

RESEARCH

Open Access



Analysis of *Vibrio cholerae* genomes identifies new type VI secretion system gene clusters

Cristian V. Crisan^{1,2,3†}, Aroon T. Chande^{2,3,4,5†}, Kenneth Williams^{1,2,3}, Vishnu Raghuram^{1,2,3}, Lavanya Rishishwar^{2,3,4,5}, Gabi Steinbach^{2,4,6}, Samit S. Watve⁷, Peter Yunker^{2,4,6}, I. King Jordan^{2,3,4,5} and Brian K. Hammer^{1,2,3*} 

Abstract

Background: Like many bacteria, *Vibrio cholerae* deploys a harpoon-like type VI secretion system (T6SS) to compete against other microbes in environmental and host settings. The T6SS punctures adjacent cells and delivers toxic effector proteins that are harmless to bacteria carrying cognate immunity factors. Only four effector/immunity pairs encoded on one large and three auxiliary gene clusters have been characterized from largely clonal, patient-derived strains of *V. cholerae*.

Results: We sequence two dozen *V. cholerae* strain genomes from diverse sources and develop a novel and adaptable bioinformatics tool based on hidden Markov models. We identify two new T6SS auxiliary gene clusters and describe Aux 5 here. Four Aux 5 loci are present in the host strain, each with an atypical effector/immunity gene organization. Structural prediction of the putative effector indicates it is a lipase, which we name TleV1 (type VI lipase effector *Vibrio*). Ectopic TleV1 expression induces toxicity in *Escherichia coli*, which is rescued by co-expression of the TliV1a immunity factor. A clinical *V. cholerae* reference strain expressing the Aux 5 cluster uses TleV1 to lyse its parental strain upon contact via its T6SS but is unable to kill parental cells expressing the TliV1a immunity factor.

Conclusion: We develop a novel bioinformatics method and identify new T6SS gene clusters in *V. cholerae*. We also show the TleV1 toxin is delivered in a T6SS manner by *V. cholerae* and can lyse other bacterial cells. Our web-based tool can be modified to identify additional novel T6SS genomic loci in diverse bacterial species.

Keywords: *Vibrio cholerae*, Type VI secretion clusters, Bacterial toxins

Background

Vibrio cholerae is a globally dispersed, Gram-negative bacterium that naturally resides on chitinous surfaces in marine habitats. When ingested, some strains of *V. cholerae* can cause the fatal cholera diarrheal disease in humans. While relatively rare in developed countries, it is estimated that nearly 3,000,000 cases and 100,000 deaths from cholera occur annually, with the disease endemic to areas of the Middle East and Southern Asia [1, 2]. Patient-derived strains (referred to as clinical strains) of *V. cholerae* possess

virulence factors that help colonize the intestine and infect the human host [3]. *V. cholerae* strains also possess other mechanisms to colonize hosts and persist in aquatic niches [4]. An important defense employed by *V. cholerae* against other prokaryotic and eukaryotic cells is the type VI secretion system (T6SS), a protein delivery system that punctures membranes of neighboring cells and delivers toxic effectors (Fig. 1a) [5, 6].

Found in approximately 25% of all Gram-negative bacterial species, the T6SS apparatus consists of a membrane complex that spans both membranes and the periplasm of the host cell [7, 8]. A baseplate complex with homology to phage components attaches to the inner membrane and is thought to interact with other components of the apparatus [8, 9]. The T6SS functions through an ATP-dependent contractile mechanism facilitated by VipA/B

* Correspondence: bhammer@gatech.edu

†Cristian V. Crisan and Aroon T. Chande contributed equally to this work.

¹Center for Microbial Dynamics and Infection, Georgia Institute of Technology, 310 Ferst Drive, Atlanta, GA 30332-0230, USA

²Institute for Bioengineering and Biosciences, Georgia Institute of Technology, 310 Ferst Drive, Atlanta, GA 30332-0230, USA

Full list of author information is available at the end of the article



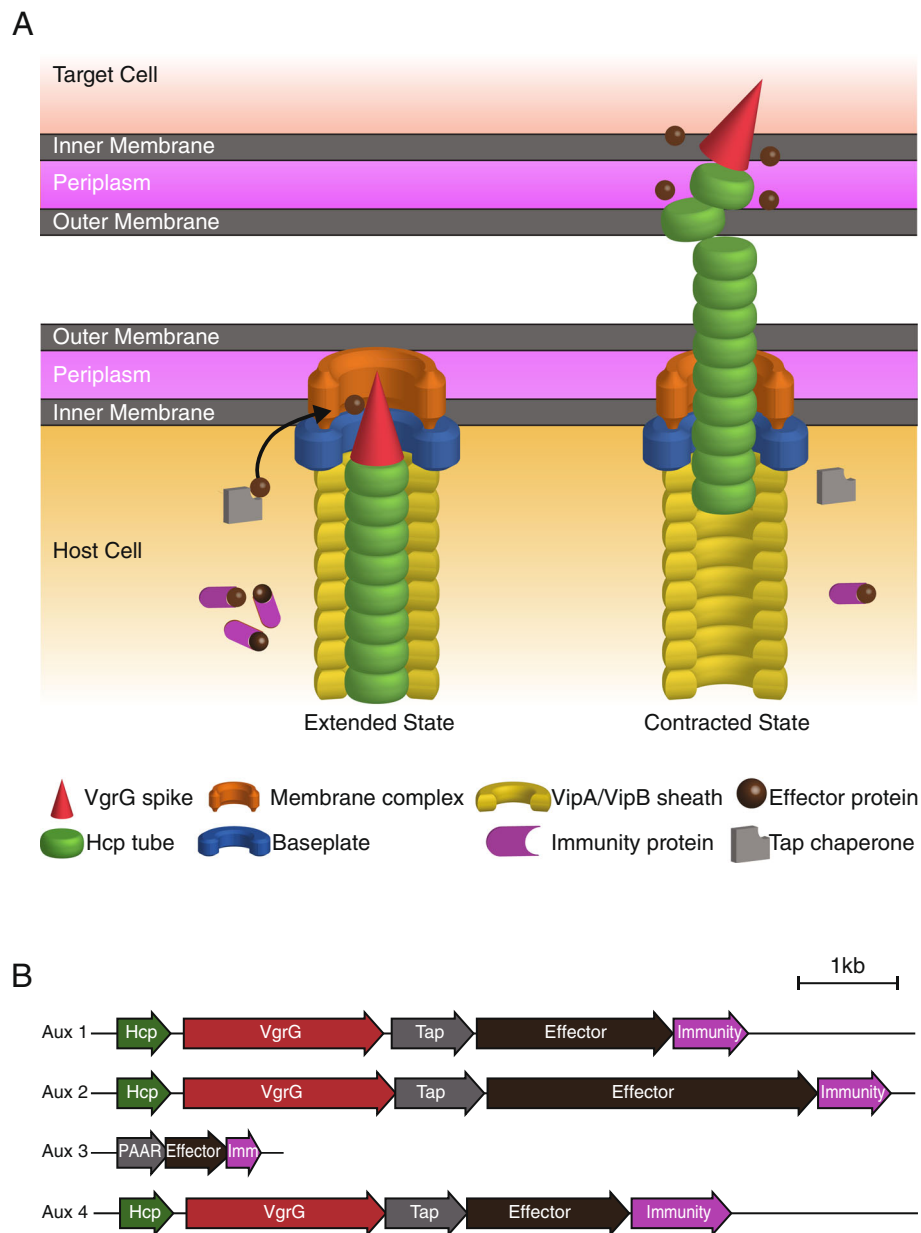


Fig. 1 Type VI secretion system of *Vibrio cholerae*. **a** Diagrammatic depiction of the T6SS apparatus extension and contraction in *V. cholerae*. The apparatus is composed of a membrane-spanning and a baseplate complex, an outer contractile sheath (VipA/B), and a needle complex (Hcp and VgrG). Effectors can interact directly with VgrG or PAAR proteins, may require chaperones for delivery on the apparatus, or be carried as cargo in the T6SS apparatus. **b** Aux clusters 1, 2, and 4 share a canonical *hcp*, *vgrG*, *tap*, *effector*, *immunity* gene organization in all strains where they are found

sheath proteins [10–12]. Hemolysin-coregulated protein (Hcp) hexamers form the inner tube of the apparatus and are exported into the extracellular milieu following a contraction of the outer sheath [5, 6, 13, 14]. The tip of the apparatus is comprised of secreted VgrG proteins that interact with T6SS toxic proteins (called effectors) to aid in their delivery [15]. PAAR proteins, found in some bacterial species harboring the T6SS, associate with VgrGs and are thought to sharpen the tip while also diversifying the cargo delivered by the T6SS [16, 17].

