Title

Deep-sea hydrothermal vent bacteria related to human pathogenic Vibrio species

Journal

Proceedings of the National Academy of Sciences of the United States of America, 112(21)

ISSN

0027-8424

Authors

Hasan, Nur A
Grim, Christopher J
Lipp, Erin K
et al.

Publication Date

2015-05-26

DOI

10.1073/pnas.1503928112

Peer reviewed

Deep-sea hydrothermal vent bacteria related to human pathogenic *Vibrio* species

Nur A. Hasan[a,b,c], Christopher J. Grim[a,c,1], Erin K. Lipp[d], Irma N. G. Rivera[e], Jongsik Chun[f], Bradd J. Haley[a], Elisa Taviani[a], Seon Young Choi[a,b], Mozammel Hoq[g], A. Christine Munk[h], Thomas S. Brettin[h], David Bruce[h], Jean F. Challacombe[h], J. Chris Detter[h], Cliff S. Han[h], Jonathan A. Eisen[i], Anwar Huq[a,j], and Rita R. Colwell[a,b,c,k,2]

[a]Maryland Pathogen Research Institute, [c]University of Maryland Institute for Advanced Computer Studies, and [j]Institute for Applied Environmental Health, University of Maryland, College Park, MD 20742; [b]CosmosID, College Park, MD 20742; [d]Environmental Health Science, College of Public Health, University of Georgia, Athens, GA 30602; [e]Department of Microbiology, Institute of Biomedical Sciences, University of São Paulo, CEP 05508-900 São Paulo, Brazil; [f]School of Biological Sciences and Institute of Microbiology, Seoul National University, Seoul 151-742, Republic of Korea; [g]Department of Microbiology, University of Dhaka, Dhaka-1000, Bangladesh; [h]Genome Science Group, Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545; [i]University of California Davis Genome Center, Davis, CA 95616; and [k]Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, MD 21205

Contributed by Rita R. Colwell, April 15, 2015 (sent for review September 5, 2014; reviewed by John Allen Baross, Richard E. Lenski, and Carla Pruzzo)

Vibrio **species are both ubiquitous and abundant in marine coastal waters, estuaries, ocean sediment, and aquaculture settings worldwide. We report here the isolation, characterization, and genome sequence of a novel *Vibrio* species, *Vibrio antiquarius*, isolated from a mesophilic bacterial community associated with hydrothermal vents located along the East Pacific Rise, near the southwest coast of Mexico. Genomic and phenotypic analysis revealed *V. antiquarius* is closely related to pathogenic *Vibrio* species, namely *Vibrio alginolyticus*, *Vibrio parahaemolyticus*, *Vibrio harveyi*, and *Vibrio vulnificus*, but sufficiently divergent to warrant a separate species status. The *V. antiquarius* genome encodes genes and operons with ecological functions relevant to the environment conditions of the deep sea and also harbors factors known to be involved in human disease caused by freshwater, coastal, and brackish water vibrios. The presence of virulence factors in this deep-sea *Vibrio* species suggests a far more fundamental role of these factors for their bacterial host. Comparative genomics revealed a variety of genomic events that may have provided an important driving force in *V. antiquarius* evolution, facilitating response to environmental conditions of the deep sea.**

Vibrio | hydrothermal vent | genomics | EX25

With more than 110 recognized species, the genus *Vibrio* comprises a diverse group of heterotrophic bacteria, of which many are known pathogens, causing disease in animals and humans (1, 2). *Vibrio cholerae* is the most notorious because it is the causative agent of cholera. *Vibrio vulnificus* and *Vibrio parahaemolyticus* cause severe illness in humans and are associated with consumption of contaminated seafood (3, 4). *Vibrio harveyi* (5), *Vibrio anguillarum* (6, 7), and *V. parahaemolyticus* (8) continue to cause substantial economic losses to the aquaculture industry worldwide.

Vibrios demonstrate a wide range of niche specialization: for example, free-living, attached to biotic and abiotic surfaces, and resident in both estuarine and marine habitats (9). The deep sea constitutes the largest habitat of the biosphere that supports microbial communities across three domains of life and represents an environment where physiochemical parameters—such as low temperature, high salinity, and high pressure—modulate community structure (10, 11). Several studies have shown the presence of physiologically, metabolically, and phylogenetically diverse mesophilic microbial communities in the deep sea, including *Vibrio* species (12–15). Barotolerant *Vibrio* spp. have been isolated from deep-sea sediment and from the gut microflora of invertebrates and fish collected from a variety of deep-sea habitats, including hydrothermal vents (16, 17). For example, strains of *Vibrio*, *Aeromonas*, and *Pseudomonas* spp. were isolated from mud-water samples collected at a depth of 4,940 m, 150 miles east of Cape Canaveral, Florida (18). Several culture-dependent and -independent studies have confirmed the ubiquity of vibrios, and suggested *Vibrio* populations generally comprise approximately 1% (by molecular techniques) of the total bacterioplankton in estuaries (19), in contrast to culture-based studies demonstrating that vibrios can comprise up to 10% of culturable marine bacteria (20). Clearly, vibrios are ubiquitous and abundant in the aquatic environment on a global scale, including both seawater and sediment (19, 21–25), and repeatedly shown to be present in high densities in and on marine organisms, such as corals (26), fish (27–29), mollusks (30), seagrass, sponges, shrimp (28, 31), and zooplankton (16, 17, 28, 32, 33).

During dives of the deep-sea submersibles *Alvin* and *Nautile* in 1999 along the East Pacific Rise, southwest of the Mexico coast, samples of water surrounding sulfide chimneys of a hydrothermal vent community were collected and four mesophilic bacterial isolates were cultured, which were subsequently tested for phenotypic traits, including growth on *V. cholerae* selective thiosulfate-citrate-bile-salts-sucrose (TCBS) agar (Oxoid). The sampling locations from where these four mesophilic bacteria were isolated are described in Table 1. Using single-gene phylogenies, the four isolates were identified as *Shewanella algae*, *V. harveyi*, and two novel *Vibrio* species, designated *Vibrio* sp. EX25 and *Vibrio* sp. EX97. Among them, *Vibrio* sp. EX25 showed phenotypic

Significance

During *Alvin* and *Nautile* dives in 1999, samples were collected from water surrounding sulfide chimneys of a hydrothermal vent along the East Pacific Rise and four mesophilic bacteria were isolated, including a novel *Vibrio* species, *Vibrio antiquarius*. Genomic, functional, and phylogenetic analyses indicate an intriguing blend of genomic features related to adaptation and animal symbiotic association, and also revealed the presence of virulence genes commonly found in *Vibrio* species pathogenic for humans. The presence of these virulence genes in an ecologically distinct *Vibrio* species was surprising. It is concluded that pathogenicity genes serve a far more fundamental ecological role than solely causation of human disease.

Author contributions: N.A.H., A.H., and R.R.C. designed research; N.A.H., C.J.G., E.K.L., I.N.G.R., M.H., A.H., and R.R.C. performed research; J.C., A.H., and R.R.C. contributed new reagents/analytic tools; N.A.H., C.J.G., J.C., B.J.H., E.T., S.Y.C., A.C.M., T.S.B., D.B., J.F.C., J.C.D., C.S.H., J.A.E., and R.R.C. analyzed data; and N.A.H., J.A.E., A.H., and R.R.C. wrote the paper.

Reviewers: J.A.B., University of Washington; R.E.L., Michigan State University; and C.P., University of Genoa.

The authors declare no conflict of interest.

[1]Present address: Center for Food Safety and Applied Nutrition, US Food and Drug Administration, Laurel, MD 20708.

[2]To whom correspondence should be addressed. Email: rcolwell@umiacs.umd.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1503928112/-/DCSupplemental.

Table 1. Locations where samples were collected and from which four mesophilic bacterial isolates were obtained

Sample	Location	Depth (m)	Source	Strain no.
1	9°N	2,520	Sulfide chimney	EX25
2	9°N	2,500	Sulfide chimney	EX97
3	13°N	2,596	Sulfide chimney	BB4
4	13°N	2,602	Sulfide chimney	A6.mk

and genetic similarity to *Vibrio alginolyticus*, *V. parahaemolyticus*, and *V. cholerae*, all of which are human pathogens. Because it appeared to be a new *Vibrio* species derived from a novel habitat and closely related to human pathogenic *Vibrio* spp., we sequenced the whole genome of EX25 to understand its evolutionary lineage and determine its gene content, specifically those genes associated with pathogenicity in humans.

Results and Discussion

Phylogenetic analysis based on 16S rRNA showed that three of the four isolates belonged to the genus *Vibrio* (BB4, EX25, and EX97) and the fourth isolate, A6, was identified as *S. algae*. *Vibrio* BB4 branched with *V. harveyi* (Fig. 1). Although isolate EX97 was clustered with *V. parahaemolyticus,* EX25 branched independently of the *V. parahaemolyticus*, and *V. alginolyticus*–*Vibrio campbellii* clade. Both isolates were sucrose-positive, oxidase-positive, and required NaCl for growth. The isolates grew well under both micro aerophilic [Gas Pak anaerobic system (H_2 and CO_2); Beckton-Dickson] and strict anaerobic conditions (80% N_2, 10% CO_2, and 10% H_2). In 1% tryptone amended with NaCl, EX25 grew at NaCl concentrations up to 10% and in alkaline peptone water visible growth also occurred at temperatures as high as 50 °C. Because protein-encoding genes evolve faster than rRNA genes, additional trees were constructed based on the protein-encoding genes *tox*R (cholera toxin transcriptional activator), *omp*W (outer membrane protein W), and other highly conserved genes among *Vibrio* spp. to yield better resolution in phylogenetic analyses. The constructed phylogenetic trees were in good agreement with 16S tree. DNA-DNA hybridization was also conducted to understand the genomic relatedness of the two isolates to 41 type strains of Vibrionaceae, and the highest relative branching ratio of 42% (Fig. S1) was obtained when EX25 was hybridized with type strains of *V. parahaemolyticus* and *V. alginolyticus*.

Taxonomy of *Vibrio* EX25 Based on Average Nucleotide Identity. Species delineation was determined by average nucleotide identity (ANI) between genomes (34), with highest ANI observed between *Vibrio* EX25 and *V. alginolyticus* 12G01 (91%) and *V. parahaemolyticus* RIMD 2210633(84%), both of which is below the hypothesized species demarcation threshold value of 95% (35, 36) (Fig. S2*A*), indicating EX25 as a separate species closely related to *V. alginolyticus* and *V. parahaemolyticus.* A tetra nucleotide signature correlation index, an alignment-free parameter helpful in deciding whether a given pair of organisms belongs to the same species (36), demonstrated EX25 to have highest correlation with *V. alginolyticus*, yet below the same species threshold (Fig. S2*B*). These findings, along with16S rDNA and DNA-DNA hybridization results, confirm separate species designation for *Vibrio* sp. EX25, for which the species name *Vibrio antiquarius* is proposed (i.e., a bacterium of the aquatic environment derived from the antiquity of the deep sea).

