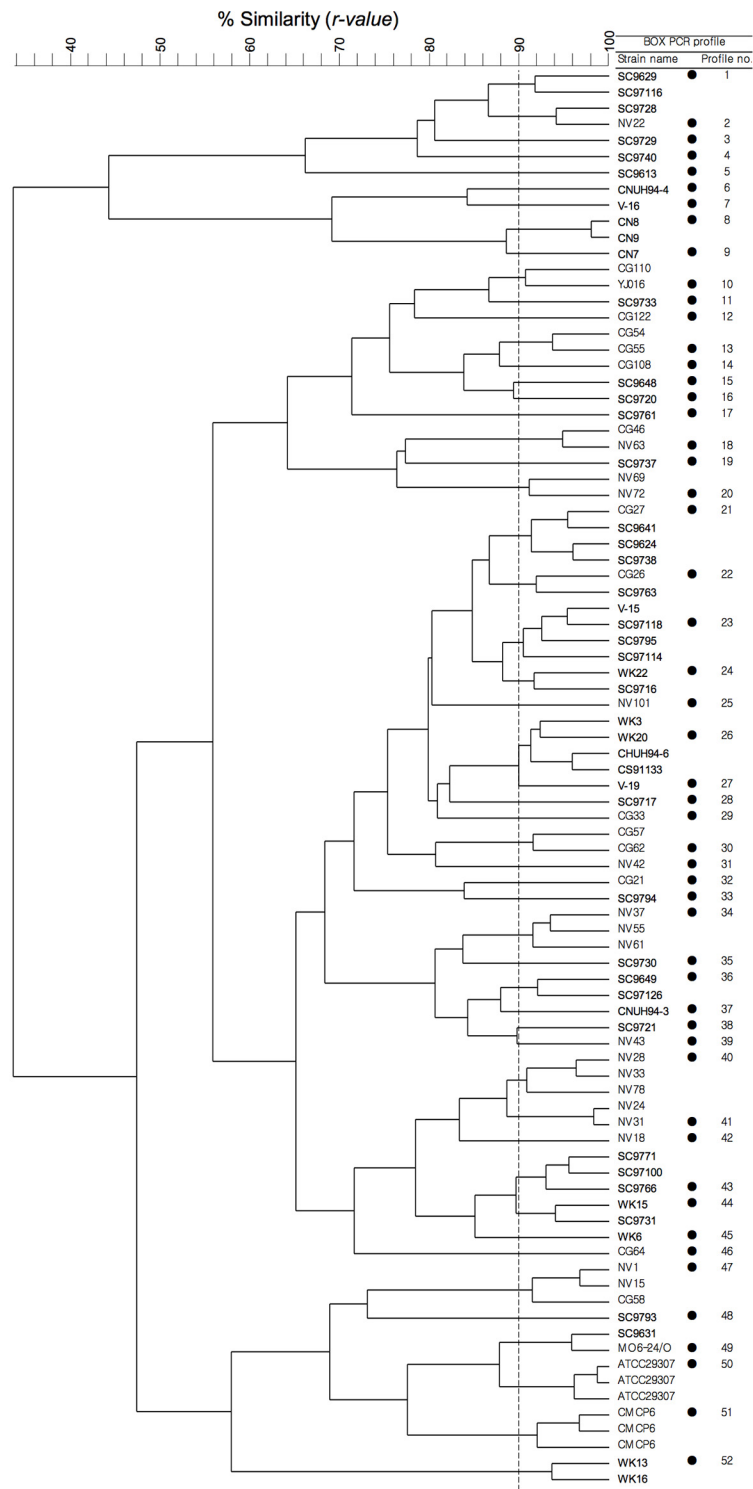
Fig 1. UPGMA cluster analysis of BOX-PCR genomic fingerprints of *Vibrio vulnificus* strains. *r* values are expressed as percentages. The dashed line indicates the cutoff level (90%). Closed circles denote representative strains selected from each of the unique fingerprinting profiles for subsequent analyses.

doi:10.1371/journal.pone.0142657.g001

gene was excluded from further analysis. Specifically, the 16S rRNA genotype B, CPS genotype 1, *vcg* genotype C, and positive PCR results for *pilF*, *viuB*, and *vuuA* have all been reported to be associated with pathogenicity [30, 37, 71–75]. Thus, these genotypes were designated as 'P' genotypes in the present study. In contrast, the alternative genotypes for these loci (16 rRNA genotype A, CPS genotype 2, *vcg* genotype E, and negative PCR results for *pilF*, *viuB*, and *vuuA*) were all designated as 'N' genotypes. Although a single genotype was not assumed to indicate the virulence potential of VV, our designations (P vs N genotypes) have an advantage over the original designations (e.g., A or B, 1 or 2, C or E, positive or negative) in that they provide an unified designation scheme with an indication (P, pathogenic; N, nonpathogenic) of the possible relationship to the virulence potential of VV strain based on the current knowledge.

The majority (> 50%) of the VV strains in our collection were found to have P genotypes for all loci tested (16S rRNA type B, 88.5%; CPS type 1, 51.9%; *vcg* type C, 90.4%; *pilF* +, 88.5%; *viuB* +, 59.6%; *vuuA* +, 62.2%). No hybrid ribotypes for the 16S rRNA genotype (type AB) [108] were observed. The CPS genotype could not be determined for 10 strains (19%) because no PCR amplification products were obtained using the primer pairs suggested by Han et al. [107]. As reported in numerous studies [30–36], all clinical strains in our collection exhibited P genotypes for the 16S rRNA gene, *pilF*, *vcg*, and *vuuA*. However, some (40% and 20%) of the clinical strains exhibited N genotypes for *viuB* and CPS, respectively. This finding is particularly interesting because even the well-studied clinical strains CMCP6 and YJ016 exhibited N genotypes for these two loci. In contrast, the clinical strain MO6-24/O exhibited P genotypes for all loci tested. Another noteworthy finding was that many of the environmental strains tested exhibited P genotypes for a number of loci. Less than half of the environmental strains exhibited N genotypes for *viuB* and *vuuA* loci (40.5% and 43.2%, respectively), and only a small fraction of the environmental strains exhibited N genotypes for the 16S rRNA gene (16.2%), CPS (32.4%), *pilF* (16.2%), and *vcg* (13.5%).

Although many studies have reported that VV genotype correlates with strain origin (i.e., P genotypes with clinical sources and N genotypes with environmental sources) [32, 34, 37, 71, 72, 108–111], our genotyping findings contradict this conclusion. Statistical analysis using the chi-square (χ^2) test ($\alpha = 0.05$) indicated that the association between genotype and strain source was not significant for any genotype, with the exception of *vuuA* (Table 3). For *vuuA*, Fisher's exact test also failed to reject the null hypothesis of random association ($p < 0.05$). While a significant association ($p < 0.05$, χ^2 test; $p < 0.05$, Fisher's exact test) was observed between the *vuuA* genotype and the strain origin, the degree of association expressed as a

Table 3. Associations between *Vibrio vulnificus* strain origin (clinical or environmental) and genotype.

Gene[a]	χ^2 test *p* value	Fisher's exact test *p* value	Association index (simple matching coefficient)
16S rRNA	0.097	0.165	0.404
CPS	0.113	0.180	0.571
pilF	0.097	0.165	0.404
vcg	0.134	0.305	0.385
viuB	0.971	≈ 1	0.462
vuuA	**0.002** [b]	**0.002**	**0.596**

[a] Genes used for genotyping.