In sequenced *V. cholerae* strains, most structural and regulatory T6SS components are encoded on a single locus on chromosome II, referred to as the large cluster (Fig. 1b). Additional components, including Hcp proteins, are encoded on two auxiliary clusters—auxiliary clusters 1 and 2 (Aux 1 and 2, respectively). Each of the three clusters also encodes a VgrG (Fig. 1b) [12, 18]. The VgrG encoded on the large cluster contains an additional C-terminal domain with antibacterial (lysozyme-like) activity, while the VgrG found on Aux 1 contains an

anti-eukaryotic (actin-crosslinking) C-terminus domain in some strains [8, 19]. Terminal genes of canonical T6SS auxiliary clusters encode a secreted effector and a cognate immunity protein. Loss of immunity proteins makes cells susceptible to T6SS attacks from neighboring siblings [20, 21]. Both auxiliary clusters also encode T6SS adaptor proteins (TAPs) that are thought to be critical in loading specific effectors onto the T6SS apparatus and have been used as genomic markers to identify novel T6SS effectors [15, 22]. An additional cluster discovered later, Aux 3, lacks *hcp*, *vgrG*, and *tap* open reading frames but contains genes coding for an effector (*tseH*) and an immunity protein (*tsiH*) [23]. The Aux 3 cluster also contains a *paar* gene whose product may allow the effector to be secreted by the VgrG of another cluster for delivery [23].

Regulation of T6SS genes in *V. cholerae* varies. Clinical strains, such as C6706 and A1552, show little T6SS activity in a rich growth medium [24–26]. Expression of genes encoded on the large T6SS cluster is upregulated by the QstR protein, which integrates signals from three other regulators: CytR (responding to nucleoside starvation), HapR (responding to quorum-sensing signals), and TfoX (responding to chitin oligomers) [27–30]. By contrast, the majority of *V. cholerae* that have no history of human pathogenicity (referred to as environmental strains) express the T6SS constitutively in rich growth medium and can kill other bacterial cells in a contact-dependent manner [25]. The regulation (if any) of the T6SS genes in those strains is currently not understood.

Results

Genome sequencing and assessment of diversity across isolates

Average nucleotide identity (ANI) was used to assess the genetic variation for environmental and clinical strains of *V. cholerae* from this study and for publicly available *V. cholerae* genomes from NCBI [31]. Strain and assembly information are summarized in Additional file 2: Table S1. ANI revealed six clusters of *V. cholerae* strains, with clinical strains clustering together and environmental strains forming several distinct clusters (Fig. 2). SIO (BH2680), the out-group, had ANI values close to 0.96 and is at the edge of the *V. cholerae* species boundary.

T6SS module typing and annotation

Canonical *V. cholerae* T6SS loci have conserved synteny, which was used to localize searches around *vgrG* sequences to reduce the required number of BLAST searches. Initial annotation using BLAST against previously reported effector sequences was partially successful [23, 32]. Large, Aux 1 and Aux 2 *vgrG* alleles were successfully annotated in most strains, with occasional mis-annotation of *vgrG-1* alleles as *vgrG-2* and vice versa. Using this approach, we confirmed the presence of all

three canonical T6SS loci (the large, Aux 1, and 2 clusters) in all sequenced isolates and the presence of the Aux 3 cluster in 30% of the isolates (Fig. 2).

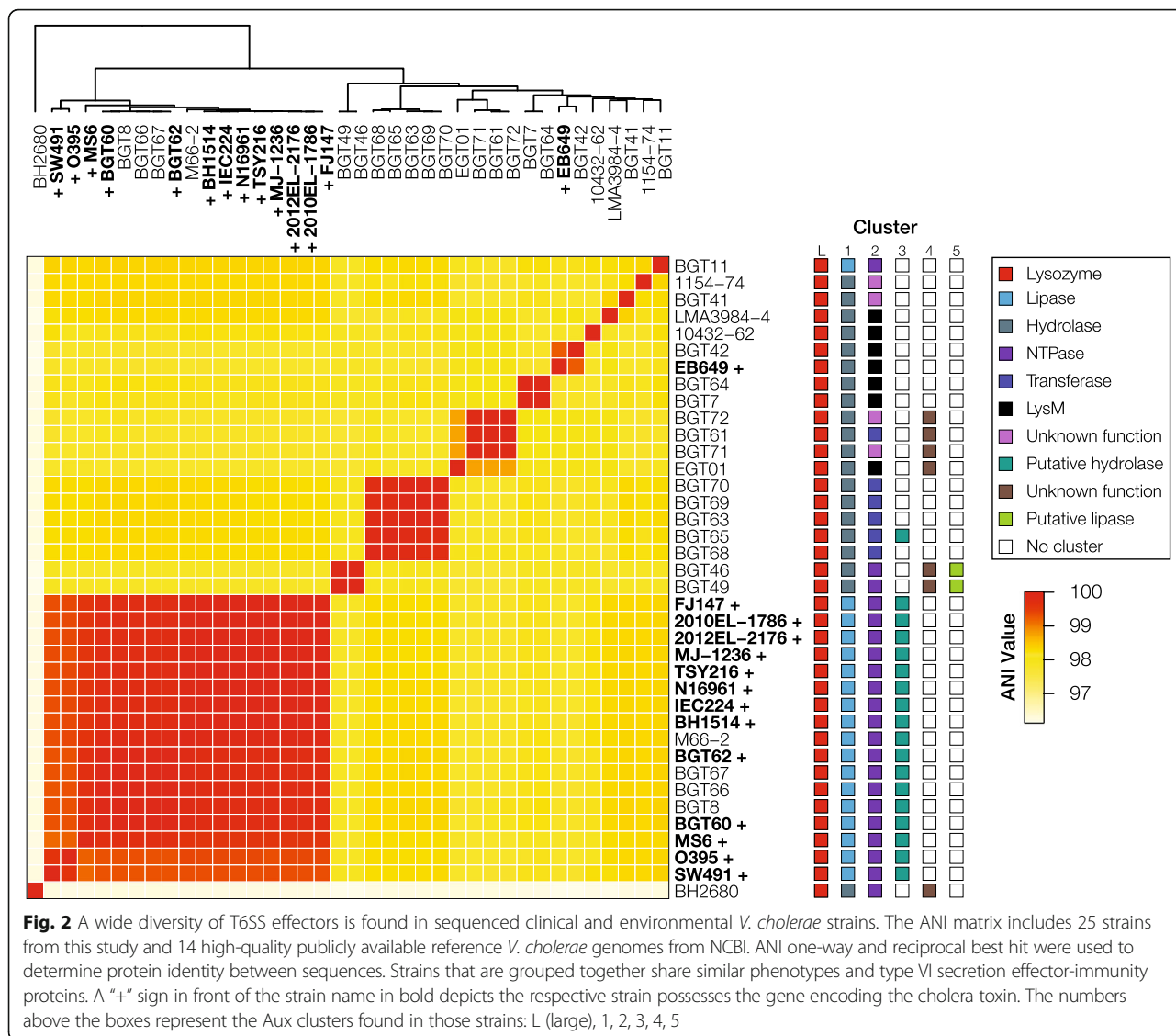
The conserved gene order was then used to aid in effector assignment and to identify several putative novel effectors for Aux 1 and 2. All effectors were typed and placed into classes based on conserved structural and/or functional domains (Fig. 2). T6SS effector proteins in Aux 1 were classified as lipases or hydrolases (with a DUF2235 domain). Most Aux 2 effectors were assigned as NTPases, transferases, and “LysM-like” proteins. Several Aux 2 effectors (found in strains 1154-74, BGT41, BGT71, and BGT72) contain no conserved domains for typing and are dissimilar from other reported effectors and were denoted as having an “unknown function” (Fig. 2).