Whole Genome Phylogeny. The phylogeny of *V. antiquarius* was inferred by constructing a genome-relatedness neighbor-joining tree, using homologous alignment of 522 orthologous protein-coding genes of 36 *Vibrio* genomes, as a strict measure of the core *Vibrio* genome. The evolutionary tree (Fig. 2) showed fully resolved bifurcating patterns, with varying levels of diversity as evidenced by tree branch lengths, placing *V. antiquarius*, *V. alginolyticus*, *V. parahaemolyticus*, *Vibrio* sp. AND4, and *V. harveyi* in a monophyletic clade. *V. antiquarius* EX25 branched with *V. alginolyticus*, with *V. parahaemolyticus* as an outgroup to both of them. Strains of *V. cholerae* and *V. vulnificus* were each monophyletic within the species. This finding corroborated findings by shared gene content, ANI, and 16S phylogeny, and strongly indicates that at least for the core of these genomes, they share a common ancestry that excludes *V. cholerae*.

Genome Features. A combination of Sanger sequencing and 454 pyrosequencing yielded high-quality assembly of the EX25genome, with an asymmetrical, two-chromosome structure observed, consistent for *Vibrio* genomes (37). The larger (C-I) and smaller (C-II) chromosomes comprised 3.26 and 1.83 Mb, respectively (Fig. S3 and Table S1), with 45% G+C composition. Chromosomal distribution of the genes followed a pattern typical of vibrios, with C-I predominantly carrying genes for viability and growth, and C-II containing genes associated with adaptation to environmental change. A total of 4,529 protein-coding sequences (CDS) were identified (2,846 at C-I and 1,683 at C-II) with 27.5% of CDSs annotated as hypothetical proteins. Additionally, the *V. antiquarius* EX25 genome contained a superintegron (SI) cassette spanning approximately 113 kb on C-I (Table S1).

Comparative Genomics. Multigenome comparison (Fig. 3) was carried out using reannotated eight reference *Vibrio* genomes (38). Pairwise reciprocal BLAST analysis revealed EX25 shared higher predicted CDSs with *V. parahaemolyticus* (3,973, 87.7%), *V. alginolyticus* (3,943, 87%), *V. harveyi* (3,567, 79%), and *V. vulnificus* (3,309, 73%) (Fig. S4). CDSs shared with *V. parahaemolyticus* (3,973) and *V. vulnificus* (3,309) corresponded to 82% and 84.5% of the *V. parahaemolyticus* and *V. alginolyticus* genomes. Analyses of shared gene content indicated *V. antiquarius*, *V. alginolyticus*, and *V. parahaemolyticus* are about equidistant (Fig. S4). Additionally, a large number of genomic regions (~23 on C-I and ~17 on C-II) were found with significant mismatch (Fig. 2) compared with other reference *Vibrio* genomes. These mismatches occurred most likely because of the insertion of genomic islands and acquisition of mobile genetic elements in those regions, resulting in strain specific CDSs. Interestingly, C-II

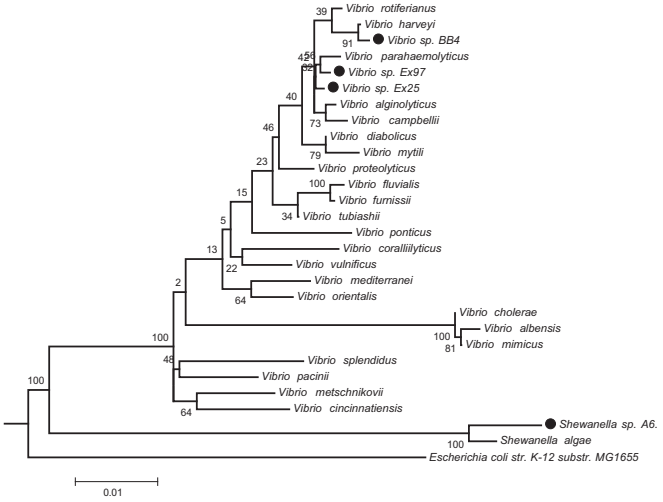


Fig. 1. 16S rRNA phylogeny of the East Pacific Rise isolates and other *Vibrio* species by neighbor-joining tree. Bootstrap consensus tree was inferred from 5,000 replicates, representing evolutionary history of the taxa. Scale bar represents 0.01 nucleotide substitutions per sequence position.

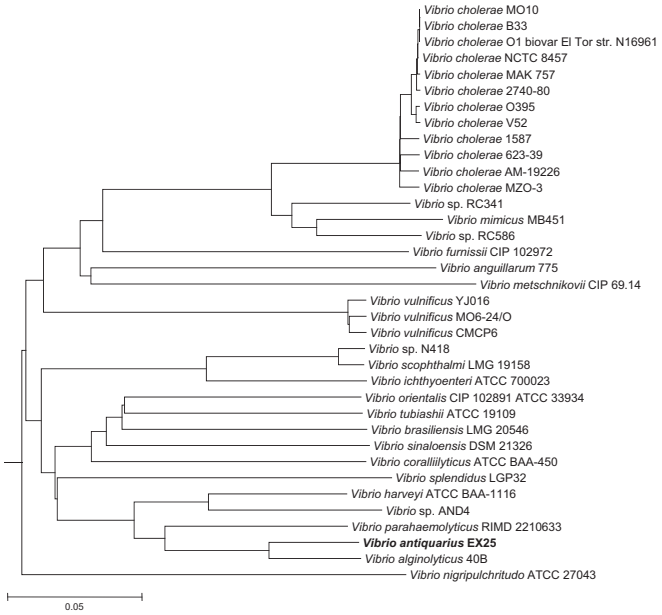


Fig. 2. Core genome phylogeny of *V. antiquarius* EX25. Neighbor-joining tree was constructed based on alignment of homologous sequences of 525 conserved ORFs. Scale bar represents 0.05 substitutions per site.

contained more strain-specific CDSs than C-I. These chromosomal regions contributed significantly to the 586 CDSs (13% of the *V. antiquarius* genome), which had no reciprocal match in the *V. alginolyticus* genome. Approximately 43% (256 CDSs) of these CDSs represent either hypothetical proteins or proteins of unknown function. The presence of an integrase gene, together with a G+C content atypical to the chromosomal G+C, and at least seven major regions of disagreement, including pre-CTX prophage, SI, *Vibrio* pathogenicity island 2 (VPI-2), ORF: 4331–4339, ORF: 4301–4312, ORF: 3898–3902, and ORF: 1829–1846, suggests these regions are subject to horizontal gene transfer or chromosomal integration via phage, but might have a necessary function in the deep-sea habitat. The *V. alginolyticus* genome contained 723 CDSs (15% of its genome) without any reciprocal match with *V. antiquarius*.

Genome Plasticity. Genomic islands (GI's), notably pathogenicity islands, contribute to the evolution and diversification of microbial life. The *V. antiquarius* genome encoded >70 genomic islands as predicted by Island Viewer (39) (Fig. S5 and Table S2). Among the GIs predicted by multiple methods, 21 were located on C-I and 12 were on C-II. Average size of the islands in C-I was 14 kb, with G+C content ranging from 36 to 40%, which is lower than the overall chromosomal G+C of 45%. The 12 GIs on C-II had a G+C content ranging from 38 to 42%, with 10.7-kb average size. Part of the GI-I (7%) displayed homology (74%) with the O-antigen gene cluster of *Escherichia coli* serotype O98: K?:H8 and *Shigella dysenteriae* strain M13547, whereas many of GI-II–encoded proteins are related to capsular polysaccharide biosynthesis and exhibit ~68% sequence similarity with the *E. coli* capsule transport proteins. A total of 222 strain-specific CDSs identified in the genome of *V. antiquarius* (4.9% of the total CDSs) had no reciprocal match in the genomes of *V. alginolyticus*, *V. parahaemolyticus*, *V. vulnificus*, *V. harveyi*, or *V. cholerae*. Among *V. antiquaries*-specific protein-coding sequences, 116 (52.25%) are hypothetical or proteins of unknown function. Linear pairwise comparison of the *V. antiquarius* EX25 genome demonstrated several intra- and interchromosomal rearrangements compared with *V. parahaemolyticus*, *V. harveyi*, *V. cholerae*, and *V. vulnificus* (Fig. S6). Compared with *V. parahaemolyticus* and *V. harveyi*, C-I of *V. antiquarius* demonstrated a high degree of synteny, compared with C-II (Fig. S7). Such rearrangements are in agreement with the supposition that extensive genome plasticity is common in *Vibrio* species, particularly on C-II (40). Additionally, the *V. antiquarius* genome contains many perfect and approximate tandem repeats. Using the tandem-repeats finder (41), 64 and 20 tandem repeats were identified in the EX25 genome, with period lengths of 6–417 and 6–429 bases in C-I and C-II, respectively (Table S3). Many tandem repeats were in protein-encoding genes exhibiting high mutation rates. The *V. antiquarius* genome also contained insertion sequences (IS elements) throughout its genome (Table S4). Therefore, these might be important processes in *V. antiquarius* evolution to facilitate faster adaptation or quicker response to rapidly changing and challenging environmental conditions of the deep sea.

Predicted Biology of *V. antiquarius*. The genome of *V. antiquarius* EX25 encodes a number of genes that are predicted to protect the bacterium from the environmental conditions of the deep sea and are illustrated in Table 2. Functions encoded in the genome include cytochromes for reduction of O_2 to H_2O_2 and cytochrome C_{551} peroxidase, which detoxifies peroxide, multiple catalase genes, and a superoxide dismutase for tolerating high O_2 concentrations, and genes encoding alkyl hydroperoxide reductase to scavenge endogenous hydrogen peroxide (Table 2). The ability to scavenge endogenous hydrogen peroxide was absent in the other *Vibrio* genomes and is the major antioxidant enzyme of the endosymbiont of the tubeworm, *Riftia pachyptila*, inhabiting deep-sea vents (42). Metalloendopeptidases and zinc-dependent carboxypeptidases were also present in the genome of *V. antiquarius*, useful in functioning in high concentrations of heavy metals, including zinc, present in the deep-sea vent environment (43).