[b] Significant associations are designated in bold.

doi:10.1371/journal.pone.0142657.t003

simple matching coefficient indicated that only around 60% of all VV strains showed the expected association (*vuuA*-P types from clinical sources and *vuuA*-N types from environmental sources). A marginal association (p = 0.097) was observed for the 16S rRNA and *pilF* genotypes, with degrees of association < 50%. Our results are consistent with those of a recent study of VV strains isolated in Japan [38]. In this study, P genotypes for the 16S rRNA gene, CPS, and *vcg* dominated both in clinical strains and in environmental strains. In addition, the prevalence of the P genotype for the 16S rRNA gene in environmental VV strains isolated in Korea has also been reported [112]. The genotypes of VV might exhibit distinct geographical patterns of distribution, as noted by Reynaud et al. [36]. This hypothesis could explain the discrepancy between our results and those of previous studies; however, this hypothesis still requires further testing.

Associations among genotypic characteristics

We next determined whether the genotypes of different loci correlated with one another. Significant (p < 0.05, χ^2 test; p < 0.05, Fisher's exact test) associations were observed between the genotypes of all tested loci, with the exception of CPS (Table 4 and S2 Table). The inter-relationships between the 16S rRNA gene, the *pilF* genotype, and the *vcg* genotype were particularly strong. Strikingly, the genotypes of these three loci agreed more than 90% with each other. Consistent with our results, a significant association between the 16S rRNA gene and the *vcg* genotype was reported in a study by Thiaville et al. [34], which was also performed with biotype 1 strains. In their study, 16S rRNA type A and *vcg* type E strains were defined as 'profile 1' strains, whereas 16S rRNA type B and *vcg* type C strains were defined as 'profile 2' strains. They concluded that profile 2 strains were more likely to cause lethal systemic infection, but suggested that genotype alone could not sufficiently predict the pathogenicity of a given VV strain because of the many observed exceptions (e.g., nonpathogenic strains with profile 2). In our study, all clinical strains were profile 2 strains; however, most environmental strains (83.8%) were also profile 2 strains.

Multilocus sequence typing (MLST)

MLST of our VV strains revealed 23–35 (average, 28) allelic types (ATs) per locus analyzed (S1 Fig and S3 Table), resulting in a total of 44 unique combinations of ATs. Each unique combination of ATs was defined as an MLST sequence type (ST) in this study (S3 Table). The nucleotide diversity (π) of each gene (average, 0.021; range, 0.015–0.030) (Table 5) was comparable to MLST data previously obtained for pathogenic bacteria, including VV [36, 95, 113–116]. Among the genes analyzed, the highest nucleotide diversity was observed for *pyrC* (π = 0.030),

Table 4. Associations between genotypic characteristics (lower left half, χ^2 test p value; upper right half, simple matching coefficient).

Gene[a]	16S rRNA	CPS	*pilF*	*vcg*	*viuB*	*vuuA*
16S rRNA		0.690	**0.923**[b]	**0.981**	**0.712**	**0.769**
CPS	0.052		0.643	0.690	0.595	0.643
pilF	**< 0.001**	0.666		**0.904**	**0.673**	**0.808**
vcg	**< 0.001**	0.052	**< 0.001**		**0.692**	**0.750**
viuB	**0.002**	0.495	**0.023**	**0.004**		0.596
vuuA	**0.003**	0.430	**< 0.001**	**0.012**	0.346	

[a] Genes used for genotyping.

[b] Significant associations are designated in bold.

doi:10.1371/journal.pone.0142657.t004

Table 5. Characteristics of the genes analyzed by MLST.

Group	Gene	No. allelic types (ATs)		No. of polymorphic sites	No. of parsimonious site	Average nucleotide diversity (π^a)	dN[b]	dS[c]	dN/dS	Tajima's D[d]
		Total	No. of singletones							
Total	glnA	21	12 (57%)	30	19	0.020	0.001	0.081	0.007	-0.200
	glp	27	16 (59.3%)	53	43	0.023	0.001	0.097	0.008	0.244
	gyrB	25	10 (40%)	38	26	0.018	0.001	0.075	0.011	0.180
	mdh	23	13 (56.5%)	35	25	0.016	0.001	0.058	0.011	0.299
	pyrC	32	24 (75%)	71	53	0.030	0.002	0.116	0.020	0.413
	recA	35	23 (65.7%)	59	40	0.015	0.000	0.061	0.001	-0.455
	vvhA	23	10 (43.5%)	36	31	0.024	0.001	0.101	0.009	0.674
	Average	26.6	15.4 (57.9%)	46	33.9	0.021	0.001	0.084	0.010	
Lineage A	glnA	4	2 (50%)	5	0	0.006	0.000	0.026	0.000	-0.797
	glp	6	6 (100%)	15	7	0.010	0.000	0.044	0.000	-0.034
	gyrB	5	4 (80%)	9	2	0.007	0.000	0.030	0.000	-0.526
	mdh	4	3 (75%)	14	1	0.011	0.001	0.039	0.026	-0.624
	pyrC	6	6 (100%)	11	3	0.007	0.001	0.026	0.026	-0.440
	recA	6	6 (100%)	14	4	0.007	0.000	0.027	0.000	-0.666
	vvhA	4	3 (75%)	5	1	0.006	0.003	0.015	0.180	-0.212
	Average	5	4.3 (86%)	10.4	2.6	0.008	0.001	0.030	0.033	
Lineage B	glnA	17	10 (58.8%)	19	15	0.015	0.000	0.062	0.006	0.266
	glp	20	9 (45%)	40	28	0.017	0.001	0.071	0.013	-0.318
	gyrB	19	5 (26.3%)	26	15	0.012	0.000	0.052	0.004	-0.475
	mdh	19	10 (52.6%)	26	19	0.014	0.000	0.054	0.004	0.461
	pyrC	25	17 (68%)	58	36	0.022	0.003	0.082	0.031	-0.320
	recA	28	16 (57.1%)	41	23	0.009	0.000	0.038	0.003	-1.005
	vvhA	18	6 (33.3%)	22	19	0.015	0.001	0.064	0.008	0.335
	Average	20.9	10.4 (49.8%)	33.1	22.1	0.015	0.001	0.060	0.010	

[a] Average pairwise nucleotide difference per site.

[b] Average number of nonsynonymous substitutions per nonsynonymous site.

[c] Average number of synonymous substitutions per synonymous site.

[d] No Tajima's D value was significantly deviated from zero ($p > 0.10$).

doi:10.1371/journal.pone.0142657.t005

which also contained the largest number (n = 71) of polymorphic sites ([Table 5]). On the other hand, *recA* showed the lowest nucleotide diversity ($\pi = 0.015$), even though it had the largest number (n = 35) of ATs. Despite its low nucleotide diversity, the high allelic diversity of *recA* can be explained by the even distribution over many polymorphic sites of a relatively small number of nucleotide substitutions in the *recA* sequence, which also resulted in *recA* having the highest number of singleton ATs. The dN/dS ratios for all genes analyzed were much lower than 1 (average, 0.010; range, 0.001–0.020), indicating that synonymous substitutions predominated over nonsynonymous substitutions ([Table 5]). Thus, we conclude that negative (purifying) selection acted against amino acid substitutions in these genes (negative selection on the protein level). This conclusion was expected, because six of the seven genes analyzed by MLST were functionally constrained housekeeping genes. A putative virulence-associated gene, *vvhA*, was also shown to be under negative selection pressure (dN/dS = 0.009). The Tajima's D values for all genes analyzed fell within a confidence interval of zero (e.g., -1.8 < 95% CI < +2.0 and -1.6 < 90% CI < +1.7 for 28 ATs), under the assumption that the Tajima's D values followed a β distribution [85], and hence were considered insignificant ([Table 5]). Thus, mutations on the

DNA level with no effect on fitness can explain most of the observed polymorphisms. In fact, as expected from the very low dN values (0.000–0.002), the amino acid sequences encoded by each gene showed at least 99.5% pairwise identity.