Hidden Markov models for effector prediction and annotation of new T6SS loci

To investigate whether sequenced *V. cholerae* strains contain additional, non-canonical T6SS loci, hidden Markov models (HMMs) were built for degenerate *hcp*, *vgrG*, and DUF2235 hydrolase domains. Using a degenerate *hcp* HMM, an additional *hcp*-like allele was identified in six environmental strains: BGT46, BGT49, BGT61, BGT71, BGT72, and EGT01. The degenerate *vgrG* HMM identified an additional pseudo-*vgrG* in the same six strains, in-frame, and directly downstream of the *hcp*-like CDS. Furthermore, the gene directly downstream of the pseudo-*vgrG* contains a DUF4123 domain found in *tap* genes. Predicted effector, immunity, and *paar* genes were also observed downstream of the *tap* gene. A similar cluster previously identified in other *V. cholerae* isolates was annotated in this study as auxiliary cluster 4 (Aux 4) [33]. Aux 4 is distinct in structure, content, and genomic localization from Aux 3 and is present in strains containing both Aux 1 and Aux 2 clusters.

T6SS Predictor: a web tool for prediction of *V. cholerae*-specific T6SS proteins

We also developed a tool for rapid prediction and annotation of putative T6SS loci and proteins. T6SS Predictor utilizes the profile HMMs developed for Hcp, VgrG, TAP, and proteins from each effector class to annotate cluster components individually. Genomic localization and low-stringency BLAST searches using consensus sequences for each cluster/effector combination are used to assign predicted proteins to a particular cluster. Effectors are annotated using a combination of profile HMM typing and BLAST to the custom Conserved Domains Database used in this study. The large cluster is not annotated by T6SS Predictor. In our testing, using strains sequenced in this study, strains from Unterweger et al., and the other reference strains used in this study (Fig. 2), T6SS Predictor reliably predicts and annotates Aux 1, 2, and 3 in clinical and environmental strains



and predicts Aux 4 and 5 VgrG proteins and effectors in environmental strains [32]. T6SS Predictor attempts to return visualizations of each locus annotated; however, contig breaks sometimes prevent the correct ordering of proteins. As a result, an annotated FASTA file containing all the predicted, putative T6SS components is provided as well.

Aux 5 clusters have an atypical genomic organization

A profile HMM constructed for Aux 1 DUF2235 effectors (hydrolases) identified new putative T6SS loci in two related strains (BGT46 and 49, Fig. 2). This cluster is annotated as auxiliary cluster 5 (Aux 5) and is distinct in content and genomic organization from Aux 1, 2, 3, and 4 (Additional file 1: Figure S1). Aux 5 is present in *V. cholerae* strains that encode the Aux 1, 2, and 4 clusters (Fig. 2, colored boxes). The genomic organization of

Aux 5 clusters is different than that of canonical T6SS auxiliary clusters in *V. cholerae* (Fig. 3a). Specifically, no open reading frames are found immediately downstream of predicted Aux 5 effectors. Instead, two genes containing DUF3304 domains found in other T6SS immunity proteins are present upstream of each effector gene.

PacBio sequencing of strain BGT49 identified an Aux 5 cluster at four distinct genomic locations (Fig. 3a, Additional file 1: Figure S1). All four Aux 5 loci (Aux 5a, b, c, d) have the same gene organization and share more than 93% nucleotide homology (Fig. 3a).

TleV1 is toxic to *E. coli* cells and can be used in intra-species T6SS-mediated competition

Each predicted effector encoded within the four Aux 5 clusters contains a DUF2235 hydrolase domain, found in other T6SS-associated effectors from *Pseudomonas*

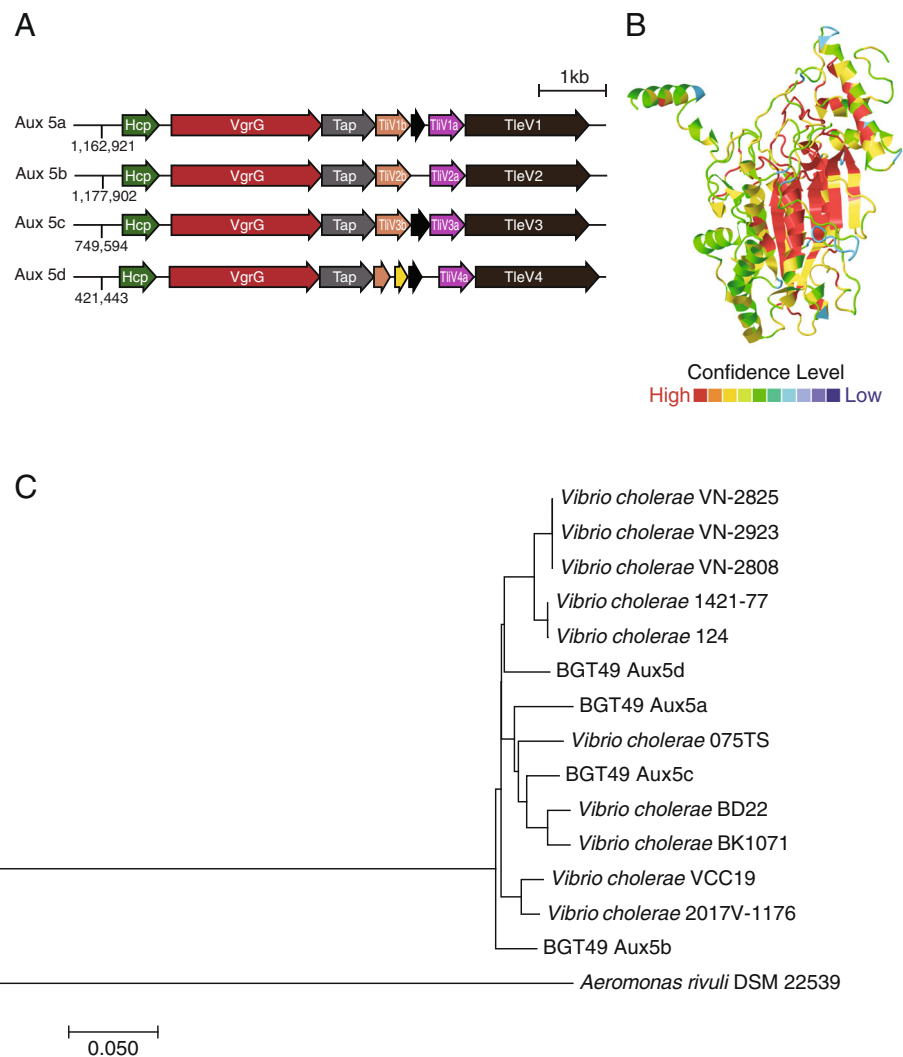


Fig. 3 Aux 5 clusters have an atypical gene organization and encode predicted lipases. **a** The novel Aux 5 cluster sequences from the four distinct genomic loci were aligned. The atypical Aux cluster organization is observed in all four Aux 5 clusters. Numbers at the beginning of the clusters represent the genomic position on the BGT49 chromosome. **b** The structure of TleV1 was predicted using the Phyre2 webserver. The *Pseudomonas aeruginosa* Tle1 crystal structure is predicted with high confidence to be homologous to the Aux 5 putative effector. The color scheme depicts the alignment confidence of the Phyre2 model to the Tle1 crystal structure. The image was obtained using JSmol. **c** Aux 5-like clusters were identified in other *V. cholerae* strains and an *Aeromonas* strain and analyzed phylogenetically

aeruginosa, *Escherichia coli*, and *Burkholderia thailandensis* [34]. Phyre2 predicts with high confidence the putative effector found on cluster Aux 5a is a homolog of the T6SS effector Tle1 from *P. aeruginosa*, despite sharing only 19% primary sequence identity (Fig. 3b, Additional file 1: Figure S2) [35, 36]. These results reveal that the effectors belong to the larger family of Tle1 lipases that can target phospholipids and destabilize membranes. We named the putative effectors found within the Aux 5 clusters TleV 1–4 (type VI lipase effector *Vibrio* 1–4) (Fig. 3a).