Manganese is used as a reliable tracer of hydrothermal vent emissions (44–46) and, unlike the other vibrios, the genome of *V. antiquarius* contains a gene annotated as multicopper oxidase, an enzyme essential for manganese oxidation, and laccase-like activity (47). *V. antiquarius* also contains delta-9 fatty acid

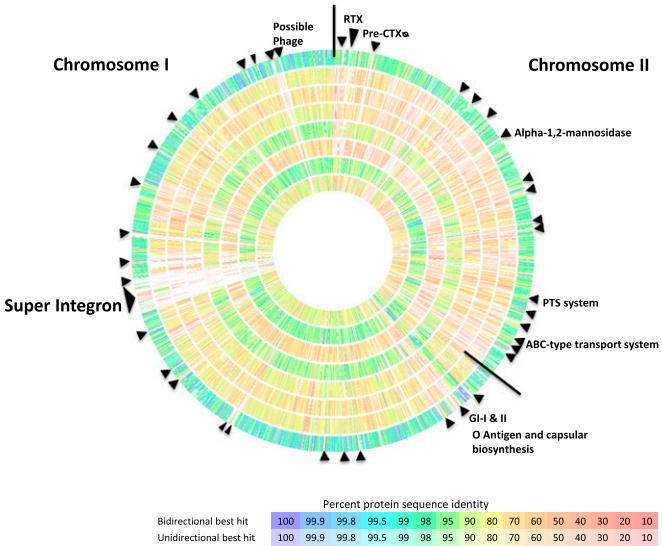


Fig. 3. Genome comparison of *V. antiquarius* EX25 with other *Vibrio* genomes. From outer ring to inner ring; *V. alginolyticus* 12G01, *V. cholerae* N16961, *V. fischeri* MJ11, *Vibrio furnissii* CIP 102972, *V. harveyi* ATCC BAA-1116, *V. hollisae* CIP 102972, *V. parahaemolyticus* RIMD 2210633, *V. vulnificus* CMCP6. Horizontal lines were drawn to separate chromosomes. Solid black arrows indicate areas of rearrangement or insertion site of genomic islands.

Table 2. Predicted characteristics of *V. antiquarius* EX25 genome

Predicted biology	Chromosome (occurrence)	Function
Response to environment		
Cytochrome *bd*	C-II (2)	Protection against O_2 and H_2O_2
Cytochrome C_{551} peroxidase	C-II	
Catalase	C-II (2)	Tolerating high oxygen concentrations
Superoxide dismutase	C-I (2)	
Alkyl hydroperoxide reductase	C-II	Scavenge endogenous hydrogen peroxide, a trait absent in the other Vibrio genomes
Methionine sulfoxide reductases	C-I	Oxidative damage repair
Metalloendopeptidases	C-II (2)	Functioning in high concentrations of heavy metals, including zinc
Zinc-dependent carboxypeptidases	C-II	
Cell sensing system		
LuxP and LuxQ	C-I	Autoinducer-2 (AI-2) mediated quorum sensing, biofilm formation, virulence, and other metabolic functions
LuxS and LuxN	C-II	
Biofilm-related pathways		
Ornithine and arginine decarboxylase	C-I	Polyamine biosynthesis
C-di-GMP phosphodiesterase mbaA	C-I	Norspermidine
syp gene cluster	C-I	Gene clusters mediating biofilm formation
Polar flagellum cluster, P	C-I	Flagellar clusters
Lateral flagellum cluster, LF	C-II	
Others		
rmf and *sulA*	C-I, C-II	Persister cells
Multicopper oxidase	C-I	Mn(II) oxidation
δ-9 fatty acid desaturase	C-II	Fatty acid unsaturation, essential for growth under high pressure
Universal stress proteins	C-II (3)	Function not defined

desaturase. Fatty acid unsaturation is a critical cellular process shown to be essential for growth under high pressure by increasing the rigidity of membranes and genes like delta-9 fatty acid desaturase are presumably up-regulated to increase membrane unsaturation and fluidity (48).

Virulence Factors. The genome of *V. antiquarius* encodes a type III secretion system (T3SS), responsible for enabling injection of effector proteins directly into target host eukaryotic cells (49). T3SS genes induce severe diarrhea in models of cholera infection (50) and are frequently found in *V. parahaemolyticus* (51). Unlike *V. parahaemolyticus*, *V. alginolyticus*, *V. harveyi*, and *V. vulnificus*, the genome of *V. antiquarius* also contains two clusters of type VI secretion systems (T6SS) genes on C-II. However, the two effector molecules, VgrG and Hcp, and regulatory proteins, vasK and vasH, are located in different clusters. To date, T6SSs have been defined as required for virulence or survival of a bacterium in a eukaryotic host (52–54). However, Weber et al. (55) reported T6SS in *V. anguillarum* regulates stress response, suggesting T6SS has an ecological rather than pathological consequence for its host (i.e., an adaptive mechanism). The *V. antiquarius* genome also contains tight adherence (*tad*) locus genes (*rcp*, *rcpA*, and *tadZ*), functioning in colonization of surfaces and biofilm formation (56).

V. antiquarius encodes the thermo labile hemolysin (*tlh*) gene in C-II, with 97% and 84% nucleotide similarity to that of *V. alginolyticus* and *V. parahaemolyticus*, respectively, but lacks *tdh* and *trh* of *V. parahaemolyticus*. Multiple putative proteases and other genes, whose products are predicted to encode hemolysins, are present in the genome of *V. antiquarius* EX25, along with homologs of ToxR, ToxS, and Integron Integrase *Intl*4, whose role in virulence is well established in *V. cholerae* and other *Vibrio* species. Other genes involved in pathogenicity are also present, including type IV pilin, *pil*A, which encodes proteins expressed during human infection, mannose-sensitive hemagglutinin, and RTX toxin. The genome contains a homolog of *vvhA* and, in a recent study, Smith and Oliver (57) suggested a role for haemolysin (*vvhA*) in *V. vulnificus* to aid in osmoregulation and cold-shock response. The pre-CTX prophage of *V. cholerae* is indicated by the presence of the accessory cholera enterotoxin, *ace*, Zonula occludence toxin, *zot*, RstA phage-related protein, and RstB phage-related integrase genes on C-II. An insertion of bacteriophage genes was noted between *rstA* and *ace*, with 96% homology to bacteriophage BfO4K68 (Fig. 4*A*). This region is most likely horizontally transferrable via bacteriophage BfO4K68. Although the genome of *V. antiquarius* does not contain the VPI-encoding receptor for CTX prophage, it does have an approximately 27.4-kb contiguous region on C-I (Fig. 4*B*), with 86% nucleotide sequence similarity to VPI-2, found in both clinical and environmental strains of *V. cholerae*, a region spanning VC1758 to VC1772 of the canonical VPI-2 and encoding a type I restriction-modification system and five

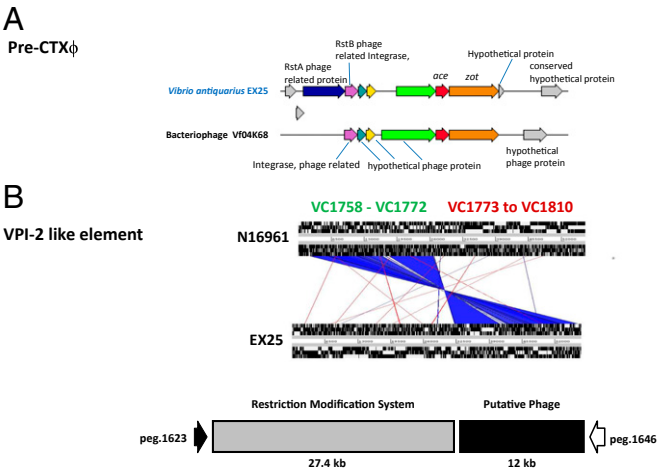


Fig. 4. Schematic representation of (*A*) Pre-CTXphi and (*B*) VPI-2 like element, identified in the genome of *Vibrio antiquarius* EX25.

hypothetical proteins. However, sections VC1773 to VC1810 were not found; instead there is a 12-kb region that includes six hypothetical proteins and a phage integrase. The 12-kb insert did not show significant match with any sequences in the National Center for Biotechnology Information GenBank database, except for a 103-bp region with 78% sequence similarity to a transcriptional regulator of *V. parahaemolyticus*. Recent analyses of the VC1773 to VC1810 region of this island suggest this is a hot-spot for novel DNA insertion (58). The presence of virulence factors, including two reported in *V. cholerae*, namely pre-CTXΦ and VPI-2 in the noncholera *Vibrio antiquarius* from a deep-sea environment, suggests their multifaceted role outside the human host; that is, ecological function in the natural habitat and alternate evolutionary origins apart from their core genome.

Metagenomic Survey. To determine whether *V. antiquarius* may be present in other environmental habitats, *V. antiquarius* ORFs were queried against publicly available environmental metagenomic datasets. Using a conventional Blast search, *V. antiquarius* EX25-specific sequences were detected in 89 shotgun metagenomic datasets and comprise saltern (60%), marine (20%), coral (3%), and human gut metagenomes (2%) (Fig. S8). Distribution of *V. antiquarius* ORFs in these metagenomes suggests ubiquity of *V. antiquarius* in the natural environment.

Summary

Because vibrios are autochthonous to a diverse and wide range of aquatic niches, it was of interest to investigate their potential presence in the deep-sea hydrothermal vents, the only deep-sea environment having high enough temperature to be supportive of mesophilic growth. Analysis of the samples collected from hydrothermal vents did reveal the presence of vibrios in this environment, and interestingly, isolates demonstrated some phenotypic and genotypic similarity to *Vibrio* species pathogenic for humans. Whole-genome sequencing and subsequent comparative and phylogenomics analysis of one of the deep-sea *Vibrio* isolates, EX25, revealed that it belongs to a new *Vibrio* species, for which we propose the name, *Vibrio antiquarius*. The genome of *V. antiquarius* encodes many genes that can be interpreted as contributing to its being native to the deep-sea environment, including genes (i.e., metalloendopeptidases, zinc-dependent carboxypeptidases, and so forth) that are indicative of its potential association with deep-sea animals. Several studies (12–15) in the past have shown the presence of diverse mesophilic microbial communities, including *Vibrio* species, in various deep-sea environments; however, information on the microbial communities associated with the deep-sea vent animals appeared to be very limited. Our study indicates that mesophilic vibrios are likely to be present in the mesophilic environment of the deep sea, particularly in association with the inhabiting animals. Additionally, the genome contained homologs of many virulence genes that are commonly found in *Vibrio* species pathogenic to human and other animals. Wide distribution of virulence genes among coastal, estuarine, and riverine *Vibrio* species, including *V. parahaemolyticus*, *V. cholerae* non-O1, *Vibrio mimicus*, *Vibrio hollisae*, *Vibrio fluvialis*, and *V. alginolyticus* are well known (59–61); however, finding these virulence genes in a deep-sea *Vibrio* sp. raises a significant question whether pathogenicity genes, in addition to pathogenicity for humans and other animals, are in fact providing ecological functions in the natural environment. Recent studies have shown that several virulence factors and pathways in *Vibrio* species that have a role in pathogenicity for humans may also have roles in the aquatic environment (62), and some of the virulence genes might be relevant for basic metabolic processes, establishing the symbiosis (63), or modulating prey/predator relationships (64) in their natural ecosystems. For example, the GbpA ligands in *V. cholerae*, which are involved in the intestinal colonization, have also been reported to mediate bacterial attachment to the chitinous surfaces and biofilm formation in the aquatic environment (65, 66). The tracheal cytotoxin produced by *Vibrio fischeri* has also been reported to be involved in the symbiotic relationship of *V. fischeri* with bobtail squid (63). Similarly, the metalloendopeptidases and zinc-dependent carboxypeptidases in *V. antiquarius* genome also encode genes (i.e., metalloendopeptidases, and zinc-dependent carboxypeptidases, and so forth) that could be indicative of its potential association with deep-sea animals. Therefore, it is possible that some of these pathogenicity genes function in the commensal relationship that *V. cholerae* and other vibrios have with zooplankton, notably copepods (67), encoding attachment, signaling, and interactions in aquatic communities, including the deep sea, and therefore primarily may play an ecological role in the natural environment. The presence of these genes in ecologically and phylogenetically diverse *Vibrio* species also suggests that these affiliations between commensals are likely very old and indicate a likely common evolution of *Vibrio* species into pathogens of humans and marine animals. Clearly, a new perspective is needed for understanding the intersecting roles of *Vibrios* in the environment and as a pathogen for humans and marine animals.