When all 44 STs (concatenated sequences) were subjected to phylogenetic reconstruction, two major phylogenetic groups, each of which exhibited a bootstrap confidence level > 80%, were observed in both the NJ tree and the ML tree (Fig 2A). This subdivision of VV into two intra-species lineages is consistent with previous studies [32, 36, 37, 71, 113]. One group was comprised of six STs that originated exclusively from environmental sources, while the other group was comprised of 37 STs that originated either from environmental sources or from clinical sources (Table 1 and S3 Table). We designated the former group as MLST lineage A and the latter group as MLST lineage B; these two MLST lineages corresponded to previously recognized lineages of VV [115] (S2 Fig). We did not consider ST-30 (representative strain, SC9733) as belonging to either of these monophyletic groups because of the low bootstrap confidence level (< 50%) of the internal node of ST-30. In spite of the close relationship between ST-30 (average genetic distance, 0.016 ± 0.002) and the MLST B lineage, no assignment was made for ST-30 in order to avoid making either lineage polyphyletic. As mentioned above for the strain sources in each of the MLST lineages, the association between MLST lineage (A or B) and isolation source (clinical or environmental) was not significant ($p = 0.146$, χ^2 test; $p = 0.309$, Fisher's exact test). However, significant ($p < 0.05$, χ^2 test; $p < 0.05$, Fisher's exact test) associations were observed between MLST lineage and genotype. Specifically, the N-type and P-type strains predominated in MLST lineages A and B, respectively (Pearson's association coefficient Φ = 0.409–1.000). The MLST A and B lineages were comprised exclusively of N-type strains and P-type strains, respectively, especially for the 16S rRNA gene. The average nucleotide diversity of MLST lineage B ($\pi = 0.015 \pm 0.004$) was significantly greater than that of MLST lineage A ($\pi = 0.008 \pm 0.002$) ($p < 0.05$, t-test) (Table 5) as reflected in their phylogenetic branching patterns (Fig 2A). The inter-group genetic distance between lineages A and B (0.033 ± 0.002) exceeded the average nucleotide diversities of both lineages. The average dN/dS ratio of MLST lineage A was slightly higher than that of MLST lineage B (Table 5), although this difference was not significant ($p > 0.05$, t-test). Interestingly, MLST lineage B appeared to be further divided into two sub-lineages, one containing the reference strains CMCP6 (ST-17) and MO6-24/O (ST-1) and the other one containing the reference strain YJ016 (ST-22). However, these sub-groupings exhibited low bootstrap confidence levels (< 50%).

Linkage analysis

We found that VV was in linkage disequilibrium; that is, VV exhibited a clonal population structure with infrequently recombining 'clonal' STs as opposed to frequently recombining 'sexual' STs. When all 44 STs were tested for linkage disequilibrium between alleles using the index of association (I_A) [88], both the conventional and standardized I_A values were significantly different from zero ($I_A = 0.643$, $p < 0.05$; $I^S_A = 0.107$, $p < 0.05$), indicating nonrandom association between alleles (linkage disequilibrium) (Table 6). Maynard-Smith et al. described a continuum of bacterial population structures, ranging from strictly clonal (e.g., *Neisseria gonorrhoeae*, $I_A = 0.04$) to panmictic (e.g., *Pseudomonas syringae*, $I_A = 18.35$) [88]. Compared with these data, VV appears to exhibit weakly clonal population structure. However, the degrees of linkage disequilibrium (I_A) estimated by Maynard-Smith et al. [88] were calculated based on electrophoretic types (ETs), which were determined by multilocus enzyme electrophoresis (MLEE). Since these values ultimately depend on amino acid sequence, the degree of linkage disequilibrium measured for VV in the present study could be underestimated (skewed toward linkage equilibrium). This is because our I_A values were calculated based on STs that
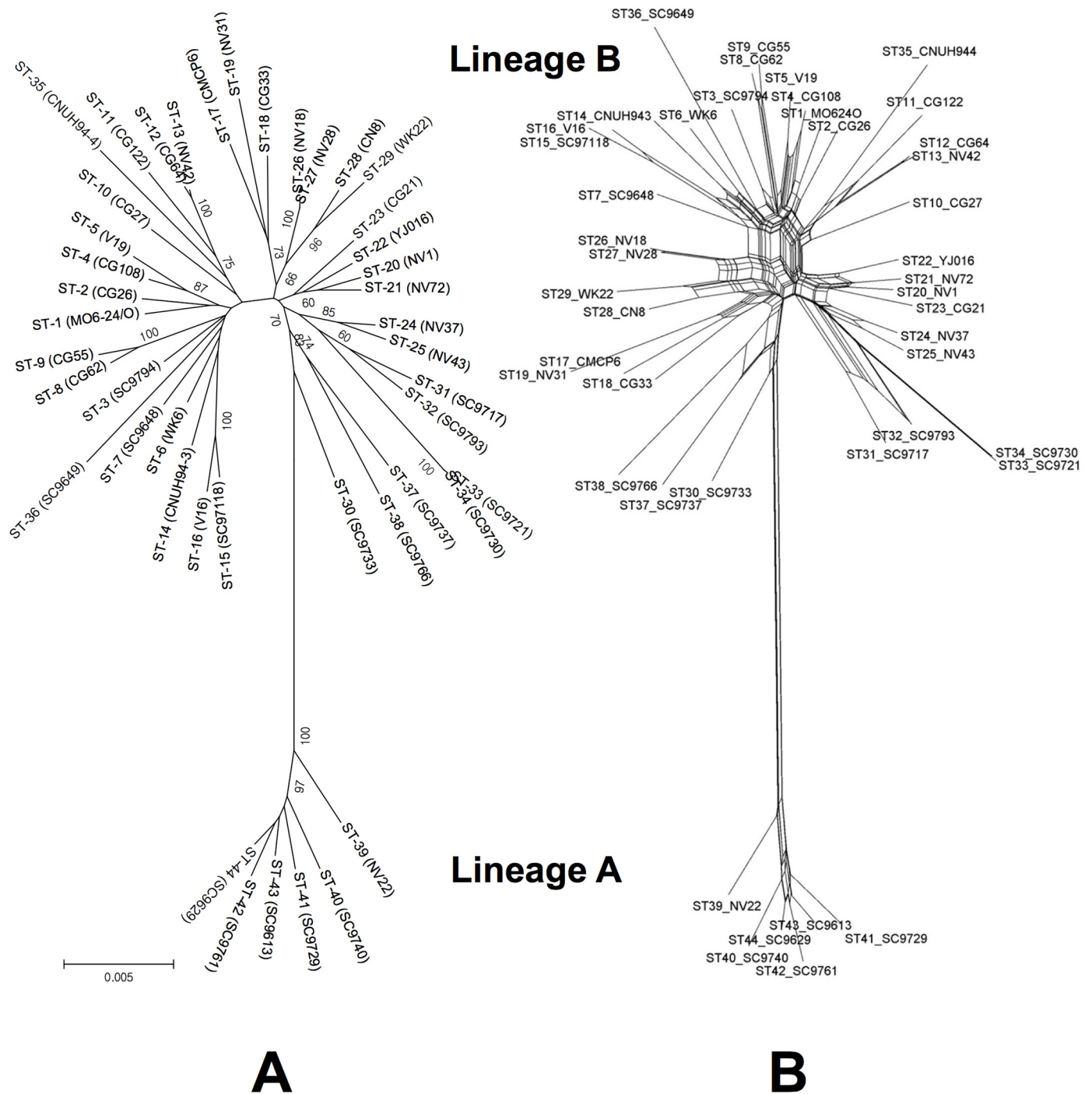
Fig 2. Phylogenetic relationships of *V. vulnificus* strains based on the concatenated sequences of seven genetic loci. The phylogenetic distances of each concatenated sequence were calculated using the Jukes-Cantor (JC69) model, and the trees were constructed using the neighbor-joining (NJ) algorithm and the neighbor-net network (NN) algorithm (panels A and B, respectively). The numbers at the nodes in the NJ tree indicate the bootstrap scores (as percentages) and are shown for frequencies at or above the threshold of 50%. Bootstrap scores are not shown in the NN tree for tree legibility, but are given in the text. The scale bar represents the expected number of substitutions per nucleotide position.

doi:10.1371/journal.pone.0142657.g002

Table 6. Multilocus linkage disequilibrium analysis of *Vibrio vulnificus*.