To experimentally validate the activity of the Aux 5a cluster, the toxicity of TleV1 was first assessed. The

wild-type *tleV1* gene was expressed in *Escherichia coli* cells under the control of the arabinose-inducible pBAD promoter. Based on the predicted structure and previous similar studies showing that Tle1 lipases have activity when delivered to the periplasm, TleV1 was also expressed in *E. coli* cells with an N-terminal periplasmic Tat (twin-arginine translocation pathway) signal [34, 37]. When its expression was induced by arabinose, TleV1 was most cytotoxic when delivered to the periplasm, but also had moderate toxicity in the cytoplasm (Fig. 4a).

To determine whether TleV1 can be loaded onto the T6SS and be delivered to target cells, the entire Aux 5a cluster was integrated by allelic exchange methods into

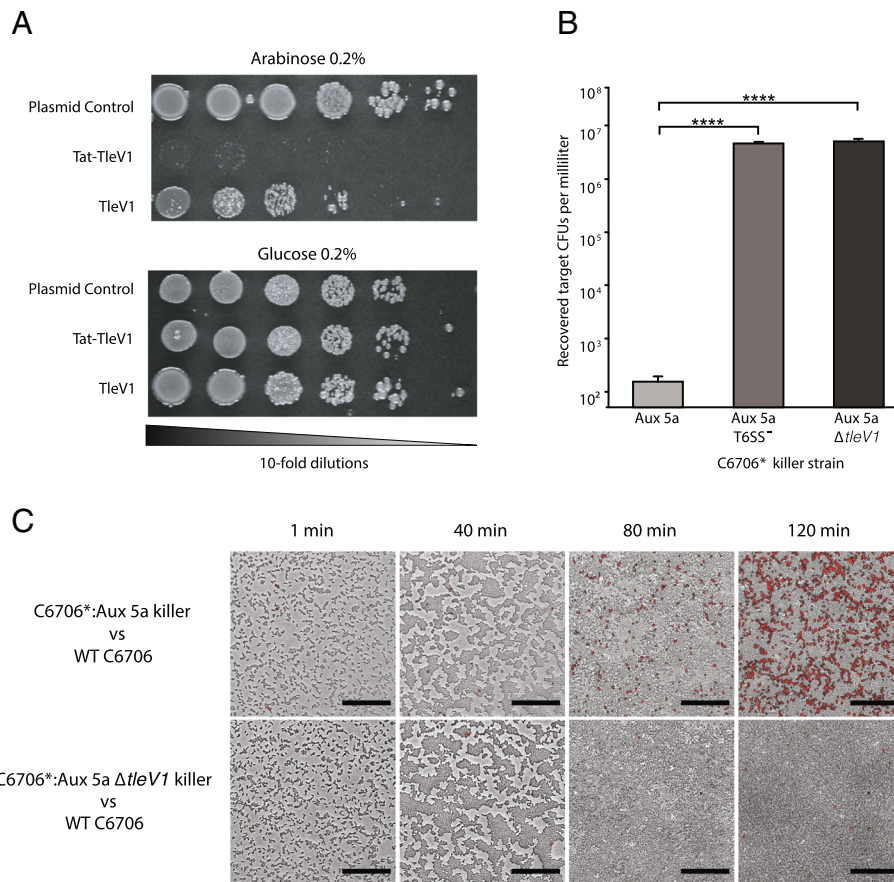


Fig. 4 TleV1 is toxic to both *E. coli* and *V. cholerae* cells. **a** Wild-type and periplasmic Tat-tagged *tleV1* genes were expressed in *E. coli* cells under the control of the pBAD promoter. Cells carrying the effector were then spotted on glucose 0.2% or arabinose 0.2% plates (and antibiotic to maintain the plasmid). **b** The clinical wild-type C6706 *V. cholerae* strain was competed with C6706* (a C6706 strain that constitutively expresses QstR and the T6SS apparatus) with an integrated Aux 5a cluster (C6706*:Aux 5a) at its *lacZ* gene locus. A T6SS⁻ C6706*:Aux 5a mutant and a C6706*:Aux 5a $\Delta tleV1$ mutant were also competed against the WT C6706 target. A one-way ANOVA with post-hoc Tukey HSD test was performed, **** $p < 0.0001$. **c** Competitions between WT C6706 vs. C6706*:Aux 5a and WT C6706 vs. C6706*:Aux 5a $\Delta tleV1$ were visualized using propidium iodide (staining red cells with a compromised membrane) as an indicator for cellular lysis. Black scale bars represent 40 μ m

the *lacZ* gene locus of a clinical *V. cholerae* C6706 strain, which we denote C6706*:Aux 5a [29]. C6706*:Aux 5a expresses the T6SS constitutively because the gene encoding QstR is under the control of the P_{tac} constitutive promoter. A competition killing assay was then performed using the *V. cholerae* C6706 strain with the integrated Aux 5a cluster (C6706*:Aux 5a) as the killer strain and wild-type C6706 as the target strain. C6706*:Aux 5a outcompeted wild-type C6706 and reduced the number of surviving wild-type C6706 by almost 5 orders of magnitude (Fig. 4b). A C6706*:Aux 5a strain with a deletion in the essential T6SS membrane complex *vasK* gene was unable to outcompete wild-type C6706, showing that Aux 5a-mediated killing was T6SS-dependent [38]. Furthermore, when *tleV1* was deleted from C6706*:Aux 5a, the strain was also unable to outcompete wild-type C6706.

We identified Aux 5-like clusters in 10 other *V. cholerae* strains and an *Aeromonas* strain (Fig. 3c, Additional file 2: Table S2). As shown in Fig. 3c, each of the four Aux 5 clusters present in *V. cholerae* strain BGT49 share more similarity to clusters from other strains than they share among themselves. This suggests they may have been horizontally acquired by BGT49 as separate events, although no phage or integrase genes were detected in the vicinity of the Aux 5a cluster (Additional file 1: Figure S3). Furthermore, the GC content of the cluster is similar to the regions flanking it (Additional file 1: Figure S3).

To test the hypothesis that a cluster could be obtained by natural transformation, a kanamycin resistance cassette was introduced immediately downstream of the Aux 5a gene cluster in BGT49 using chitin-induced natural transformation. Genetic manipulation

of BGT49 is difficult because the strain was refractory to plasmid uptake by standard methods like mating or electroporation. Kanamycin-marked BGT49 genomic DNA was then used in a second natural transformation event to integrate the Aux 5 cluster into the genome of C6706. The C6706 strain containing the Aux 5 cluster was then able to successfully kill wild-type C6706 strain in a T6SS-dependent manner (data not shown). However, we observed that during transformation, more than one Aux 5 cluster was transferred to the C6706 strain.