Materials and Methods

Sample Collection and Isolation of Cultures. Samples were collected from two sites on the East Pacific Rise, 9°50′N, 104°17′W by Deep Submergence Vehicle *Alvin*, and 13°N (12°49′N, 103°56′W) at a depth of 2,500 m by Deep Submergence Vehicle *Nautile*. Further description of sampling procedures and isolation sites has been described elsewhere (68). Samples were inoculated into heterotrophic, anaerobic seawater-based media immediately after sample retrieval onboard the mother ship. Samples collected from 9°N were inoculated into MSH medium (69), with 4 mM FeS, 4 mM H_2S, 0.5 g/L yeast extract, 0.5 g/L trypticase peptone, with and without 5 mM acetate. Samples from 13°N were inoculated into MSH medium amended with 0.05 g/L of yeast extract, 0.05 g/L trypticase peptone, 20 mM acetate, 0.3 g/L coenzyme M, and 15 mM of iron pyrophosphate. All samples were incubated at 25 °C (9°N) or 30 °C (13°N). Four pure cultures, obtained by serial dilution and colonies picked from roll tubes, were designated A6 and BB4 (13°N), and EX25 and EX97 (9°N). The cultures were enriched aerobically in alkaline peptone water (APW) and Luria-Bertani broth (LB) and spread on TCBS. Yellow colonies (sucrose positive) on TCBS were streaked on the same agar medium (APW or LB) used for selection. The pure cultures were subjected to biochemical tests using API20E strips (Biomerieux Vitek). Salt tolerance was assayed in nutrient broth containing 3%, 6%, and 8% (wt/vol) NaCl.

DNA–DNA Hybridization. Relatedness of the isolates to reference strains of *Vibrio* spp. was also determined by total DNA–DNA hybridization, using random-primer labeling and chemiluminescent detection with DIG High Prime DNA Labeling and Detection Starter Kit II (Roche). Ten replicates of each strain (10 ng) were hybridized with genomic DNA from 41 reference-type strains of Vibrionaceae species. Four additional strains of *V. cholerae* and two of *V. mimicus* were included as reference strains and their relative binding ratios calculated. The reference strains were also probed against one another, serving as controls.

Immunological identification, using anti–DIG-AP and CSPD (chemiluminescent substrate) was performed, according to manufacturer's protocol (Roche). Blots were exposed to X-ray film for 20 min to 4 h, depending on the signal obtained. Developed film was scanned using a densitometer (Personal Densitometer SI) and dots quantitatively evaluated using ImageQuaNT software for Windows NT (Molecular Dynamics, v4.2a).

Genome Sequencing. The genome of *Vibrio* sp. EX25 was sequenced by the Joint Genome Institute, and all general aspects of sequencing performed at the Joint Genome Institute can be found at jgi.doe.gov/. Draft sequences were obtained from a blend of Sanger and 454 sequences and involved paired-end Sanger sequencing on 8-kb plasmid libraries to 5× coverage, with 20× coverage of 454 data accomplished and optional paired end Sanger sequencing on 35-kb fosmid libraries to 1–2× coverage (depending on repeat complexity). ThePhred/Phrap/Consed software package (www.phrap.com) was used for sequence assembly and quality assessment (70, 71). After the shotgun

stage, reads were assembled with parallel phrap (High Performance Software). Draft assemblies were based on 39,974 total reads. Repeat resolution was performed using Dupfinisher (72). Gaps between contigs were closed by editing in Consed and several targeted finishing reactions, including transposon bombs (73), primer walks on clones, primer walks on PCR products, and adapter PCR reactions. Gene-finding and annotation were achieved using the RAST server (38). The completed genome sequences of *Vibrio* EX25 contained 42,569 reads, achieving average eightfold sequence coverage, with error rate of less than 1 in 100,000. 16S rRNA GenBank accession numbers for *Vibrio* sp. BB4, EX25, EX97, and *S. algae* A6 are AF 319768; AF319769, AF319770, and AF 319767, respectively.

Comparative Genomics. Genome-to-genome comparison was performed using three approaches. First, nucleotide sequences as whole contigs were directly aligned using the MUMmer (74). Second, ORFs of a given pair of genomes were reciprocally compared with each other, using BLASTN, BLASTP, and TBLASTX (ORF-dependent comparison). Third, a bioinformatic pipeline was developed to identify homologous regions of a given query ORF. Initially, a segment on a target contig homologous to a query ORF was identified using BLASTN. This potentially homologous region was expanded in both directions by 2,000 bp, after which nucleotide sequences of the query ORF and selected target homologous region were aligned, using a pairwise global alignment algorithm (75), and the resultant matched region in the subject contig was extracted and saved as a homolog (ORF-independent comparison). Orthologs and paralogs were differentiated by reciprocal comparison. In most cases, both ORF-dependent and independent comparisons yielded the same orthologs, although the ORF-independent method performed better for draft sequences of low quality where sequencing errors, albeit rare, hampered identification of correct ORFs. Orthologous regions were used to generate phylogenetic trees. The set of orthologous regions for each CDS were aligned using CLUSTALW2. The resultant multiple alignments were concatenated to form genome scale alignments which were then used to generate the neighbor-joining (76) phylogenetic tree.

ACKNOWLEDGMENTS. We thank Dr. A. L. Reyesenbach and Dr. Krista Longnecker for advice and assistance in collecting isolates and 16S RNA analysis. This study was supported by National Institutes of Health Grant 1R01A139129-01 and National Oceanic and Atmospheric Administration, Oceans and Human Health Initiative Grant S0660009 (to R.R.C.). Funding for genome sequencing was provided by the Office of the Chief Scientist and National Institute of Allergy and Infectious Diseases Microbial Sequencing Centers (N01-AI-30001 and N01-AI-40001).