	V_O	V_E	I_A (*p* value)	I^S_A (*p* value)
All sequence types (STs)	0.539	0.328	**0.643 (*p* < 0.001)[a]**	**0.107 (*p* < 0.001)**
Lineage A STs	0.543	0.498	0.091 (*p* = 0.619)	0.015 (*p* = 0.572)
Lineage B STs	0.676	0.425	**0.592 (*p* < 0.001)**	**0.099 (*p* < 0.001)**

V_O, observed variance of the number of loci at which two STs differ (K) [88].

V_E, expected variance of K [88].

I_A, index of association (I_A = [V_O/V_E]-1) [88].

I^S_A, standardized index of association ([I^S_A = I_A/[L-1], where L = the number of loci) [89].

[a] Significant linkage disequilibrium groups are marked in bold.

doi:10.1371/journal.pone.0142657.t006

were determined by MLST, which depends on nucleotide sequence. Boyd et al. [117] reported that changes of around 26 nucleotides were required to influence the electrophoretic mobility of housekeeping enzymes. Despite this underestimation, the I_A for VV determined in the present study was significantly deviated from zero. We hypothesize that VV would exhibit a highly clonal population structure if its I_A value was determined from ETs, because only a very limited number of ETs are available for each locus due to the high level of amino acid sequence identity between ATs (99.8–100%, almost fixed alleles on the protein level).

Consistent with our results, Bisharat et al. [115] reported a clonal population structure for VV, although a different association statistic was used for the linkage analysis. A closely related species, *V. cholerae*, has also been shown to exhibit a clonal population structure [118] with an I_A value (I^S_A = 0.143, p < 0.05) close to that observed in the present study. The clonal population structure observed for the VV STs in the present study could be explained by insufficient recombination between STs, resulting in a high level of association between alleles, as proposed by Maynard-Smith et al. [119]. The rate of recombination of large chromosomal segments might not be sufficient to randomize the genomes or separate the clonal association of the VV strains analyzed in the present study. Interestingly, when the two MLST lineages were considered separately, the STs belonging to MLST lineage A showed no evidence of linkage disequilibrium (I^S_A = 0.015, *p* > 0.05), while MLST lineage B still showed significant linkage disequilibrium (I^S_A = 0.099, *p* < 0.05). However, we cannot exclude the possibility that the linkage equilibrium observed for MLST lineage A might be due to insufficient sampling, as reflected by the high frequency of singleton ATs.

Inconsistent with the observed results of linkage analysis, ancestral recombination events in the genes used in MLST were detected with two different tests for recombination, the PHI test for overall evidence of recombination and Sawyer's runs test for individual recombination events [91, 94] (Table 7). The PHI test detected significant evidence of recombination (permutation *p* value < 0.05) in *glp*, *gyrB*, *mdh*, *pyrC*, and *vvhA* when all STs were tested. The same result was observed for MLST lineage B, but the PHI statistic could not be calculated for a number of genes in MLST lineage A due to the lack of phylogenetically informative (parsimonious) sites (Table 5). Sawyer's test, a more conservative test for recombination events, also detected evidence of recombination for four genes (*glnA*, *glp*, *gyrB*, and *pyrC*) (Table 7). However, breakpoint analysis identified only a limited number of recombination events. Although 1–25 DNA fragments (segments) were shared by paired sequences via ancestral gene conversions, only 1–4 (and usually only one) recombination events were identified within MLST lineage B and between MLST lineages A and B. Moreover, Sawyer's test did not detect any recombination events within MLST lineage A. This finding supports our skeptical view for the linkage equilibrium observed for MLST lineage A.

Table 7. Analysis of overall recombination (PHI test) and individual recombination events (Sawyers' runs test).

Gene	PHI test *p* value			Sawyer's runs test[a]						
	Total	Within lineages		Simulated *p* value[b]	No. of fragments[c]	No. of recombination events[d]				
		A	B			Total	Within lineages		Between lineages	Multiple[e]
							A	B	A-B	
glnA	0.213	ND[f]	0.836	0.013	1	1	0	1	0	0
glp	**0.001**[g]	0.004	**< 0.001**	**0.001**	25	1	0	0	0	1
gyrB	**< 0.001**	ND	**< 0.001**	0.005	8	1	0	0	1	0
mdh	0.041	ND	**0.003**	0.561	0	0	0	0	0	0
pyrC	**< 0.001**	1.000	**< 0.001**	**0.002**	12	4	0	3	1	0
recA	0.490	0.256	0.901	0.109	0	0	0	0	0	0
vvhA	**0.002**	ND	**< 0.001**	0.274	0	0	0	0	0	0
Sum					46	7	0	4	2	1

[a] Gene conversion test for recombination between ancestors of sequences in an alignment (global inner recombination) [94, 95]. Global outer recombination events were not detected.

[b] Global *p* value obtained using 10,000 permutations.

[c] Fragment shared by two sequences in the alignment via ancestral gene conversion [94]. Only fragments that significantly (*p* < 0.05) implied recombination events were counted.

[d] Groups of fragments linked to the same 5' and/or 3' breakpoints were classified as a single recombination event as described by Nightingales et al. [95] and den Bakker et al. [96].

[e] Recombination events both between and within lineages.

[f] Not determined due to the lack of parsimonious sites.

[g] Significant (*p* < 0.05) support for recombination.

doi:10.1371/journal.pone.0142657.t007

Concluding Remarks

We hypothesize that the clonal population structure of VV, which comprised STs with some recombination potential, may be influenced by ecological factors that restrict the breakdown of linkage disequilibrium. Consistent with this hypothesis, *V. cholerae* and *Haemophilus influenza* have been shown to maintain clonal population structures, even though these organisms are naturally transformable [88, 118, 120]. Alternatively, VV could be comprised of distinct eco-types [121], between which recombination is limited. According to the ecotype model of bacterial species, in which distinct ecotypes evolve primarily by periodic selection or selective sweep, alleles are likely to be in significant linkage disequilibrium [122]. The two MLST lineages (or certain ST groups) of VV might correspond to two (or multiple) ecotypes of VV, either of which could be a derived/ancestral ecotype. Although Tajima's D values in the present study were not significantly deviated from zero (Table 5), the positive values observed for all VV STs are consistent with this hypothesis because Tajima's D test tends to produce positive values for a sub-dividing population. Neighbor-Net network analysis further revealed that recombination between MLST lineages A and B was relatively rare compared with recombination within MLST lineage B (Fig 2B). Considering the star-shaped topology and the many cyclic branches near the root of MLST lineage B, this lineage may have undergone rapid clonal expansion after a period of frequent recombination. We speculate that MLST lineage B, or at least some STs in MLST lineage B, correspond to the pathogenic lineage (or pathogenic ecotypes) of VV, since all of our clinical strains belonged to MLST lineage B, and P genotypes predominated in MLST lineage B. The predominance of lineage B strains with P genotype characteristics in our strain

collection also suggests the prevalence of potentially pathogenic strains in the marine environment in East Asia. However, for this suggestion to be experimentally supported, the pathogenicity of our strains should be determined using an animal model. In addition, a more extensive MLST analysis of VV, incorporating many more genetic loci, would help clarify the population structure of VV. Inferring phylogenetic relationships from the core genome of VV as constructed from both clinical strains and environmental strains would also help meet this goal. Until recently, studies of VV have focused mainly on clinical strains, especially the well-characterized strains CMCP6, MO6-24/O, and YJ016. It is therefore very important for future studies to include a comprehensive group of environmental strains and to integrate in vivo pathogenicity results with genomic data. This approach will enable a better understanding of the evolutionary patterns and pathogenesis of this notorious marine pathogen.