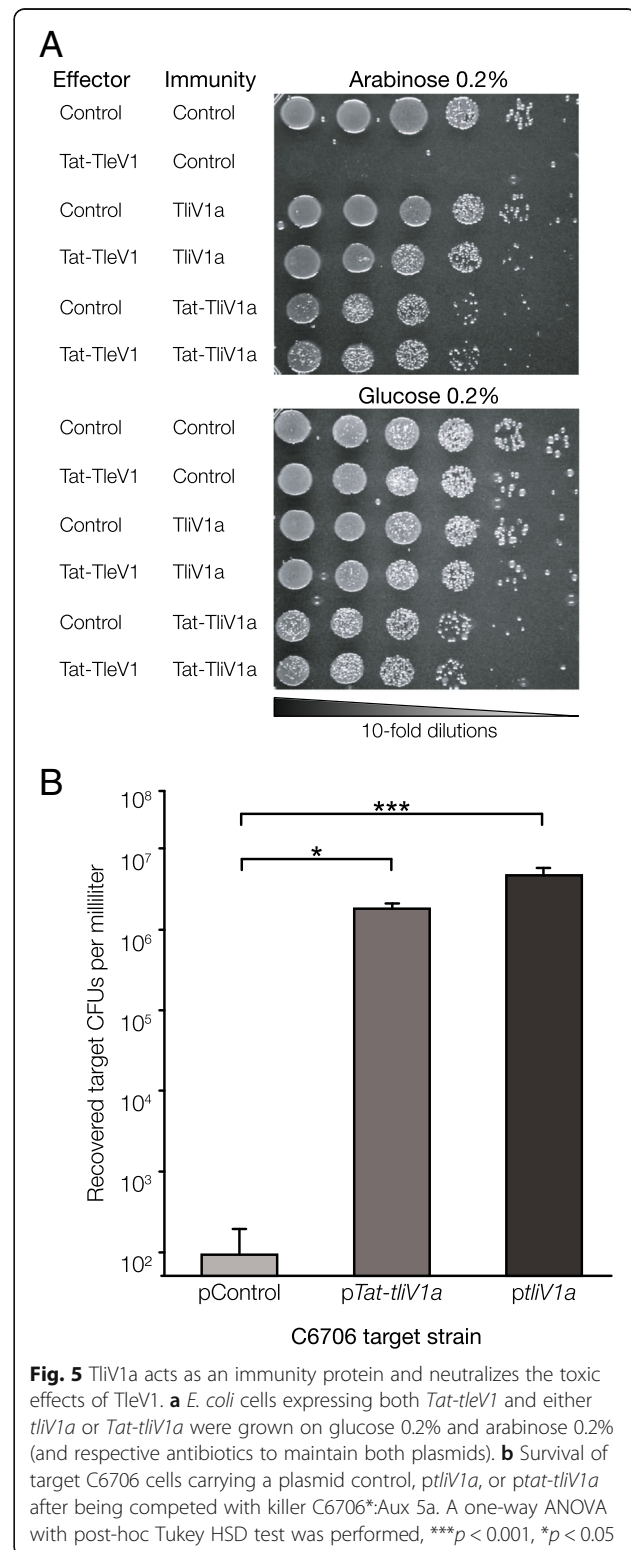
To determine whether TleV1 is toxic to cells in a manner consistent with a lipase, we examined killing induced by TleV1 using confocal microscopy (Nikon A1plus). Propidium iodide, which stains the DNA of dead cells with a compromised membrane, was used to observe cell lysis. A large number of dead cells were detected when C6706*:Aux 5a killer cells were mixed with target wild-type C6706 cells (Fig. 4c). A few dead cells were observed at low cell density with little cell contact, but substantial killing occurred after 2 h, when cells became densely packed. By contrast, in competitions where killer C6706*:Aux 5a cells had a $\Delta tleV1$ deletion, only an occasional dead cell was detected throughout the time period. This result suggests TleV1 acts as a bactericidal effector when delivered into target cells.

TliV1a can neutralize the toxic effects of TleV1

Unlike other T6SS auxiliary clusters in *V. cholerae*, where a single immunity gene is usually found downstream of an effector gene, two alleles coding for predicted immunity proteins were found upstream of each effector in all four Aux 5 clusters. For Aux 5a, we named the two genes upstream of the *tleV1* effector as *tliV1a* and *tliV1b* (type VI lipase immunity *Vibrio* 1a and 1b) (Fig. 3a). To test whether immunity gene *tliV1a*, encoded directly upstream of *tleV1*, can prevent self-intoxication of *E. coli* cells expressing TleV1, wild-type TliV1a, or periplasmically directed Tat-TliV1a were expressed from a second plasmid in the same cells under the control of the P_{tac} promoter. Survival of *E. coli* cells expressing both Tat-TleV1 and TliV1a, or both Tat-TleV1 and Tat-TliV1a, was comparable to the survival of cells containing control plasmids, indicating that co-expression of the immunity gene can neutralize the toxicity of TleV1 (Fig. 5a), as shown for other effector-immunity pairs [21, 32].

To confirm that TliV1a can behave as an immunity protein, both TliV1a and Tat-TliV1a were expressed in wild-type C6706 *V. cholerae*. C6706 expressing either TliV1a or Tat-TliV1a was then competed as a target with killer C6706*:Aux 5a. Expression of either TliV1a or Tat-TliV1a significantly rescued survival of C6706

cells as compared to C6706 cells expressing a control plasmid (Fig. 5b). These results indicate TliV1a can act as an immunity protein and prevent cellular intoxication caused by TleV1.



Discussion

Whole-genome comparison and diversity assessment

Average nucleotide identity (ANI) has replaced DNA-DNA hybridization as the species typing tool in the genomic era [39]. BLAST-based ANI (ANIb) has a strict species cutoff, with ANIb values < 0.96 indicating different species [39–41]. As expected, clinical strains of *V. cholerae* clustered together likely due to their clonal nature [42–44]. ANIb values greater than 99% are often used for subspecies or strain delineation, further supporting the clonal origins of clinical samples [45]. Dot plots between strains in different ANI clusters show few, small (< 20 kb) rearrangements and many, small unique genomic regions (data not shown), consistent with frequent horizontal gene transfer, perhaps due to natural competence in *V. cholerae* [25, 46, 47].

Although *V. cholerae* strains BGT61, 71, and 72 are genetically similar and were collected in the same year (1978), they were isolated from locations more than 5000 miles apart (Additional file 2: Table S1). The results suggest *V. cholerae* may be widely distributed by environmental and human factors into diverse environmental reservoirs. EGT01 is genetically similar to BGT61, 71, and 72 but was collected 33 years later (2011) from gray water (water from non-sewage, home water sources) in Haiti following the 2010 cholera outbreak yet shares many of the same genomic features. EGT01 also encodes two bacterial CRISPR systems absent from the other strains, including one upstream of a T6SS cluster [31].

Comparison with other T6SS annotation methods

Prior studies noted the difficulty involved with accurate identification and classification of diverse T6SS proteins. Unterweger et al. used a common approach, “uclust-then-BLAST,” in which predicted proteins are clustered (typically by 95% identity) followed by bi-directional best-hit BLAST searches [32]. This technique is well suited for gene finding and annotation of well-characterized, conserved sequences. BLAST-based approaches, as used in this study and by Unterweger et al., are also able to accurately annotate sequences with high conservation, allowing for rapid identification of canonical T6SS proteins [32]. However, because of BLAST’s reliance on direct sequence comparisons and relatively high stringency in matching criteria, this approach is not well suited for exploratory annotation, especially in cases where large sequence divergence is expected. Less stringent BLAST searches can produce tens of off-target hits, such as the many transmembrane proteins that partially match VgrGs, which require significant manual curation. Manual curation is further complicated due to contig breaks, which can make unambiguous assignments of putative loci more difficult.

An existing annotation tool, SecReT6, adopts a similar clustering and BLAST approach with the addition of

profile HMMs for rapid sussing out prior to BLAST [48]. SecReT6’s T6SS effector database only contains T6SS alleles from clinical strains but as shown by this study and Unterweger et al., clinical strains typically contain the same effectors at Aux 1 and 2 [32]. Thus, using clinical strains as the basis for effector typing under-represents the known sequence diversity of effector proteins and restricts annotations by SecReT6 to a limited set of *V. cholerae* effectors. The database contains 76 secreted effector proteins, covering the large cluster VgrG, lipase class Aux 1 proteins, and NPPase/transferase class Aux 2 proteins. SecReT6 is unable to identify T6SS loci in the environmental strains in this study without lipase or NPPase Aux 1 and 2 effectors, respectively, and does not detect Aux 4 or 5 effectors. Additionally, such tools are unable to provide annotations of divergent structural proteins, such as the *hcp* and *vgrG* alleles found in Aux 4 and 5, and the effectors at those loci, preventing their discovery.

The classification approach taken here differs from that used by Unterweger et al., which relied on comparing relatively large “effector modules” containing multiple variable proteins (C-terminus of VgrG, TAP, effector, and immunity), instead of comparing like to like (e.g., TAP proteins to other TAP proteins) [32]. Unterweger et al. classified Aux 1 and 2 effectors into three and five categories, respectively, and the large cluster VgrG into seven categories. Our analysis suggests there are two Aux 1 and four Aux 2 categories based on the predicted effector activity.

Discovery, characterization, and validation of novel T6SS gene loci

All clinical *V. cholerae* isolates sequenced to date contain the same three or four T6SS genomic loci (a large cluster and two or three auxiliary clusters), and the variability of effector sequences within clinical *V. cholerae* strains is limited. By contrast, the sequenced environmental strains described here and by Bernardy et al. contain a wider diversity of effector sequences in both auxiliary clusters [25]. HMMs based on degenerate *hcp* and *vgrG* genes revealed novel T6SS gene loci in environmental strains.