1. Farmer JJ, Janda JM, Brenner FW, Cameron DN, Birkhead KM (2005) Genus I. *Vibrio pacini* 1854. *Bergey's Manual of Systematic Bacteriology*, eds Brenner DJ, Kreig NR, Staley JT (Springer Science Business Media, New York), 2 Ed, Vol 2, pp 494–546.
2. Pruzzo C, Huq A, Colwell RR, Donelli G (2005) Pathogenic *Vibrio* species in the marine and estuarine environment. *Ocean and Health Pathogens in the Marine Environment*, eds Belkin S, Colwell RR (Springer, New York), pp 217–252.
3. Hülsmann A, et al. (2003) RpoS-dependent stress response and exoenzyme production in *Vibrio vulnificus*. *Appl Environ Microbiol* 69(10):6114–6120.
4. Wong HC, Wang P, Chen SY, Chiu SW (2004) Resuscitation of viable but non-culturable *Vibrio parahaemolyticus* in a minimum salt medium. *FEMS Microbiol Lett* 233(2): 269–275.
5. Owens L, Busico-Salcedo N (2006) *Vibrio harveyi*: Pretty problems in paradise. *The Biology of the Vibrios*, eds Thompson FL, Austin B, Swings J (ASM Press, Washington, DC), pp 266–280.
6. Crosa JH, Actis LA, Tolmasky ME (2006) The biology and pathogenicity of *Vibrio anguillarum* and *Vibrio ordalii*. *The Biology of the Vibrios*, eds Thompson FL, Austin B, Swings J (ASM Press, Washington, DC), pp 251–265.
7. Miyamoto N, Eguchi M (1997) Response to low osmotic stress in a fish pathogen, *Vibrio anguillarum*. *FEMS Microbiol Lett* 22(3):225–231.
8. Austin B, Zhang XH (2006) *Vibrio harveyi*: A significant pathogen of marine vertebrates and invertebrates. *Lett Appl Microbiol* 43(2):119–124.
9. Reen FJ, Almagro-Moreno S, Ussery D, Boyd EF (2006) The genomic code: Inferring Vibrionaceae niche specialization. *Nat Rev Microbiol* 4(9):697–704.
10. Jannasch HW (1995) Microbial interaction with hydrothermal fluids. *Sea Floor Hydrothermal Systems: Physical, Chemical, Biological, and Geological Interactions*, eds Humphris SE, Zierenberg RA, Mullineaux LS, Thomson RE (American Geophysical Union, Washington, DC), pp 273–296.
11. Karl DM, Wirsen CO, Jannasch HW (1980) Deep-sea primary production at the galapagos hydrothermal vents. *Science* 207(4437):1345–1347.
12. Moyer CL, Dobbs FC, Karl DM (1995) Phylogenetic diversity of the bacterial community from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii. *Appl Environ Microbiol* 61(4):1555–1562.
13. Raguénès G, Christen R, Guezennec J, Pignet P, Barbier G (1997) *Vibrio diabolicus* sp. nov., a new polysaccharide-secreting organism isolated from a deep-sea hydrothermal vent polychaete annelid, *Alvinella pompejana*. *Int J Syst Bacteriol* 47(4):989–995.
14. Reysenbach AL, Longnecker K, Kirshtein J (2000) Novel bacterial and archaeal lineages from an in situ growth chamber deployed at a Mid-Atlantic Ridge hydrothermal vent. *Appl Environ Microbiol* 66(9):3798–3806.
15. Sievert SM, Kuever J, Muyzer G (2000) Identification of 16S ribosomal DNA-defined bacterial populations at a shallow submarine hydrothermal vent near Milos Island (Greece). *Appl Environ Microbiol* 66(7):3102–3109.
16. Ohwada K, Tabor PS, Colwell RR (1980) Species composition and barotolerance of gut microflora of deep-sea benthic macrofauna collected at various depths in the atlantic ocean. *Appl Environ Microbiol* 40(4):746–755.
17. Tabor PS, Ohwada K, Colwell RR (1981) Filterable marine bacteria found in the deep sea: Distribution, taxonomy, and response to starvation. *Microb Ecol* 7(1):67–83.
18. Schwarz JR, Walder JD, Colwell RR (1974) Deep-sea bacteria: Growth and utilization of hydrocarbons at ambient and in situ pressure. *Appl Microbiol* 28(6):982–986.
19. Heidelberg JF, Heidelberg KB, Colwell RR (2002) Bacteria of the gamma-subclass Proteobacteria associated with zooplankton in Chesapeake Bay. *Appl Environ Microbiol* 68(11):5498–5507.
20. Eilers H, Pernthaler J, Glöckner FO, Amann R (2000) Culturability and in situ abundance of pelagic bacteria from the North Sea. *Appl Environ Microbiol* 66(7):3044–3051.
21. Barbieri E, et al. (1999) Occurrence, diversity, and pathogenicity of halophilic *Vibrio* spp. and non-O1 *Vibrio cholerae* from estuarine waters along the Italian Adriatic coast. *Appl Environ Microbiol* 65(6):2748–2753.
22. DeLoney-Marino CR, Wolfe AJ, Visick KL (2003) Chemoattraction of *Vibrio fischeri* to serine, nucleosides, and N-acetylneuraminic acid, a component of squid light-organ mucus. *Appl Environ Microbiol* 69(12):7527–7530.
23. Reidl J, Klose KE (2002) *Vibrio cholerae* and cholera: Out of the water and into the host. *FEMS Microbiol Rev* 26(2):125–139.
24. Urakawa H, Kita-Tsukamoto K, Ohwada K (1999) Reassessment of the taxonomic position of *Vibrio iliopiscarius* (Onarheim et al. 1994) and proposal for *Photobacterium iliopiscarium* comb. nov. *Int J Syst Bacteriol* 49(Pt 1):257–260.
25. Olive DM, Bean P (1999) Principles and applications of methods for DNA-based typing of microbial organisms. *J Clin Microbiol* 37(6):1661–1669.
26. Roque A, Molina-Aja A, Bolán-Mejía C, Gomez-Gil B (2001) In vitro susceptibility to 15 antibiotics of vibrios isolated from penaeid shrimps in Northwestern Mexico. *Int J Antimicrob Agents* 17(5):383–387.
27. Arias CR, Garay E, Aznar R (1995) Nested PCR method for rapid and sensitive detection of *Vibrio vulnificus* in fish, sediments, and water. *Appl Environ Microbiol* 61(9):3476–3478.
28. Thompson JR, et al. (2004) Diversity and dynamics of a north Atlantic coastal *Vibrio* community. *Appl Environ Microbiol* 70(7):4103–4110.
29. Grisez L, Chair M, Sorgeloos P, Ollevier F (1996) Mode of infection and spread of *Vibrio anguillarum* in turbot *Scophthalmus maximus* larvae after oral challenge through live feed. *Dis Aquat Organ* 26(3):181–187.
30. Sawabe T, et al. (2002) Fluorescent amplified fragment length polymorphism and repetitive extragenic palindrome-PCR fingerprinting reveal host-specific genetic diversity of *Vibrio* halioticoli-like strains isolated from the gut of Japanese abalone. *Appl Environ Microbiol* 68(8):4140–4144.
31. Vandenberghe J, et al. (1999) Vibrios associated with *Litopenaeus vannamei* larvae, postlarvae, broodstock, and hatchery probionts. *Appl Environ Microbiol* 65(6): 2592–2597.
32. Tamplin ML, Gauzens AL, Huq A, Sack DA, Colwell RR (1990) Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl Environ Microbiol* 56(6):1977–1980.
33. Heidelberg JF, et al. (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406(6795):477–483.
34. Konstantinidis KT, Ramette A, Tiedje JM (2006) Toward a more robust assessment of intraspecies diversity, using fewer genetic markers. *Appl Environ Microbiol* 72(11): 7286–7293.
35. Konstantinidis KT, Tiedje JM (2005) Towards a genome-based taxonomy for prokaryotes. *J Bacteriol* 187(18):6258–6264.
36. Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 106(45):19126–19131.
37. Okada K, Iida T, Kita-Tsukamoto K, Honda T (2005) Vibrios commonly possess two chromosomes. *J Bacteriol* 187(2):752–757.
38. Aziz RK, et al. (2008) The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* 9(1):75.
39. Langille MGI, Brinkman FSL (2009) IslandViewer: An integrated interface for computational identification and visualization of genomic islands. *Bioinformatics* 25(5): 664–665.
40. Han H, et al. (2008) Genome plasticity of *Vibrio parahaemolyticus*: microevolution of the 'pandemic group'. *BMC Genomics* 9(1):570.
41. Benson G (1999) Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* 27(2):573–580.
42. Markert S, et al. (2007) Physiological proteomics of the uncultured endosymbiont of *Riftia pachyptila*. *Science* 315(5809):247–250.
43. Karl DM, et al. (1988) A microbiological study of Guaymas basin high-temperature hydrothermal vents. *Deep-Sea Res* 35(5):777–791.
44. Webb SM, Dick GJ, Bargar JR, Tebo BM (2005) Evidence for the presence of Mn(III) intermediates in the bacterial oxidation of Mn(II). *Proc Natl Acad Sci USA* 102(15): 5558–5563.
45. Edmonds HN, et al. (2003) Discovery of abundant hydrothermal venting on the ultraslow-spreading Gakkel ridge in the Arctic Ocean. *Nature* 421(6920):252–256.
46. Dick GJ, Lee YE, Tebo BM (2006) Manganese(II)-oxidizing *Bacillus* spores in Guaymas Basin hydrothermal sediments and plumes. *Appl Environ Microbiol* 72(5):3184–3190.

47. Ridge JP, et al. (2007) A multicopper oxidase is essential for manganese oxidation and laccase-like activity in *Pedomicrobium sp*. ACM 3067. *Environ Microbiol* 9(4):944–953.
48. Vezzi A, et al. (2005) Life at depth: *Photobacterium profundum* genome sequence and expression analysis. *Science* 307(5714):1459–1461.
49. Huber KE, Waldor MK (2002) Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* 417(6889):656–659.
50. Hasan NA, et al. (2012) Genomic diversity of 2010 Haitian cholera outbreak strains. *Proc Natl Acad Sci USA* 109(29):E2010–E2017.
51. Park KS, et al. (2004) Functional characterization of two type III secretion systems of *Vibrio parahaemolyticus*. *Infect Immun* 72(11):6659–6665.
52. Bingle LE, Bailey CM, Pallen MJ (2008) Type VI secretion: A beginner's guide. *Curr Opin Microbiol* 11(1):3–8.
53. Cascales E (2008) The type VI secretion toolkit. *EMBO Rep* 9(8):735–741.
54. Filloux A, Hachani A, Bleves S (2008) The bacterial type VI secretion machine: Yet another player for protein transport across membranes. *Microbiology* 154(Pt 6): 1570–1583.
55. Weber B, Hasic M, Chen C, Wai SN, Milton DL (2009) Type VI secretion modulates quorum sensing and stress response in *Vibrio anguillarum*. *Environ Microbiol* 11(12): 3018–3028.
56. Planet PJ, Kachlany SC, Fine DH, DeSalle R, Figurski DH (2003) The widespread colonization island of *Actinobacillus actinomycetemcomitans*. *Nat Genet* 34(2):193–198.
57. Smith B, Oliver JD (2006) In situ and in vitro gene expression by *Vibrio vulnificus* during entry into, persistence within, and resuscitation from the viable but nonculturable state. *Appl Environ Microbiol* 72(2):1445–1451.
58. Murphy RA, Boyd EF (2008) Three pathogenicity islands of *Vibrio cholerae* can excise from the chromosome and form circular intermediates. *J Bacteriol* 190(2):636–647.
59. Nishibuchi M (1996) [Origin of the pathogenic vibrios in the environment: inference from the studies on the molecular genetics of *Vibrio cholerae* and *Vibrio parahaemolyticus*.]. *Nihon Saikingaku Zasshi* 51(3):823–832. Japanese.
60. Sechi LA, Duprè I, Deriu A, Fadda G, Zanetti S (2000) Distribution of *Vibrio cholerae* virulence genes among different *Vibrio* species isolated in Sardinia, Italy. *J Appl Microbiol* 88(3):475–481.
61. Klein SL, Gutierrez West CK, Mejia DM, Lovell CR (2014) Genes similar to the *Vibrio parahaemolyticus* virulence-related genes *tdh*, *tlh*, and *vscC2* occur in other Vibrionaceae species isolated from a pristine estuary. *Appl Environ Microbiol* 80(2): 595–602.
62. Vezzulli L, Guzmán CA, Colwell RR, Pruzzo C (2008) Dual role colonization factors connecting *Vibrio cholerae*'s lifestyles in human and aquatic environments open new perspectives for combating infectious diseases. *Curr Opin Biotechnol* 19(3):254–259.
63. McFall-Ngai M (2014) Divining the essence of symbiosis: Insights from the squid-vibrio model. *PLoS Biol* 12(2):e1001783.
64. Martínez JL (2013) Bacterial pathogens: From natural ecosystems to human hosts. *Environ Microbiol* 15(2):325–333.
65. Chiavelli DA, Marsh JW, Taylor RK (2001) The mannose-sensitive hemagglutinin of *Vibrio cholerae* promotes adherence to zooplankton. *Appl Environ Microbiol* 67(7): 3220–3225.
66. Reguera G, Kolter R (2005) Virulence and the environment: A novel role for *Vibrio cholerae* toxin-coregulated pili in biofilm formation on chitin. *J Bacteriol* 187(10): 3551–3555.
67. Huq A, et al. (1983) Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. *Appl Environ Microbiol* 45(1):275–283.
68. Nercessian O, Reysenbach AL, Prieur D, Jeanthon C (2003) Archaeal diversity associated with in situ samplers deployed on hydrothermal vents on the East Pacific Rise (13 degrees N). *Environ Microbiol* 5(6):492–502.
69. Boone DR, Johnson RL, Liu Y (1989) Diffusion of the interspecies electron carriers H2 and formate in methanogenic ecosystems and its implications in the measurement of Km for H2 or formate uptake. *Appl Environ Microbiol* 55(7):1735–1741.
70. Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8(3):186–194.
71. Gordon D, Abajian C, Green P (1998) Consed: A graphical tool for sequence finishing. *Genome Res* 8(3):195–202.
72. Han CS, Chain P (2006) Finishing repeat regions automatically with Dupfinisher. *International Conference on Bioinformatics and Computational Biology*, eds Arabnia HR, Valafar H (CSREA Press, Livermore, CA), pp 141–146.
73. Goryshin IY, Reznikoff WS (1998) Tn5 in vitro transposition. *J Biol Chem* 273(13): 7367–7374.
74. Kurtz S, et al. (2004) Versatile and open software for comparing large genomes. *Genome Biol* 5(2):R12.
75. Myers EW, Miller W (1988) Optimal alignments in linear space. *Comput Appl Biosci* 4(1):11–17.
76. Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4(4):406–425.