Supporting Information

S1 Fig. Phylogenetic relationships of *V. vulnificus* strains based on the 16S rRNA gene (A), *glnA* (B), *glp* (C), *gyrB* (D), *mdh* (E), *pyrC* (F), *recA* (G), and *vvhA* (H). The phylogenetic distances of each concatenated sequence were calculated using the Jukes-Cantor (JC69) model, and the trees were constructed using the neighbor-joining (NJ) method. The numbers at the nodes in the NJ trees indicate the bootstrap scores (as percentages) and are shown for frequencies at or above the threshold of 50%. The scale bar represents the expected number of substitutions per nucleotide position.
(TIFF)

S2 Fig. Phylogenetic tree showing that MLST lineages A and B identified in this study correspond to previously recognized lineages of *V. vulnificus*. The tree includes five additional reference strains (strains 101/4, 2322, 99–796 DP-E7, ATL6-1306, and ATL-71503) belonging to MLST lineage A in addition to strains in our collection (the strains included in [Fig 2]). The lineage information and multi-locus (*glnA*, *glp*, *gyrB*, *mdh*, *pyrC*, *recA*, and *vvhA*) sequences of the additional reference strains were obtained from Bisharat et al. (115) and genome sequences deposited in the NCBI GenBank database (GenBank accession numbers JQDT01000016, JQDS01000033, JSWN01000066, JSWO01000069 and JSWP01000065) (124, 125), respectively. All reference strains are marked by closed circles. The phylogenetic distances were calculated using the JC69 model, and the tree was constructed using the NJ method. The numbers at the nodes indicate the bootstrap scores (as percentages) and are shown for frequencies at or above a threshold of 50%. The scale bar represents the expected number of substitutions per nucleotide position.
(TIFF)

S1 Table. Genes used for multilocus sequence typing (MLST) of *V. vulnificus* strains.
(DOCX)

S2 Table. Associations between genotypic characteristics (lower left half, Fisher's exact test p value; upper right half, simple matching coefficient).
(DOCX)

S3 Table. Allelic profiles of the *V. vulnificus* strains selected based on their BOX-PCR results.
(DOCX)

Acknowledgments

and HUFS (program no. HUFS-2014). H-J Kim was supported in part by the NRF (grant no. NRF-2012R1A1A2004270) and KEITI (GAIA project no. 20130005501).

Author Contributions

Conceived and designed the experiments: JCC. Performed the experiments: HJK. Analyzed the data: HJK JCC. Contributed reagents/materials/analysis tools: JCC. Wrote the paper: HJK JCC.

References

1. Jones MK, Oliver JD. *Vibrio vulnificus*: disease and pathogenesis. Infect Immun. 2009; 77:1723–1733. doi: 10.1128/IAI.01046-08 PMID: 19255188

2. Horseman MA, Surani S. A comprehensive review of *Vibrio vulnificus*: an important cause of severe sepsis and skin and soft-tissue infection. Int J Infect Dis. 2011; 15:e157–166. doi: 10.1016/j.ijid.2010.11.003 PMID: 21177133

3. Strom MS, Paranjpye RN. Epidemiology and pathogenesis of *Vibrio vulnificus*. Microbes Infect. 2000; 2:177–188. PMID: 10742690

4. Linkous DA, Oliver JD. Pathogenesis of *Vibrio vulnificus*. FEMS Microbiol Lett. 1999; 174:207–214. PMID: 10339810

5. Hlady WG, Klontz KC. The epidemiology of *Vibrio* infections in Florida, 1981–1993. J Infect Dis. 1996; 173:1176–1183. PMID: 8627070

6. WHO. Risk assessment of *Vibrio vulnificus* in raw oysters: interpretative summary and technical report. Geneva, Switzerland: 2005.

7. Oliver JD. Wound infections caused by *Vibrio vulnificus* and other marine bacteria. Epidemiol Infect. 2005; 133:383–391. PMID: 15962544

8. CDC. Summary of human *Vibrio* isolates reported to CDC, 2007. Atlanta, GA: 2007.

9. Tsao CH, Chen CC, Tsai SJ, Li CR, Chao WN, Chan KS, et al. Seasonality, clinical types and prognostic factors of *Vibrio vulnificus* infection. J Infect Dev Ctries. 2013; 7:533–540. doi: 10.3855/jidc.3008 PMID: 23857388

10. Hsueh PR, Lin CY, Tang HJ, Lee HC, Liu JW, Liu YC, et al. *Vibrio vulnificus* in Taiwan. Emerg Infect Dis. 2004; 10:1363–1368. PMID: 15496235

11. Inoue Y, Matui T, Ono T. An outbreak of *Vibrio vulnificus* infection in Kumamoto, Japan, 2001. Arch Dermatol. 2004; 140:888–889. PMID: 15262710

12. Osaka K, Komatsuzaki M, Takahashi H, Sakano S, Okabe N. *Vibrio vulnificus* septicaemia in Japan: an estimated number of infections and physicians' knowledge of the syndrome. Epidemiol Infect. 2004; 132:993–996. PMID: 15473164

13. Cazorla C, Guigon A, Noel M, Quilici ML, Lacassin F. Fatal *Vibrio vulnificus* infection associated with eating raw oysters, New Caledonia. Emerg Infect Dis. 2011; 17:136–137. doi: 10.3201/eid1701.100603 PMID: 21192878

14. Ruppert J, Panzig B, Guertler L, Hinz P, Schwesinger G, Felix SB, et al. Two cases of severe sepsis due to *Vibrio vulnificus* wound infection acquired in the Baltic Sea. Eur J Clin Microbiol Infect Dis. 2004; 23:912–915. PMID: 15599654

15. Torres L, Escobar S, Lopez AI, Marco ML, Pobo V. Wound infection due to *Vibrio vulnificus* in Spain. Eur J Clin Microbiol Infect Dis. 2002; 21:537–538. PMID: 12172745

16. Franca JC, Raboni SM, Sanfelice E, Polido D, Gentili A, Marques F. *Vibrio vulnificus* infection in Southern Brazil-case report. An Bras Dermatol. 2013; 88:424–426. doi: 10.1590/abd1806-4841.20131780 PMID: 23793212

17. Oliver JD. Vibrio vulnificus. Doyle MP, editor. New York: Marcel Dekker; 1989.

18. Yoshida S, Ogawa M, Mizuguchi Y. Relation of capsular materials and colony opacity to virulence of *Vibrio vulnificus*. Infect Immun. 1985; 47:446–451. PMID: 2578434

19. Simpson LM, White VK, Zane SF, Oliver JD. Correlation between virulence and colony morphology in *Vibrio vulnificus*. Infect Immun. 1987; 55:269–272. PMID: 2432016

20. McPherson VL, Watts JA, Simpson LM, Oliver JD. Physiological effects of the lipopolysaccharide of *Vibrio vulnificus* on mice and rats. Microbios. 1991; 67:141–149. PMID: 1779875

21. Park KH, Kim JS, Lee YR, Moon YJ, Hur H, Choi YH, et al. Low-density lipoprotein protects *Vibrio vulnificus*-induced lethality through blocking lipopolysaccharide action. Exp Mol Med. 2007; 39:673–678. PMID: 18059143

22. Gray LD, Kreger AS. Purification and characterization of an extracellular cytolysin produced by *Vibrio vulnificus*. Infect Immun. 1985; 48:62–72. PMID: 3980095

23. Gray LD, Kreger AS. Mouse skin damage caused by cytolysin from *Vibrio vulnificus* and by *V. vulnificus* infection. J Infect Dis. 1987; 155:236–241. PMID: 3805762

24. Lee JH, Kim MW, Kim BS, Kim SM, Lee BC, Kim TS, et al. Identification and characterization of the *Vibrio vulnificus rtxA* essential for cytotoxicity in vitro and virulence in mice. J Microbiol. 2007; 45:146–152. PMID: 17483800

25. Kim YR, Lee SE, Kook H, Yeom JA, Na HS, Kim SY, et al. *Vibrio vulnificus* RTX toxin kills host cells only after contact of the bacteria with host cells. Cell Microbiol. 2008; 10:848–862. PMID: 18005241