The Aux 4 cluster contains a canonical T6SS auxiliary cluster gene order and encodes a predicted effector (Tse4). A TMHMM prediction found no transmembrane helices and predicted the effector to be non-cytoplasmic. SWISS MODEL and Phyre2 do not predict any significant homology to known structures for Tse4, but I-TASSER suggests the effector could adopt a similar fold to pilin proteins found in *Streptococcus* species [35, 49, 50]. The cluster was found in other *V. cholerae* strains and is homologous to a cluster previously described [33]. The activity of Aux 4 is beyond the scope

of this study, but the ability of the cluster to be used in a T6SS-dependent manner and the biochemical function of the effector protein are currently being investigated.

The novel Aux 5 T6SS cluster present in two sequenced *V. cholerae* strains (BGT46 and BGT49) was identified using a DUF2235 HMM. The cluster is also found in 10 other *V. cholerae* strains and an *Aeromonas* strain (Fig. 2c, Additional file 2: Table S2). The 11 strains were isolated from diverse geographic locations over three decades and 2 *V. cholerae* strains are human isolates, unlike BGT46 and BGT49 (Additional file 2: Table S2). Short-read Illumina-based genome assembly of strain BGT49 was insufficient to resolve the gene order of the Aux 5 cluster. Subsequent sequencing of BGT49 using long-read PacBio technology confirmed the presence of *hcp*, *vgrG*, and *tap* open reading frames and confirmed this locus is not an assembly artifact.

In BGT49, sequences with high homology related to the Aux 5 cluster are found at four distinct genomic locations. The four Aux 5 genetic loci each encodes a predicted effector carrying a DUF2235 hydrolase domain, found in other lipases. The genetic organization of the novel cluster is different from other *V. cholerae* T6SS clusters. The Aux 5 clusters contain two putative immunity genes containing a DUF3304 domain upstream of the putative effector. A truncated vestigial gene with limited sequence homology to *tleV* genes is also observed between the two immunity genes at all four Aux 5 loci (Fig. 2a). Phyre2 and I-TASSER predict TleV1 is most similar to Tle1 from *P. aeruginosa*, suggesting TleV1 belongs to the Tle1 family of T6SS lipases [34–36, 49]. TleV1 and the other three TleV alleles lack a GX SXG-conserved catalytic motif associated with Tle1 lipases but contain a GXDLG motif [34].

Expression of TleV1 in the cytoplasm induced moderate toxicity in *E. coli* cells, but TleV1 was highly toxic when expressed in the *E. coli* periplasm, consistent with its designation as a Tle1-like lipase. This effect could be observed because TleV1 has a catalytic activity when present in both the cytoplasm and the periplasm. Alternatively, TleV1 could have a cryptic signal that exports the wild-type protein to the periplasm even in the absence of an exogenous signal, as proposed for other T6SS effectors [51]. Expression of the predicted immunity gene upstream of TleV1 was able to neutralize the toxicity of the effector in both *E. coli* and *V. cholerae* cells. Both the cytoplasmic and the periplasmic versions of TliV1a were able to rescue survival of both *E. coli* and *V. cholerae* cells. TliV1a is not predicted by SignalP-5.0 or PSORTb 3.0 to contain motifs for transport to the periplasm [52, 53]. It is possible that the immunity factor is not transported to the periplasm when Tat-tagged, or may also act in the periplasm, although its transport mechanism into that compartment remains unknown. A

second putative immunity protein found within the cluster, TliV1b, shares approximately 86% homology to TliV1a and contains a DUF3304 domain. It is also not predicted to encode a periplasmic signal sequence. However, unlike TliV1a, expression of TliV1b was not sufficient to rescue survival of *E. coli* expressing the TleV1 effector or the survival of *V. cholerae* cells when competed against C6706*:Aux 5 killer cells (Additional file 1: Figure S4). We hypothesize the truncated sequence found between *tliV1a* and *tliV1b* could be a vestigial remnant of an effector that was exchanged with the current *tliV1a-tleV1* immunity-effector module. TliV1b could have been an immunity protein specifically for the effector that is now lost. Alternatively, *tliV1b* could encode an immunity protein efficient at neutralizing effectors found in the other Aux 5 clusters.

Thomas et al. have previously shown experimentally that different effectors within *V. cholerae* auxiliary clusters can be swapped among strains [54]. Kirchberger et al. have also proposed that effector modules and *tap* genes can be swapped and acquired [55]. However, to our knowledge, this study is the first to experimentally show that an additional non-native T6SS auxiliary cluster can be acquired and used by a *V. cholerae* strain to kill kin cells lacking the immunity protein.

Conclusion

Competition within microbial communities is an important aspect of the life cycle of *V. cholerae* and other benign and pathogenic microbes. Twenty-six *V. cholerae* strains were sequenced, and hidden Markov models were used to probe for novel gene clusters associated with T6SS activity in *V. cholerae* isolates. Using the bioinformatics tool we developed, a novel cluster, named auxiliary cluster 5 (Aux 5), was discovered, and the effector encoded within the cluster was toxic when expressed in *E. coli* cells. The entire cluster was transferred into a different *V. cholerae* strain and conferred T6SS-dependent competitive advantage to the recipient strain. We propose that the tool we have developed is better suited than previous methods for discovering novel T6SS effectors in *V. cholerae* species and may be adapted in the future to facilitate the discovery of effectors in other bacterial species.

Methods

Vibrio cholerae culture conditions, DNA extraction, and sequencing

Strains were grown overnight in LB medium (Difco) at 37 °C, with shaking. Bacterial cells were pelleted by centrifugation and the supernatant discarded. Genomic DNA was isolated using ZR Fungal/Bacterial DNA MiniPrep kit (Zymo Research) and paired-end fragment libraries constructed using Nextera XT DNA Library Preparation Kit (Illumina) with a fragment length of 300 bp.

For PacBio sequencing, DNA from cultures of *V. cholerae* strain BGT49 was extracted using the PacBio phenol-chloroform recommended protocol and cleaned using AMPure XP beads (Beckman Coulter). Purified DNA was sent to the University of Washington PacBio Sequencing Services. Raw reads were trimmed and assembled using Canu, which is designed for long-read sequencing [56]. Resulting contigs were then scaffolded using SSPACE-LongRead, and read correction was performed with short-read data from Illumina sequencing using Pilon [57, 58].

Genome sequence analysis

Strain and assembly information are summarized in Additional file 2: Table S1.

Publicly available genome sequences

Completed and publicly available *Vibrio cholerae* genome sequences were downloaded from National Center for Biotechnology Information's (NCBI) RefSeq sequence collection and additional incomplete genomes, and sequence read archives were retrieved from NCBI's GenBank collection and Pathosystems Resource Integration Center (PATRIC) [59–61]. GenBank and RefSeq accessions are listed in Additional file 2: Table S1.

Whole-genome comparisons

RefSeq genomes and genomes from this study were subjected to an all-by-all nucleotide comparison using one-way, reciprocal best hits BLAST to calculate percent identity between 1024-bp blocks generated from each genome sequence [62]. The average nucleotide identity by BLAST (ANIb) was computed for each one-way, pairwise comparison, and the lower ANI value for each given pair was retained [39–41]. A 30 × 30 symmetric matrix of ANIb values was constructed and hierarchically clustered by complete linkage and heatmap generated in R using the ggplot2 package [63, 64].

Computational characterization of T6SS

Initial identification and annotation of large and auxiliary T6SS clusters were by BLAST against a database constructed using sequences reported previously by Unterweger et al. and Altindis et al. [23, 32]. VrgG-3 and VrgG-1 and 2 alleles served as markers for putative large and auxiliary clusters, respectively. BLAST hits to effector proteins were considered true positives if within three CDS of a predicted VgrG and in the same orientation. When VgrG proteins were identified with no neighboring effector annotation, genes at the +2 and +3 CDS with respect to VgrG protein were marked for manual validation. All BLAST-identified loci were manually validated, with new *vgrG* and effector alleles incorporated into the BLAST database. This iterative

method was applied until no additional clusters were found.