Supporting Information

Hasan et al. 10.1073/pnas.1503928112

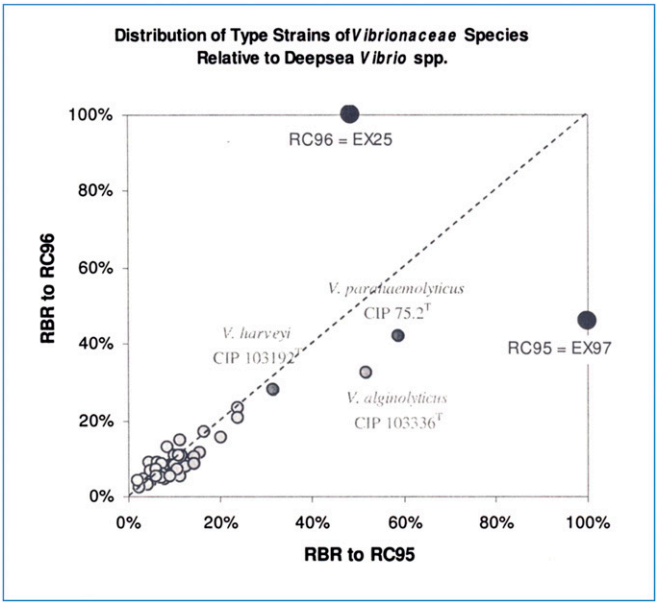


Fig. S1. DNA–DNA hybridization relative binding ratio of *Vibrio* sp. EX25 and EX97 against 42 type strains of Vibrionaceae.

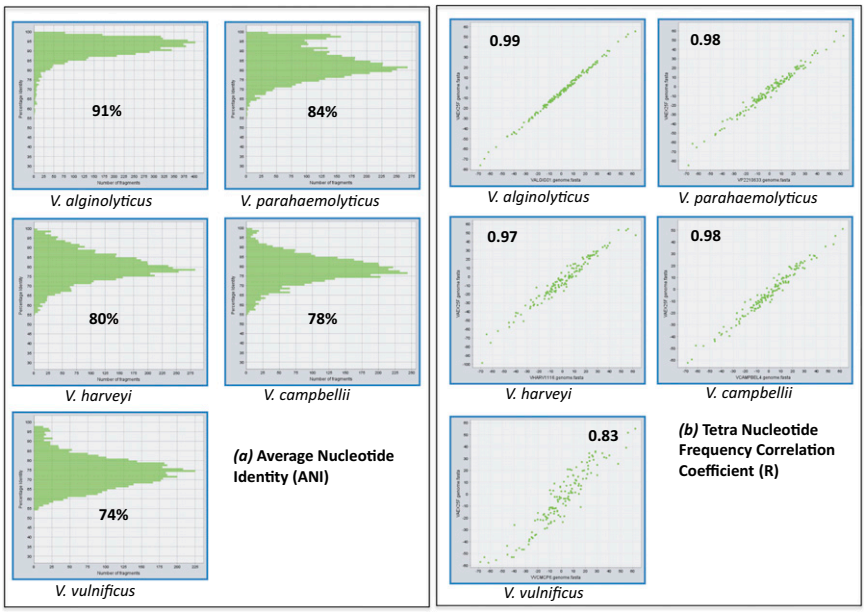
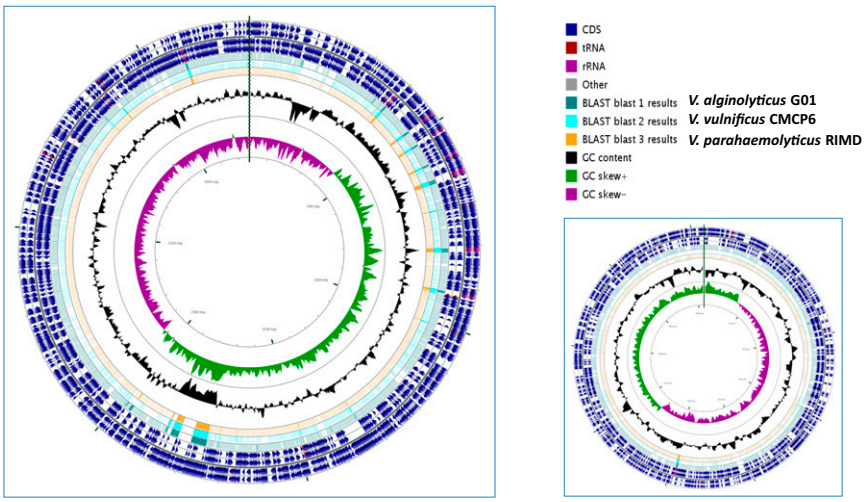


Fig. S2. Plotted values of pairwise ANI and tetra nucleotide frequency of *Vibrio antiquarius* EX25 with other *Vibrio* genome as determined by JSpecies.



***Vibrio species* EX25 Chromosome I**
Length: 3,259,580 bp

***Vibrio species* EX25 Chromosome II**
Length: 1,829,445 bp

Fig. S3. Circular maps of the chromosomes of *V. antiquarius* EX25 genome generated by CGView. From the outside in, the first-second and third-fourth circles show CDSs transcribed clockwise and anticlockwise, respectively. Circles 5, 6, and 7 display Blast bidirectional hit against *Vibrio alginolyticus* 12G01, *Vibrio parahaemolyticus* RIMD 22110633, and *Vibrio vulnificus* CMCP6; circle 8 shows GC content; and circle 9, the GC skew.

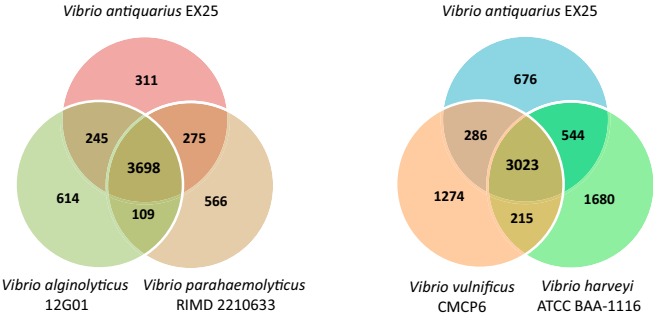


Fig. S4. Venn diagrams showing number of unique and shared orthologs between *Vibrio antiquarius* EX25 and other *Vibrio* genomes.

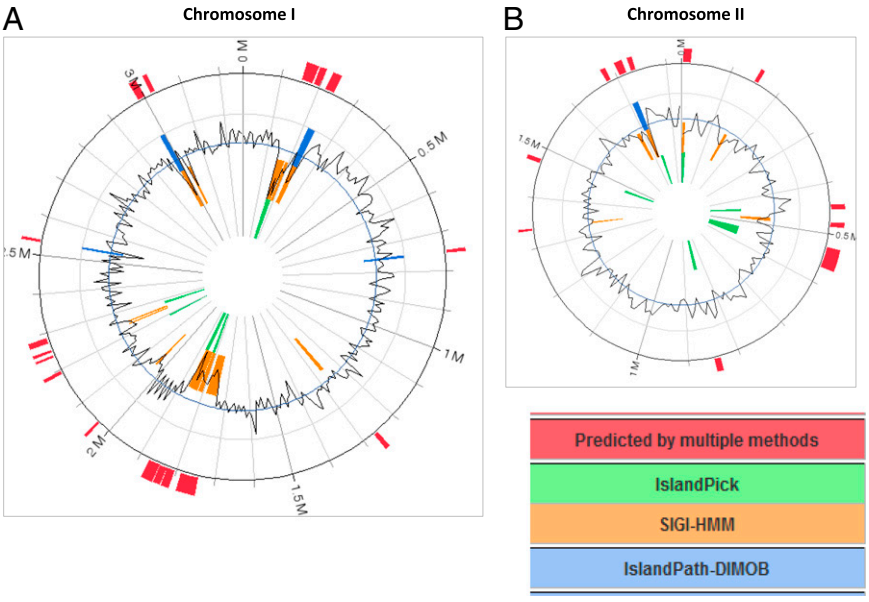


Fig. S5. Distribution of GIs in (*A*) the large chromosome (C-I) and (*B*) the small chromosome (C-II) of *V. antiquarius* EX25. Color corresponds to the GI prediction methods.

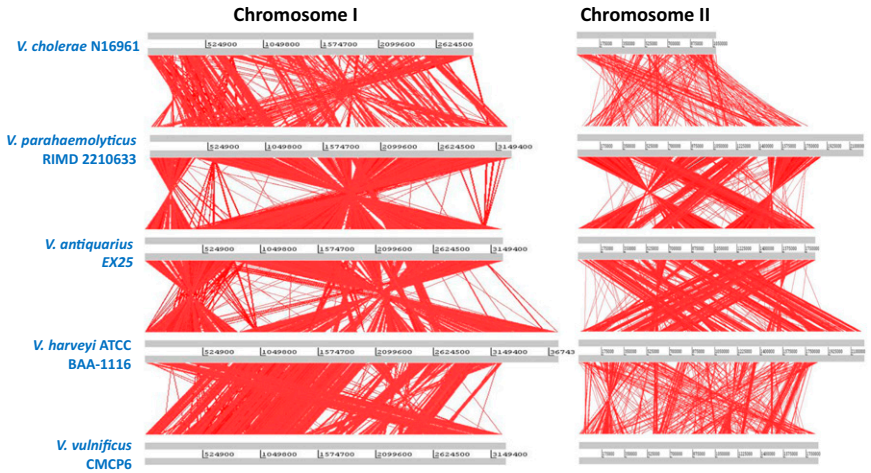


Fig. S6. Linear pairwise comparison of *V. antiquarius* EX25 genome. The analysis was performed using Artemis Comparison Tool. Regions with similarity are highlighted by connecting lines between the genomes. The gray bars represent the forward and reverse strands. EX25 has high degree of similarity with *V. parahaemolyticus* and *Vibrio harveyi*, despite of genome wise rearrangements.

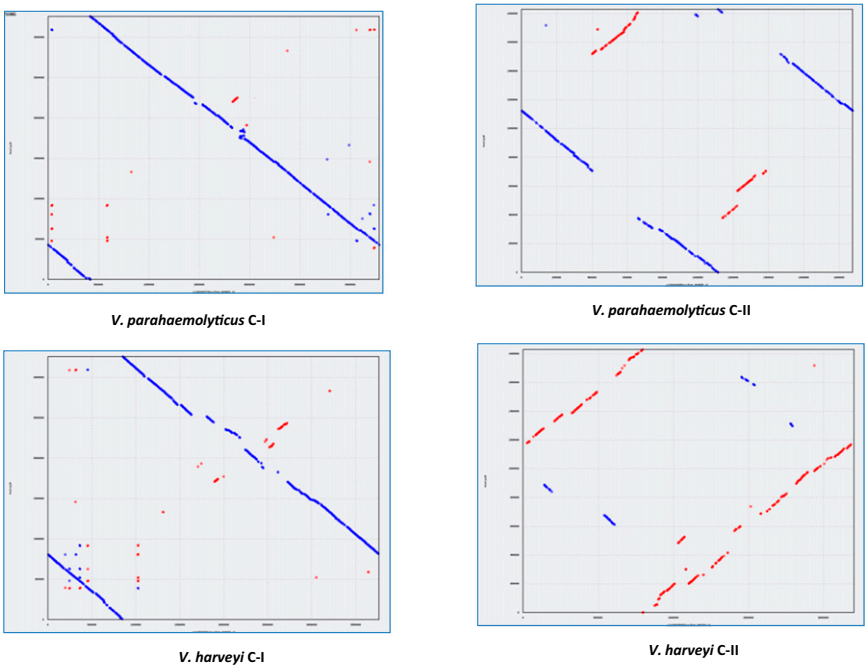


Fig. S7. Chromosome wise alignment of *V. antiquarius* EX25 genome with *V. parahaemolyticus* and *V. harveyi* genome by MUMmer demonstrates evidence of genomic inversion and rearrangement.