26. Wright AC, Simpson LM, Oliver JD. Role of iron in the pathogenesis of *Vibrio vulnificus* infections. Infect Immun. 1981; 34:503–507. PMID: 7309236

27. Paranjpye RN, Strom MS. A *Vibrio vulnificus* type IV pilin contributes to biofilm formation, adherence to epithelial cells, and virulence. Infect Immun. 2005; 73:1411–1422. PMID: 15731039

28. Kim YR, Rhee JH. Flagellar basal body *flg* operon as a virulence determinant of *Vibrio vulnificus*. Biochem Biophys Res Commun. 2003; 304:405–410. PMID: 12711330

29. Lee JH, Rho JB, Park KJ, Kim CB, Han YS, Choi SH, et al. Role of flagellum and motility in pathogenesis of *Vibrio vulnificus*. Infect Immun. 2004; 72:4905–4910. PMID: 15271959

30. Roig FJ, Sanjuán E, Llorens A, Amaro C. *pilF* polymorphism-based PCR to distinguish *Vibrio vulnificus* strains potentially dangerous to public health. Appl Environ Microbiol. 2010; 76:1328–1333. doi: 10.1128/AEM.01042-09 PMID: 20038687

31. Baker-Austin C, Lemm E, Hartnell R, Lowther J, Onley R, Amaro C, et al. *pilF* polymorphism-based real-time PCR to distinguish *Vibrio vulnificus* strains of human health relevance. Food Microbiol. 2012; 30:17–23. doi: 10.1016/j.fm.2011.09.002 PMID: 22265278

32. Chatzidaki-Livanis M, Hubbard MA, Gordon K, Harwood VJ, Wright AC. Genetic distinctions among clinical and environmental strains of *Vibrio vulnificus*. Appl Environ Microbiol. 2006; 72:6136–6141. PMID: 16957239

33. Gonzalez-Escalona N, Whitney B, Jaykus LA, DePaola A. Comparison of direct genome restriction enzyme analysis and pulsed-field gel electrophoresis for typing of *Vibrio vulnificus* and their correspondence with multilocus sequence typing data. Appl Environ Microbiol. 2007; 73:7494–7500. PMID: 17720834

34. Thiaville PC, Bourdage KL, Wright AC, Farrell-Evans M, Garvan CW, Gulig PA. Genotype is correlated with but does not predict virulence of *Vibrio vulnificus* biotype 1 in subcutaneously inoculated, iron dextran-treated mice. Infect Immun. 2011; 79:1194–1207. doi: 10.1128/IAI.01031-10 PMID: 21199909

35. Bier N, Bechlars S, Diescher S, Klein F, Hauk G, Duty O, et al. Genotypic diversity and virulence characteristics of clinical and environmental *Vibrio vulnificus* isolates from the Baltic Sea region. Appl Environ Microbiol. 2013; 79:3570–3581. doi: 10.1128/AEM.00477-13 PMID: 23542621

36. Reynaud Y, Pitchford S, De Decker S, Wikfors GH, Brown CL. Molecular typing of environmental and clinical strains of *Vibrio vulnificus* isolated in the northeastern USA. PLoS One. 2013; 8:e83357. doi: 10.1371/journal.pone.0083357 PMID: 24386187

37. Nilsson WB, Paranjype RN, DePaola A, Strom MS. Sequence polymorphism of the 16S rRNA gene of *Vibrio vulnificus* is a possible indicator of strain virulence. J Clin Microbiol. 2003; 41:442–446. PMID: 12517889

38. Yokochi N, Tanaka S, Matsumoto K, Oishi H, Tashiro Y, Yoshikane Y, et al. Distribution of virulence markers among *Vibrio vulnificus* isolates of clinical and environmental origin and regional characteristics in Japan. PLoS One. 2013; 8:e55219. doi: 10.1371/journal.pone.0055219 PMID: 23383115

39. Baizabal-Ramirez O, Negrete-Perez M, Guerrero-Daza D, Martinez-Herrera N, Aburto-Desachy Y, Mata-Miranda P. Septic shock by *Vibrio vulnificus* at the coast Gulf of Mexico. Rev Med Inst Mex Seguro Soc. 2011; 49:433–436. PMID: 21982195

40. Shapiro RL, Altekruse S, Hutwagner L, Bishop R, Hammond R, Wilson S, et al. The role of Gulf Coast oysters harvested in warmer months in *Vibrio vulnificus* infections in the United States, 1988–1996. *Vibrio* Working Group. J Infect Dis. 1998; 178:752–759. PMID: 9728544

41. Motes ML, DePaola A, Cook DW, Veazey JE, Hunsucker JC, Garthright WE, et al. Influence of water temperature and salinity on *Vibrio vulnificus* in Northern Gulf and Atlantic Coast oysters (*Crassostrea virginica*). Appl Environ Microbiol. 1998; 64:1459–1465. PMID: 9546182

42. Penman AD, Lanier DC Jr., Avara WT 3rd, Canant KE, DeGroote JW, Brackin BT, et al. *Vibrio vulnificus* wound infections from the Mississippi Gulf coastal waters: June to August 1993. South Med J. 1995; 88:531–533. PMID: 7732441

43. Inoue Y, Ono T, Matsui T, Miyasaka J, Kinoshita Y, Ihn H. Epidemiological survey of *Vibrio vulnificus* infection in Japan between 1999 and 2003. J Dermatol. 2008; 35:129–139. doi: 10.1111/j.1346-8138.2008.00432.x PMID: 18346255

44. Oishi H, Ura Y, Mitsumizo S, Nakashima M. A collective review of *Vibrio vulnificus* infection in Japan. Kansenshogaku Zasshi. 2006; 80:680–689. PMID: 17176855

45. Lee SH, Chung BH, Lee WC. Retrospective analysis of epidemiological aspects of *Vibrio vulnificus* infections in Korea in 2001–2010. Jpn J Infect Dis. 2013; 66:331–333. PMID: 23883847

46. Oliver JD, Warner RA, Cleland DR. Distribution and ecology of *Vibrio vulnificus* and other lactose-fermenting marine vibrios in coastal waters of the southeastern United States. Appl Environ Microbiol. 1982; 44:1404–1414. PMID: 7159083

47. Leon Robles A, Acedo Felix E, Gomez-Gil B, Quinones Ramirez EI, Nevarez-Martinez M, Noriega-Orozco L. Relationship of aquatic environmental factors with the abundance of *Vibrio cholerae*, *Vibrio parahaemolyticus*, *Vibrio mimicus* and *Vibrio vulnificus* in the coastal area of Guaymas, Sonora, Mexico. J Water Health. 2013; 11:700–712. doi: 10.2166/wh.2013.160 PMID: 24334844

48. Jeong HS, Kim JY, Jeon SM, Park MS, Kim SH. Genotypic characterization of *Vibrio vulnificus* clinical isolates in Korea. Public Health Res Perspect. 2011; 2(1):8–14.