Putative effector functional annotations were assigned based on conserved functional domains. Reverse, position-specific BLAST (rpsBLAST) against the Protein Families (Pfam), Cluster of Orthologous Groups (COG), and Conserved Domain Database (CDD) databases was used to identify characteristic domains [65–67].

Hidden Markov models for effector prediction and annotation of new T6SS loci

HMMs were trained on manually curated alignments of *hcp*, *vgrG*, *tap*, effector, and immunity gene sequences for each cluster type using sequences from Unterweger et al. and the strains sequenced in this study. Two additional models were created and trained for both *hcp* and *vgrG*, using sequences from other bacterial genera and in silico mutated sequences, respectively. HMMs were validated by reannotating the genomes from the study.

T6SS Predictor

T6SS Predictor is a Shiny application built in R using custom Perl scripts to predict and annotate putative loci [63, 68]. T6SS Predictor takes as input either a protein FASTA file or genomic DNA FASTA file, with the option to provide a GFF annotation file instead of relying on de novo CDS prediction. Predictions are generated in 2–5 min, and the resulting output includes an annotated locus map of any identified loci and a FASTA file with putative T6SS proteins. T6SS Predictor is available from this project's homepage: <https://vibriocholera.com> and on Github [69]. T6SS Predictor is hosted on fault-tolerant hardware located in the USA and France and served using HTTPS best practices.

Bacterial strains

V. cholerae C6706 El Tor biotype O1 strain *qstR** constitutively expresses the T6SS apparatus genes, and C6706 *qstR** $\Delta vasK$ is deficient for T6SS apparatus function. Both strains were used for integration and natural transformation experiments. *E. coli* MG1655 with deleted arabinose metabolism genes *araBAD* was used for expression of TleV1 from an arabinose-inducible promoter. Genomic DNA from the environmental *V. cholerae* strain BGT49 (V56) was used for Illumina and PacBio sequencing and for amplifying the Aux 5 cluster by PCR. Details regarding the *V. cholerae* and *E. coli* strains used are provided in Additional file 2: Table S3.

Modified *V. cholerae* strains

All C6706 *V. cholerae* genetically modified strains (both insertions and deletions) were engineered using published allelic exchange techniques [70].

Recombinant DNA techniques

Primers used in PCR experiments were obtained from Eurofins Genomics. Phusion, Taq and Q5 Polymerases (Promega and New England Biolabs), and their respective buffers were used according to the manufacturer's instructions. DNA restriction nucleases were used to digest plasmids (Promega and New England Biolabs). Gibson assembly mixes were used according to the manufacturer's instructions to construct the plasmids used in this study (New England Biolabs). All recombinant strains and constructs used in the study were tested by colony PCR and verified for accuracy by Sanger sequencing.

E. coli toxicity experiments

E. coli strains expressing the *tleV1* gene under control of arabinose-inducible pBAD promoter were cultured in LB medium with 150 µg/mL of spectinomycin and 0.2% glucose overnight. Cells were then washed three times with LB and resuspended in fresh culture medium to an OD₆₀₀ of 0.5. To assess toxicity, tenfold serial dilutions were performed, and 3 µL aliquots of cell suspensions were then spotted on agar plates containing either spectinomycin and 0.2% glucose or containing spectinomycin and 0.2% arabinose. Agar plates were incubated statically overnight at 37 °C. The same growth conditions were used for *E. coli* strains expressing both *tleV1* and *tliV1a* genes with the exception that cells were grown overnight in LB medium with 150 µg/mL of spectinomycin, 10 µg/mL of chloramphenicol, and 0.2% glucose overnight and then were spotted on agar plates containing either spectinomycin, chloramphenicol, and 0.2% glucose or containing spectinomycin, chloramphenicol, and 0.2% arabinose.

T6SS killing assays

V. cholerae strains (both killer and target) were incubated overnight with shaking in liquid LB at 37 °C. Both strains were then washed three times with LB, diluted to an OD₆₀₀ of 1 in fresh LB, and then mixed together in a 10:1 (killer to target) ratio. Aliquots (50 µL) of the mixed cell suspension were spotted on filter paper with a 0.2-µm pore size that was placed on an LB plate and incubated at 37 °C for 3 h. Each filter paper was then vortexed for 30 s in 5 mL of LB. Resuspended cells were diluted and spread on plates containing antibiotic to select for surviving target cells. Plates were then incubated at 37 °C overnight, and the number of colonies was counted.

Confocal microscopy experiments

V. cholerae strains (both killer and target) were incubated overnight with shaking in liquid LB at 37 °C. Each overnight culture was back-diluted 1:100 and incubated with shaking at 37 °C for approximately 6 h. Cell suspensions were then normalized to an OD₆₀₀ of 1 in fresh LB and mixed in a 1:1 (killer to target) ratio. An 8-µL aliquot of propidium iodide

(100 µg/mL) was added to an agar pad and allowed to dry. Next, a 1-µL aliquot of the killer:target cell mixture was spotted. Cells were imaged at 37 °C and 96–100% humidity for 5 h using an Eclipse Ti-E Nikon inverted microscope. A Perfect Focus System was used with a × 40 objective (Plan Fluor ELWD × 40 DIC M N1) to stabilize the focus in the plane of the biofilm growth during long-term imaging. A Nikon A1plus camera was used to obtain images. The images were processed in ImageJ.

Natural transformation experiments

Natural transformation experiments were performed as described by Watve et al. [71]. Briefly, *V. cholerae* overnight cultures were diluted 1:100 in fresh LB medium and allowed to achieve an OD₆₀₀ of ~ 0.3. Two milliliters of each culture was then added to a sterile crab shell fragment and was incubated overnight at 30 °C in artificial sea water medium (17 g/L of Instant Ocean, cat. no. SS115-10). Genomic DNA from donor bacteria containing an antibiotic resistance gene was added, and the cells were incubated for 24 h. Cells were then spread on plates containing antibiotic to select for transformed cells.

Additional files

Additional file 1: Figure S1. Depiction of T6SS gene clusters in the BGT49 genome. **Figure S2.** TleV1 is predicted to share structural homology to the *P. aeruginosa* Tle1 T6SS lipase. **Figure S3.** BGT49 genome region containing the Aux 5a cluster. **Figure S4.** Expression of the *tliV1b* gene does not neutralize the toxicity of Tat-TleV1. **Figure S5.** Graphical description of the annotation workflow used in this study and, in part, by the T6SS Predictor. (PDF 1330 kb)

Additional file 2: Table S1. *V. cholerae* genomes analyzed in this study. **Table S2.** *V. cholerae* and *Aeromonas* strains encoding Aux 5-like clusters. **Table S3.** Genetically modified *E. coli* and *V. cholerae* strains and plasmids used in this study. (PDF 266 kb)

Additional file 3: Review history. (DOCX 31 kb)

Acknowledgements

We would like to thank all the current and past members of the Hammer Lab who have contributed to this paper with critiques and discussions, in particular, Jacob Thomas, Harshini Chandrashekar, and Siu Lung Ng.

Review history

The review history is available as Additional file 3.

Authors' contributions

CVC and KW contributed to the experimental design and all wet lab experiments. ATC, VR, LR, and SW contributed to the experimental design and all bioinformatics analyses. GS was responsible for the confocal microscopy experiments. BKH, IKJ, and PY were responsible for the experimental design, funding, advice, and critiques. All authors read and approved the final manuscript.

Funding

This research was supported in part by the NSF (MCB-1149925), the Georgia Tech School of Biological Sciences, Georgia Tech's Soft Matter Incubator, the Intramural Research Program of the NIH, NLM, NCBI, and the IHRC-Georgia Tech Applied Bioinformatics Laboratory. GS is funded by the German National Academy of Sciences Leopoldina (LPDS 2017-03).