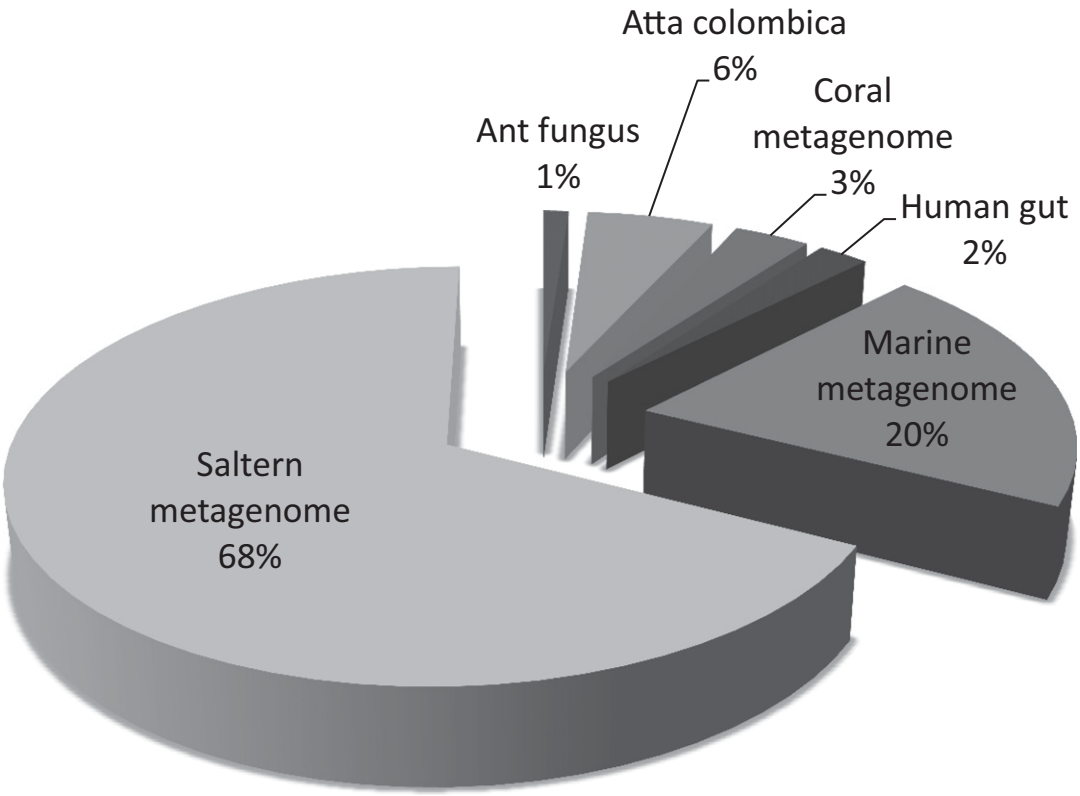


Fig. S8. Distribution *V. antiquarius* ORFs in environmental samples.

Table S1. General features of the *Vibrio antiquarius* EX25 genome compared with other *Vibrio* spp.

Feature	*V. antiquarius* EX25	*V. alginolyticus* 12G01	*V. parahaemolyticus* RIMD 2210633	*V. harveyii ATCC BAA-1116*	*V. vulnificus* CMCP6	*Vibrio cholerae* N16961
Chromosome size (Mb)	5.09	5.16*	5.17	5.97	5.12	4.03
Chromosome I	3.26	5.16	3.29	3.77	3.28	2.96
Chromosome II	1.83		1.88	2.2	1.84	1.07
Avg. G+C content (mol%)	45	44	45	45	46.4	46.7
Chromosome I	45	44	45	45	46	47
Chromosome II	45		45	45	47	46
Protein coding sequences	4,529	4,666	4,832	5,920	4,488	3,835
Chromosome I	2,846	4,666	3,080	3,546	2,926	2,742
Chromosome II	1,683		1,752	2,374	1,562	1,093
Average CDS length (bp)	942	947	950	910	918	915
Chromosome I	925	947	952	923	887	948
Chromosome II	975		946	890	977	832
Percent of coding region (%)	86.6	85.6	86	85.4	84	86
Chromosome I	87	85.6	86	85	83	87
Chromosome II	86		86	86	86	84
Ribosomal RNA operons	11	11	11	11	9	8
Chromosome I	10	11	10	10	8	8
Chromosome II	1		1	1	1	0
No. of tRNAs	124	69*	126	121	112	98
Chromosome I	111	69*	112	105	98	94
Chromosome II	13		14	16	13	4
No. of Integrons	1	0	1	?	1	1
Chromosome I	SI	0	1	?	SI	0
Chromosome II	0		0	0	0	SI

*Draft genome, therefore, numbers are approximate.

Table S2. Predicted genomic islands in *V. antiquarius* EX25 genome by Island Viewer

Chromosome	Start	End	Size	GI prediction program
I	152168	161565	9397	Predicted by multiple methods
I	152168	179105	26937	Predicted by multiple methods
I	164835	179105	14,270	Predicted by multiple methods
I	185242	198828	13,586	Predicted by multiple methods
I	216801	237262	20,461	Predicted by multiple methods
I	747045	754488	7,443	Predicted by multiple methods
I	1258282	1268568	10,286	Predicted by multiple methods
I	1743046	1787115	44,069	Predicted by multiple methods
I	1805375	1823263	17,888	Predicted by multiple methods
I	1811083	1823263	12,180	Predicted by multiple methods
I	1827588	1843329	15,741	Predicted by multiple methods
I	1846421	1873579	27,158	Predicted by multiple methods
I	1857722	1873579	15,857	Predicted by multiple methods
I	2031353	2037358	6,005	Predicted by multiple methods
I	2188274	2196673	8,399	Predicted by multiple methods
I	2236265	2240545	4,280	Predicted by multiple methods
I	2249511	2255107	5,596	Predicted by multiple methods
I	2273729	2286686	12,957	Predicted by multiple methods
I	2532043	2539396	7,353	Predicted by multiple methods
I	2983586	3000759	17,173	Predicted by multiple methods
I	3019515	3028921	9,406	Predicted by multiple methods
I	152168	161565	9,397	IslandPick
I	164835	179105	14,270	IslandPick
I	1811083	1823263	12,180	IslandPick
I	1857722	1873579	15,857	IslandPick
I	2188274	2196673	8,399	IslandPick
I	2273729	2286686	12,957	IslandPick
I	152237	178308	26,071	SIGI-HMM
I	185242	198828	13,586	SIGI-HMM
I	217871	237262	19,391	SIGI-HMM
I	1258282	1268568	10,286	SIGI-HMM
I	1743046	1787115	44,069	SIGI-HMM
I	1805375	1821178	15,803	SIGI-HMM
I	1827588	1843329	15,741	SIGI-HMM
I	1846421	1866814	20,393	SIGI-HMM
I	2031353	2037358	6,005	SIGI-HMM
I	2236265	2240545	4,280	SIGI-HMM
I	2249511	2255107	5,596	SIGI-HMM
I	2984491	2994696	10,205	SIGI-HMM
I	2994997	3000759	5,762	SIGI-HMM
I	3019515	3028921	9,406	SIGI-HMM
I	216801	236658	19,857	IslandPath-DIMOB
I	747045	754488	7,443	IslandPath-DIMOB
I	2532043	2539396	7,353	IslandPath-DIMOB
I	2983586	2998701	15,115	IslandPath-DIMOB
II	4541	18312	13,771	Predicted by multiple methods
II	149356	155892	6,536	Predicted by multiple methods
II	443598	451950	8,352	Predicted by multiple methods
II	477232	486517	9,285	Predicted by multiple methods
II	526683	537430	10,747	Predicted by multiple methods
II	537447	566943	29,496	Predicted by multiple methods
II	838513	849576	11,063	Predicted by multiple methods
II	1335191	1339216	4,025	Predicted by multiple methods
II	1469114	1477357	8,243	Predicted by multiple methods
II	1678852	1686830	7,978	Predicted by multiple methods
II	1706590	1719999	13,409	Predicted by multiple methods
II	1730672	1738710	8,038	Predicted by multiple methods
II	4541	18312	13,771	IslandPick
II	443598	451950	8,352	IslandPick
II	526683	537430	10,747	IslandPick
II	537447	566943	29,496	IslandPick
II	838513	849576	11,063	IslandPick
II	1469114	1477357	8,243	IslandPick

Table S2. Cont.