49. Chu C, Do Y, Kim Y, Saito Y, Lee SD, Park H, et al. Mathematical modeling of *Vibrio vulnificus* infection in Korea and the influence of global warming. Osong Public Health Res Perspect. 2011; 2:51–58. doi: 10.1016/j.phrp.2011.05.002 PMID: 24159451

50. KCDC. KCDC's annual report 2007. Korea: 2010.

51. Kim YR, Lee SE, Kim CM, Kim SY, Shin EK, Shin DH, et al. Characterization and pathogenic significance of *Vibrio vulnificus* antigens preferentially expressed in septicemic patients. Infect Immun. 2003; 71:5461–5471. PMID: 14500463

52. Wright AC, Simpson LM, Oliver JD, Morris JG Jr. Phenotypic evaluation of acapsular transposon mutants of *Vibrio vulnificus*. Infect Immun. 1990; 58:1769–1773. PMID: 2160432

53. Chen CY, Wu KM, Chang YC, Chang CH, Tsai HC, Liao TL, et al. Comparative genome analysis of *Vibrio vulnificus*, a marine pathogen. Genome Res. 2003; 13:2577–2587. PMID: 14656965

54. Jones JL, Ludeke CH, Bowers JC, DePaola A. Comparison of plating media for recovery of total and virulent genotypes of *Vibrio vulnificus* in U.S. market oysters. Int J Food Microbiol. 2013; 167:322–327. doi: 10.1016/j.ijfoodmicro.2013.09.017 PMID: 24184610

55. Wood RR, Arias CR. Distribution and survival of *Vibrio vulnificus* genotypes in postharvest Gulf Coast (USA) oysters under refrigeration. J Appl Microbiol. 2012; 113:172–180. doi: 10.1111/j.1365-2672.2012.05306.x PMID: 22507030

56. Gordon KV, Vickery MC, DePaola A, Staley C, Harwood VJ. Real-time PCR assays for quantification and differentiation of *Vibrio vulnificus* strains in oysters and water. Appl Environ Microbiol. 2008; 74:1704–1709. doi: 10.1128/AEM.01100-07 PMID: 18245234

57. Drake SL, Whitney B, Levine JF, DePaola A, Jaykus LA. Correlation of mannitol fermentation with virulence-associated genotypic characteristics in *Vibrio vulnificus* isolates from oysters and water samples in the Gulf of Mexico. Foodborne Pathog Dis. 2010; 7:97–101. doi: 10.1089/fpd.2009.0362 PMID: 19785540

58. Broza YY, Danin-Poleg Y, Lerner L, Broza M, Kashi Y. *Vibrio vulnificus* typing based on simple sequence repeats: insights into the biotype 3 group. J Clin Microbiol. 2007; 45:2951–2959. PMID: 17652479

59. Bisharat N, Cohen DI, Harding RM, Falush D, Crook DW, Peto T, et al. Hybrid *Vibrio vulnificus*. Emerg Infect Dis. 2005; 11:30–35. PMID: 15705319

60. Shao C-P, Hor L-I. Metalloprotease is not essential for *Vibrio vulnificus* virulence in mice. Infect Immun. 2000; 68:3569–3573. PMID: 10816513

61. Hollis DG, Weaver RE, Baker CN, Thornsberry C. Halophilic *Vibrio* species isolated from blood cultures. J Clin Microbiol. 1976; 3:425–431. PMID: 1262454

62. Kim KY, Yang HC, Tamplin ML, Choi SH. Intraspecific variation of environmental and clinical *Vibrio vulnificus* isolates as demonstrated by restriction endonuclease digestion profiles. J Microbiol Biotechnol. 1999; 9:78–83.

63. Lee JH, Lee KH, Choi SH. Enumeration of *Vibrio vulnificus* in Natural samples by colony blot hybridization. J Microbiol Biotechnol. 2001; 11(2):302–309.

64. Park JH, Cho YJ, Chun J, Seok YJ, Lee JK, Kim KS, et al. Complete genome sequence of *Vibrio vulnificus* MO6-24/O. J Bacteriol. 2011; 193:2062–2063. doi: 10.1128/JB.00110-11 PMID: 21317338

65. Tison DL, Nishibuchi M, Greenwood JD, Seidler RJ. *Vibrio vulnificus* biogroup 2: new biogroup pathogenic for eels. Appl Environ Microbiol. 1982; 44:640–646. PMID: 7138004

66. Bisharat N, Agmon V, Finkelstein R, Raz R, Ben-Dror G, Lerner L, et al. Clinical, epidemiological, and microbiological features of *Vibrio vulnificus* biogroup 3 causing outbreaks of wound infection and bacteraemia in Israel. Lancet. 1999; 354:1421–1424. PMID: 10543668

67. Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: a laboratory manual. 4th ed. New York: Cold spring harbor laboratory press; 2012.

68. Cho JC, Tiedje JM. Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. Appl Environ Microbiol. 2000; 66:5448–5456. PMID: 11097926

69. Häne BG, Jäger K, Drexler HG. The Pearson product-moment correlation coefficient is better suited for identification of DNA fingerprint profiles than band matching algorithms. Electrophoresis. 1993; 14:967–972. PMID: 8125064

70. Rademaker JLW, Louws FJ, De Bruijn FJ. Characterization of the Diversity of Ecologically Important Microbes by rep-PCR Genomic Fingerprinting. In: Akkermans ADL, van Elsas JD, De Bruijn F, editors. Molecular microbial ecology manual, suppl 3. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1998. p. 1–26.

71. Rosche TM, Yano Y, Oliver JD. A rapid and simple PCR analysis indicates there are two subgroups of *Vibrio vulnificus* which correlate with clinical or environmental isolation. Microbiol Immunol. 2005; 49:381–389. PMID: 15840964

72. Panicker G, Vickery MCL, Bej AK. Multiplex PCR detection of clinical and environmental strains of *Vibrio vulnificus* in shellfish. Can J Microbiol. 2004; 50:911–922. PMID: 15644908

73. Kim H-J, Lee H-J, Lee K-H, Cho J-C. Simultaneous detection of Pathogenic *Vibrio* species using multiplex real-time PCR. Food Control. 2012; 23:491–498.

74. Kaysner CA, DePaola A. Compendium of methods for the microbiological examination of food. In: Downes FP, Ito K, editors. Vibrio. Washington, D. C.: American Public Health Association; 2001. p. 405–420.

75. Han F, Pu S, Hou A, Ge B. Characterization of clinical and environmental types of *Vibrio vulnificus* isolates from Louisiana oysters. Foodborne Pathog Dis. 2009; 6:1251–1258. doi: 10.1089/fpd.2009.0343 PMID: 19743928

76. Jolley KA, Maiden MCJ. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC bioinformatics. 2010; 11:595. doi: 10.1186/1471-2105-11-595 PMID: 21143983

77. Gutacker M, Conza N, Benagli C, Pedroli A, Bernasconi MV, Permin L, et al. Population genetics of *Vibrio vulnificus*: identification of two divisions and a distinct eel-pathogenic clone. Appl Environ Microbiol. 2003; 69:3203–3212. PMID: 12788717

78. Kotetishvili M, Stine OC, Chen Y, Kreger A, Sulakvelidze A, Sozhamannan S, et al. Multilocus sequence typing has better discriminatory ability for typing *Vibrio cholerae* than does pulsed-field gel electrophoresis and provides a measure of phylogenetic relatedness. J Clin Microbiol. 2003; 41:2191–2196. PMID: 12734277

79. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994; 22:4673–4680. PMID: 7984417

80. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011; 28:2731–2739. doi: 10.1093/molbev/msr121 PMID: 21546353

81. Jukes Cantor. Mammalian Protein Metabolism. In: Munro HN, editor. Evolution of protein molecules. New York: Academic Press; 1969. p. 21–132.

82. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006; 22:2688–2690. PMID: 16928733

83. Jolley KA, Feil EJ, Chan M-S, Maiden MCJ. Sequence type analysis and recombinational tests (START). Bioinformatics. 2001; 17:1230–1231. PMID: 11751234

84. Nei M, Li WH. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A. 1979; 76:5269–5273. PMID: 291943

85. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989; 123:585–595. PMID: 2513255

86. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25:1451–1452. doi: 10.1093/bioinformatics/btp187 PMID: 19346325

87. Simonsen KL, Churchill GA, Aquadro CF. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics. 1995; 141:413–429. PMID: 8536987

88. Maynard-Smith J, Smith NH, O'Rourke M, Spratt BG. How clonal are bacteria? Proc Natl Acad Sci. 1993; 90:4384–4388. PMID: 8506277

89. Hudson RR. Analytical results concerning linkage disequilibrium in models with genetic transformation and conjugation. J Evol Biol. 1994; 7:535–548.

90. Martin DP, Lemey P, Posada D. Analysing recombination in nucleotide sequences. Mol Ecol Resour. 2011; 11:943–955. doi: 10.1111/j.1755-0998.2011.03026.x PMID: 21592314

91. Bruen TC, Philippe H, Bryant D. A simple and robust statistical test for detecting the presence of recombination. Genetics. 2006; 172:2665–2681. PMID: 16489234

92. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 2006; 23:254–267. PMID: 16221896

93. Chen J-M, Cooper DN, Chuzhanova N, Férec C, Patrinos GP. Gene conversion: mechanisms, evolution and human disease. Nature Rev Genet. 2007; 8:762–775. PMID: 17846636

94. Sawyer S. Statistical tests for detecting gene conversion. Mol Biol Evol. 1989; 6:526–538. PMID: 2677599

95. Nightingale KK, Windham K, Wiedmann M. Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal listeriosis cases and foods. J Bacteriol. 2005; 187:5537–5551. PMID: 16077098

96. den Bakker HC, Didelot X, Fortes ED, Nightingale KK, Wiedmann M. Lineage specific recombination rates and microevolution in *Listeria monocytogenes*. BMC Evol Biol. 2008; 8:277. doi: 10.1186/1471-2148-8-277 PMID: 18842152

97. Bryant D, Moulton V. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. Mol Biol Evol. 2004; 21:255–265. PMID: 14660700

98. Biosca EG, Amaro C, Esteve C, Alcaide E, Garay E. First record of *Vibrio vulnificus* biotype 2 from diseased European eel, *Anguilla anguilla* L. J Fish Dis. 1991; 14:103–109.

99. Zaidenstein R, Sadik C, Lerner L, Valinsky L, Kopelowitz J, Yishai R, et al. Clinical characteristics and molecular subtyping of *Vibrio vulnificus* illnesses, Israel. Emerg Infect Dis. 2008; 14:1875–1882. doi: 10.3201/eid1412.080499 PMID: 19046510

100. Broza YY, Raz N, Lerner L, Danin-Poleg Y, Kashi Y. Genetic diversity of the human pathogen *Vibrio vulnificus*: a new phylogroup. Int J Food Microbiol. 2012; 153:436–443. doi: 10.1016/j.ijfoodmicro.2011.12.011 PMID: 22227412

101. Biosca EG, Amaro C, Larsen JL, Pedersen K. Phenotypic and genotypic characterization of *Vibrio vulnificus*: proposal for the substitution of the subspecific taxon biotype for serovar. Appl Environ Microbiol. 1997; 63:1460–1466. PMID: 9097444

102. Vinuesa P, Rademaker JL, de Bruijn FJ, Werner D. Genotypic characterization of Bradyrhizobium strains nodulating endemic woody legumes of the Canary Islands by PCR-restriction fragment length polymorphism analysis of genes encoding 16S rRNA (16S rDNA) and 16S-23S rDNA intergenic spacers, repetitive extragenic palindromic PCR genomic fingerprinting, and partial 16S rDNA sequencing. Appl Environ Microbiol. 1998; 64:2096–2104. PMID: 9603820

103. Rademaker JL, Hoste B, Louws FJ, Kersters K, Swings J, Vauterin L, et al. Comparison of AFLP and rep-PCR genomic fingerprinting with DNA-DNA homology studies: *Xanthomonas* as a model system. Int J Syst Evol Microbiol. 2000; 50 Pt 2:665–677. PMID: 10758874

104. Chao A. Nonparametric estimation of the number of classes in a population. Scand Stat. 1984:265–270.

105. Colwell RK, Coddington JA. Estimating terrestrial biodiversity through extrapolation. Philos Trans R Soc Lond B Biol Sci. 1994; 345:101–118. PMID: 7972351

106. Pielou ECJ. The measurement of diversity in different types of biological collections. J Theor Biol. 1966; 13:131–144.

107. Han F, Ge B. Multiplex PCR assays for simultaneous detection and characterization of *Vibrio vulnificus* strains. Lett Appl Microbiol. 2010; 51:234–240. doi: 10.1111/j.1472-765X.2010.02887.x PMID: 20586937

108. Vickery MC, Nilsson WB, Strom MS, Nordstrom JL, DePaola A. A real-time PCR assay for the rapid determination of 16S rRNA genotype in *Vibrio vulnificus*. J Microbiol Methods. 2007; 68:376–384. PMID: 17070612

109. Chatzidaki-Livanis M, Jones MK, Wright AC. Genetic variation in the *Vibrio vulnificus* group 1 capsular polysaccharide operon. J Bacteriol. 2006; 188:1987–1998. PMID: 16484211

110. Warner E, Oliver JD. Population structures of two genotypes of *Vibrio vulnificus* in oysters (*Crassostrea virginica*) and seawater. Appl Environ Microbiol. 2008; 74:80–85. PMID: 17993556

111. Bogard RW, Oliver JD. Role of iron in human serum resistance of the clinical and environmental *Vibrio vulnificus* genotypes. Appl Environ Microbiol. 2007; 73:7501–7505. PMID: 17933938

112. Kim MS, Jeong HD. Development of 16S rRNA targeted PCR methods for the detection and differentiation of *Vibrio vulnificus* in marine environments. Aquaculture. 2001; 193:199–211.

113. Cohen AL, Oliver JD, DePaola A, Feil EJ, Boyd EF. Emergence of a virulent clade of *Vibrio vulnificus* and correlation with the presence of a 33-kilobase genomic island. Appl Environ Microbiol. 2007; 73:5553–5565. PMID: 17616611

114. Do T, Jolley KA, Maiden MC, Gilbert SC, Clark D, Wade WG, et al. Population structure of *Streptococcus oralis*. Microbiology. 2009; 155:2593–2602. doi: 10.1099/mic.0.027284-0 PMID: 19423627

115. Bisharat N, Cohen DI, Maiden MC, Crook DW, Peto T, Harding RM. The evolution of genetic structure in the marine pathogen, *Vibrio vulnificus*. Infect Genet Evol. 2007; 7:685–693. PMID: 17716955

116. Feil EJ, Cooper JE, Grundmann H, Robinson DA, Enright MC, Berendt T, et al. How clonal is *Staphylococcus aureus*? J Bacteriol. 2003; 185:3307–3316. PMID: 12754228

117. Boyd EF, Nelson K, Wang FS, Whittam TS, Selander RK. Molecular genetic basis of allelic polymorphism in malate dehydrogenase (*mdh*) in natural populations of *Escherichia coli* and Salmonella enterica. Proc Natl Acad Sci U S A. 1994; 91:1280–1284. PMID: 8108402

118. Keymer DP, Boehm AB. Recombination shapes the structure of an environmental *Vibrio cholerae* population. Appl Environ Microbiol. 2011; 77:537–544. doi: 10.1128/AEM.02062-10 PMID: 21075874

119. Maynard-Smith J, Dowson CG, Spratt BG. Localized sex in bacteria. Nature. 1991; 349:29–31. PMID: 1985260

120. Miller MC, Keymer DP, Avelar A, Boehm AB, Schoolnik GK. Detection and transformation of genome segments that differ within a coastal population of *Vibrio cholerae* strains. Appl Environ Microbiol. 2007; 73:3695–3704. PMID: 17449699

121. Cohan FM. What are bacterial species? Annu Rev Microbiol. 2002; 56:457–487. PMID: 12142474

122. Atwood KC, Schneider LK, Ryan FJ. Periodic selection in *Escherichia coli*. Proc Natl Acad Sci U S A. 1951; 37:146–155. PMID: 14808170

123. Wong H-C, Chen S-Y, Chen M-Y, Oliver JD, Hor L-I, Tsai W-C. Pulsed-field gel electrophoresis analysis of *Vibrio vulnificus* strains isolated from Taiwan and the United States. Appl Environ Microbiol. 2004; 70:5153–5158. PMID: 15345394