Chromosome	Start	End	Size	GI prediction program
II	1730672	1738710	8,038	IslandPick
II	5941	12176	6,235	SIGI-HMM
II	149356	155892	6,536	SIGI-HMM
II	477232	486517	9,285	SIGI-HMM
II	1335191	1339216	4,025	SIGI-HMM
II	1678852	1686830	7,978	SIGI-HMM
II	1710424	1719999	9,575	SIGI-HMM
II	1706590	1719999	13,409	IslandPath-DIMOB

Table S3. Tandem repeats in Chromosome I and II of *V. antiquarius* EX25 as identified by Tandem repeats finder program (1)

Indices	Period size	Copy number	Consensus size	Percent matches (%)	Percent indels (%)	Score	A	C	G	T	Entropy (0-2)
Chromosome I											
142579–143022	167	2.7	166	87	3	599	27	26	19	26	1.99
179301–179335	6	5.8	6	86	0	52	22	34	0	42	1.54
179298–179336	12	3.3	12	88	0	60	28	33	0	38	1.57
227892–227916	8	3.1	8	100	0	50	52	36	0	12	1.39
355495–355526	16	1.9	17	93	6	57	21	25	18	34	1.96
356893–357939	180	5.8	180	90	0	1,508	26	22	21	29	1.99
491105–491154	7	7.1	7	95	0	91	14	26	14	46	1.81
556005–556054	6	8.3	6	100	0	100	50	16	34	0	1.45
616158–616242	6	14.2	6	100	0	170	16	32	0	50	1.45
639705–639757	15	3.5	15	81	0	61	5	30	35	28	1.8
674360–674399	9	4.4	9	87	0	62	12	17	52	17	1.74
702147–703124	225	4.3	225	99	0	1,920	22	26	26	23	2
703047–703347	112	2.7	111	98	0	575	23	25	26	23	2
711051–711081	15	2.1	15	93	0	53	29	25	32	12	1.93
789899–789939	21	2	21	85	9	57	26	19	21	31	1.97
858826–858875	24	2.1	24	100	0	100	32	16	38	14	1.88
923681–924058	207	1.8	207	100	0	756	28	25	21	24	1.99
926231–926286	8	7	8	100	0	112	62	12	25	0	1.3
960542–960570	7	4.1	7	100	0	58	31	0	27	41	1.56
981315–981343	15	2	15	93	6	51	27	27	13	31	1.94
1009413–1010275	129	6.7	129	93	2	1,398	25	24	27	23	2
1009413–1010275	259	3.3	258	93	3	1,425	25	24	27	23	2
1009413–1010275	388	2.2	387	92	3	1,443	25	24	27	23	2
1025196–1025243	21	2.3	21	77	0	60	29	18	37	14	1.91
1027708–1028103	172	2.3	173	86	5	548	27	19	25	27	1.99
1040179–1040218	18	2.2	18	86	0	53	27	37	27	7	1.84
1040185–1040225	18	2.3	18	86	0	64	21	41	29	7	1.8
1040761–1040795	18	1.9	18	88	0	52	17	42	25	14	1.86
1106669–1106741	36	2.1	35	82	7	85	35	10	28	24	1.9
1117474–1117541	15	5.3	15	63	36	52	41	17	29	11	1.85
1117411–1117572	24	6.8	24	90	0	225	45	16	29	8	1.77
1137586–1137619	18	1.9	18	88	5	52	14	32	8	44	1.76
1168322–1168368	18	2.6	18	96	0	85	12	12	14	59	1.61
1184245–1184269	13	1.9	13	100	0	50	44	8	16	32	1.76
1367440–1367527	6	14.7	6	100	0	176	34	32	32	0	1.58
1404796–1405201	144	2.8	144	89	0	607	26	22	29	21	1.99
1405074–1405330	129	2	129	96	0	469	27	19	31	21	1.98
1405203–1405582	144	2.6	144	95	0	679	28	21	29	19	1.98
1404930–1405474	273	2	273	91	0	874	27	21	29	20	1.98
1404813–1405590	417	1.9	417	90	0	1,234	27	21	29	20	1.98
1485201–1486075	114	7.7	114	90	0	1,020	13	27	25	33	1.93
1534476–1534515	21	2	21	85	5	55	37	22	17	22	1.94
1536319–1536362	18	2.4	18	84	0	61	18	25	22	34	1.96
1609163–1609869	189	3.7	190	88	3	971	27	22	23	27	1.99
1625766–1625790	6	4.2	6	100	0	50	16	32	20	32	1.94
1625819–1625859	21	2	21	85	9	57	4	26	41	26	1.76
1683564–1683615	24	2.2	24	85	0	68	30	19	26	23	1.98
1719198–1719273	36	2.1	36	100	0	152	19	27	26	26	1.99
1941458–1941491	16	2.1	17	88	5	52	17	8	14	58	1.61
1980719–1982927	387	5.7	385	84	1	2,391	23	22	31	22	1.98
2052966–2053005	16	2.5	16	87	0	53	45	10	20	25	1.81
2148434–2148484	18	2.8	18	78	0	57	37	7	31	23	1.83
2328472–2328497	13	2	13	100	0	52	30	30	23	15	1.95
2355693–2355739	19	2.4	20	79	13	53	14	23	29	31	1.95
2525743–2525777	18	1.9	18	88	0	52	17	28	11	42	1.83
2548814–2548848	15	2.3	15	90	0	52	8	25	20	45	1.79
2676921–2677383	169	2.7	169	96	2	840	24	24	23	27	2
2807331–2807359	14	2	15	93	6	51	3	27	6	62	1.37
2854115–2854146	15	2.1	15	94	0	55	31	40	0	28	1.57
2884038–2885368	243	5.5	243	88	5	2,068	25	22	24	28	2
2978369–2978417	21	2.3	22	75	10	57	14	28	10	46	1.77

Table S3. Cont.

Indices	Period size	Copy number	Consensus size	Percent matches (%)	Percent indels (%)	Score	A	C	G	T	Entropy (0-2)
3044156–3044203	21	2.3	21	100	0	96	41	20	8	29	1.81
3070894–3070933	21	1.9	21	85	10	55	37	22	17	22	1.94
3225069–3225132	24	2.7	24	92	0	110	26	34	21	17	1.95
Chromosome II											
99867–99902	18	2	18	88	0	54	11	27	27	33	1.91
129129–129166	18	2.1	18	85	9	51	23	23	31	21	1.98
155231–155255	11	2.3	11	100	0	50	28	28	16	28	1.97
158409–158453	21	2.1	22	83	4	56	6	17	42	33	1.76
364794–365587	429	1.9	429	81	4	967	23	24	25	27	2
441514–441539	9	2.9	9	100	0	52	34	23	34	7	1.83
462016–462055	20	2	20	95	0	71	32	27	10	30	1.89
518949–518975	7	3.9	7	100	0	54	29	14	44	11	1.8
554384–555234	312	2.7	312	89	2	1,246	18	27	27	26	1.98
737306–737331	12	2.2	12	100	0	52	61	15	7	15	1.55
913197–913240	15	2.9	15	100	0	88	20	13	13	52	1.74
927868–927905	20	1.9	19	89	5	58	36	15	34	13	1.87
1104287–1104326	15	2.7	15	92	0	62	30	20	10	40	1.85
1119356–1119389	9	3.8	9	100	0	68	44	35	20	0	1.52
1229593–1229632	21	2	20	85	5	53	12	7	12	67	1.41
1305480–1306175	161	4.3	161	95	0	1,196	22	21	23	33	1.97
1441770–1441929	6	26.7	6	100	0	320	0	16	33	50	1.45
1567288–1567325	13	3	13	92	3	60	34	34	0	31	1.58
1724175–1724206	15	2.1	15	94	0	55	6	21	50	21	1.71
1786177–1788942	306	9	306	89	1	3,964	26	24	29	19	1.98

1. Benson G (1999) Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* 27(2):573–580.

Table S4. Insertion sequences in *V. antiquarius* EX25 genome

Sequences producing significant alignments	IS family	Group	Origin	Score (bits)	E (value)
Chromosome I					
ISVch1	IS481	—	*Vibrio cholerae* N16961	82	4E-12
ISShha1	ISNCY	IS1202	*Shewanella halifaxens*	52	0.004
ISCARN94	IS1595	IS1595	Metagenomic data from CARNOULES	46	0.22
ISPtu1	IS4	IS10	*Pseudoalteromonas tunicata*	44	0.85
ISAtu4	IS3	IS51	*Agrobacterium tumefaciens*	44	0.85
ISMae23	IS630		*Microcystis aeruginosa*	42	3.4
ISHwa10	IS66	ISBst12	*Haloquadratum walsbyi*	42	3.4
ISLac1	IS1182		*Lactobacillus acidophilus*	42	3.4
ISCARN79	IS1380		Metagenomic data from CARNOULES	42	3.4
ISStau5	IS66	ISBst12	*Stigmatella aurantiaca*	42	3.4
ISGme1	IS5	IS5	*Geobacter metallireducens*	42	3.4
ISSpma1	IS3	IS407	*Sphingopyxis macrogoltabida*	42	3.4
ISNisp1	IS3	IS3	*Nitrobacter* sp.	42	3.4
ISMac3	IS21		*Methanosarcina acetivorans*	42	3.4
ISSg1	ISL3	—	*Streptococcus gordonii* M5	42	3.4
ISIMb1	IS110	—	*Moraxella bovis* EPP63	42	3.4
ISCx1	ISL3	—	*Corynebacterium xerosis* M82B (pTP10)	42	3.4
IS231D	IS4	IS231	*Bacillus thuringiensis* subsp. *finitimus*	42	3.4
IS1112	IS30	—	*Xanthomonas oryzae* pv. *oryzae* PXO86Rif	42	3.4
Chromosome II					
IS231S	IS4	IS231	*Bacillus anthracis*	44	0.48
ISCARN12	IS200/IS605	IS1341	Metagenomic data from CARNOULES	42	1.9
ISAcma4	IS1		*Acaryochloris marina*	42	1.9
ISBthe3	IS4	IS50	*Bacteroides thetaiotaomicron*	42	1.9
ISH5	IS4	ISH8	*Halobacterium* sp.	42	1.9
ISHwa12	ISH3		*Haloquadratum walsbyi*	40	7.5
ISButh6	IS5		*Burkholderia thailandensis*	40	7.5
ISBf11	IS1380	IS942	*Bacteroides fragilis*	40	7.5
ISAar37	IS3	IS3	*Arthrobacter arilaitensis*	40	7.5
ISAar28	IS481		*Arthrobacter arilaitensis*	40	7.5
ISBcl1	IS1182		*Bacillus clausii*	40	7.5
ISAcma19	IS4	IS10	*Acaryochloris marina*	40	7.5
ISThsp12	IS1634		*Thiomonas* sp.	40	7.5
ISCARN111	IS110	IS1111	Metagenomic data from CARNOULES	40	7.5
ISCARN102	IS630		Metagenomic data from CARNOULES	40	7.5
ISAzo35	IS701		*Azoarcus* sp.	40	7.5
ISAzo28	IS110		*Azoarcus* sp.	40	7.5
ISCARN64	IS3	IS3	Metagenomic data from CARNOULES	40	7.5
ISSpn5	IS1380		*Streptococcus pneumoniae*	40	7.5
ISVsh1	ISAs1		*Vibrio shilonii*	40	7.5
ISFnu8	IS3		*Fusobacterium nucleatum*	40	7.5
ISRhba1	IS1595	ISNwi1	*Rhodobacterales bacterium*	40	7.5
ISDha3	IS4	IS4Sa	*Desulfitobacterium hafniense*	40	7.5
ISSod13	IS481		*Shewanella oneidensis*	40	7.5
ISC774	IS6		*Sulfolobus solfataricus*	40	7.5
ISVvu1	IS1		*Vibrio vulnificus*	40	7.5
IS231L	IS4	IS231	*Bacillus anthracis*	40	7.5
ISVpa1	IS110	IS1111	*Vibrio parahaemolyticus*	40	7.5
ISWz1	IS91	—	*Weeksella zoohelcum*	40	7.5
IS91B	IS91	—	*Escherichia coli* G7 (pRI8801)	40	7.5
IS1400	IS3	IS407	*Yersinia enterocolitica* Ye 8081	40	7.5
IS1380B	IS1380	—	*Acetobacter pasteurianus* NC11380	40	7.5
IS1380A	IS1380	—	*Acetobacter pasteurianus* NC11380	40	7.5