Availability of data and materials

T6SS Predictor is available from this project's homepage: <https://vibriocholeera.com>. Genomic sequence data that support the findings of this study have been deposited at NCBI with the project accession code PRJNA339415 [72]. The source code for the predictive webserver and data that support these findings are freely available at Github under the MIT license [69] as well as in Zenodo [73]. <https://git.vcholeerae.com/arch/vibriocholeera.com>. Accession codes for all publicly available genomes analyzed in this study can be found in Additional file 2: Table S1. Data and material not provided in the text, such as DNA primers, are available upon request.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Center for Microbial Dynamics and Infection, Georgia Institute of Technology, 310 Ferst Drive, Atlanta, GA 30332-0230, USA. ²Institute for Bioengineering and Biosciences, Georgia Institute of Technology, 310 Ferst Drive, Atlanta, GA 30332-0230, USA. ³School of Biological Sciences, Georgia Institute of Technology, 310 Ferst Drive, Atlanta, GA 30332-0230, USA. ⁴Applied Bioinformatics Laboratory, Atlanta, GA, USA. ⁵PanAmerican Bioinformatics Institute, Cali, Valle del Cauca, Colombia. ⁶School of Physics, Georgia Institute of Technology, Atlanta, GA, USA. ⁷Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, MA, USA.

Received: 3 January 2019 Accepted: 18 July 2019

Published online: 12 August 2019

References

- CDC. Foodborne Diseases Active Surveillance Network (FoodNet): FoodNet Surveillance Report for 2014 (Final Report). 2014th. 2016. Atlanta Department of Health and Human Services, CDC.; 2014. Available from: <http://www.cdc.gov/foodnet/reports/annual-reports-2014.html>
- Legros D. Global cholera epidemiology: opportunities to reduce the burden of cholera by 2030. *J Infect Dis*. 2018;218(suppl_3):S137–40.
- Blakesch M. Competence-induced type VI secretion might foster intestinal colonization by *Vibrio cholerae*. *BioEssays*. 2015;37(11):1163–8 Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/bies.201500101>.
- Logan SL, Thomas J, Yan J, Baker RP, Shields DS, Xavier JB, et al. The *Vibrio cholerae* type VI secretion system can modulate host intestinal mechanics to displace gut bacterial symbionts. *Proc Natl Acad Sci*. 2018;115(16):201720133 Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1720133115>.
- Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, et al. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc Natl Acad Sci*. 2006;103(5):1528–33 Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1345711&tool=pmcentrez&rendertype=abstract>.
- Mougous JD. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science*. 2006;312(5779):1526–30 Available from: <http://www.sciencemag.org/cgi/doi/10.1126/science.1128393>.
- Durand E, Nguyen VS, Zoued A, Logger L, Péhau-Arnaudet G, Aschtgen M-S, et al. Biogenesis and structure of a type VI secretion membrane core complex. *Nature*. 2015;523(7562):555–60 Available from: <http://www.nature.com/doi/10.1038/nature14667>.
- Pukatzki S, Ma AT, Revel AT, Sturtevant D, Mekalanos JJ. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc Natl Acad Sci U S A*. 2007;104(39):15508–13.
- Aschtgen MS, Thomas MS, Cascales E. Anchoring the type VI secretion system to the peptidoglycan: TssL, TagL, TagP... what else? *Virulence*. 2010;1(6):535–40.
- Bönemann G, Pietrosiuk A, Diemand A, Zentgraf H, Mogk A. Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. *EMBO J*. 2009;28(4):315–25.
- Bröms JE, Ishikawa T, Wai SN, Sjöstedt A. A functional VipA-VipB interaction is required for the type VI secretion system activity of *Vibrio cholerae* O1 strain A1552. *BMC Microbiol*. 2013;13:96 Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3656785&tool=pmcentrez&rendertype=abstract>.
- Zheng J, Ho B, Mekalanos JJ. Genetic analysis of anti-amoebae and anti-bacterial activities of the type VI secretion system in *Vibrio cholerae*. *PLoS One*. 2011;6(8):e23876 Available from: <http://dx.doi.org/10.1371/journal.pone.0023876>.
- Ma LS, Lin JS, Lai EM. An IcmF family protein, ImpLM, is an integral inner membrane protein interacting with ImpKL, and its Walker A motif is required for type VI secretion system-mediated Hcp secretion in *Agrobacterium tumefaciens*. *J Bacteriol*. 2009;191(13):4316–29.
- Cascales E, Cambillau C. Structural biology of type VI secretion systems. *Philos Trans R Soc Lond B Biol Sci*. 2012;367(1592):1102–11 Available from: <http://rstb.royalsocietypublishing.org/content/367/1592/1102>.
- Unterwieser D, Kostjuk B, Otjengerdes R, Wilton A, Diaz-Satizabal L, Pukatzki S. Chimeric adaptor proteins translocate diverse type VI secretion system effectors in *Vibrio cholerae*. *EMBO J*. 2015;34(16):2198–210 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26194724>.
- Shneider MM, Buth SA, Ho BT, Basler M, Mekalanos JJ, Leiman PG. PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature*. 2013;500(7462):350–3 Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3792578/>.
- Ma J, Sun M, Dong W, Pan Z, Lu C, Yao H. PAAR-Rhs proteins harbor various C-terminal toxins to diversify the antibacterial pathways of type VI secretion systems. *Environ Microbiol*. 2017;19(1):345–60.
- Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? *BMC Genomics*. 2009;10:104 Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2660368&tool=pmcentrez&rendertype=abstract>.
- Durand E, Cambillau C, Cascales E, Journet L, VgrG, Tae, Tle, and beyond: the versatile arsenal of type VI secretion effectors. *Trends Microbiol*. 2014;22(9):498–507 Available from: <http://www.sciencedirect.com/science/article/pii/S0966842X14001358>.
- Russell AB, Hood RD, Bui NK, LeRoux M, Vollmer W, Mougous JD. Type VI secretion delivers bacteriolytic effectors to target cells. *Nature*. 2011;475(7356):343–7 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21776080%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3146020>.
- Zhang H, Gao ZQ, Wang WJ, Liu GF, Xu JH, Su XD, et al. Structure of the type VI effector-immunity complex (Tae4-Tai4) provides novel insights into the inhibition mechanism of the effector by its immunity protein. *J Biol Chem*. 2013;288(8):5928–39.
- Liang X, Moore R, Wilton M, Wong MJQ, Lam L, Dong TG. Identification of divergent type VI secretion effectors using a conserved chaperone domain. *Proc Natl Acad Sci U S A*. 2015;112(29):9106–11 Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4517263/>.
- Altindis E, Dong T, Catalano C, Mekalanos J. Secretome analysis of *Vibrio cholerae* type VI secretion system reveals a new effector-immunity pair. *MBio*. 2015;6(2):e00075 Available from: <http://mbio.asm.org/content/6/2/e00075-15.full>.
- Ishikawa T, Sabharwal D, Bröms J, Milton DL, Sjöstedt A, Uhlin BE, et al. Pathoadaptive conditional regulation of the type VI secretion system in *Vibrio cholerae* O1 strains. *Infect Immun*. 2012;80(2):575–84.
- Bernardy EE, Turnsek MA, Wilson SK, Tarr CL, Hammer BK. Diversity of clinical and environmental isolates of *Vibrio cholerae* in natural transformation and contact-dependent bacterial killing indicative of type VI secretion system activity. *Appl Environ Microbiol*. 2016;82(9):2833–42 Available from: <http://aem.asm.org/content/82/9/2833.abstract>.
- Unterwieser D, Kitaoka M, Miyata ST, Bachmann V, Brooks TM, Moloney J, et al. Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. *PLoS One*. 2012;7(10):e48320.
- Jaskólska M, Stutzmann S, Stoudmann C, Blakesch M. Qstr-dependent regulation of natural competence and type VI secretion in *Vibrio cholerae*. *Nucleic Acids Res*. 2018;46(20):10619–34 Available from: <https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gky717/5068245>.
- Antonova ES, Bernardy EE, Hammer BK. Natural competence in *Vibrio cholerae* is controlled by a nucleoside scavenging response that requires CytR-dependent anti-activation. *Mol Microbiol*. 2012;86(5):1215–31.
- Watte SS, Thomas J, Hammer BK. CytR is a global positive regulator of competence, type VI secretion, and chitinases in *Vibrio cholerae*. *PLoS One*. 2015;10(11):1–18.

30. Borgeaud S, Metzger LC, Scrinari T, Blokesch M. The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science*. 2015;347(6217):63–7 Available from: <https://doi.org/10.1126/science.1260064>.
31. Watve SS, Chande AT, Rishishwar L, Mariño-Ramírez L, Jordan IK, Hammer BK. Whole-genome sequences of 26 *Vibrio cholerae* isolates. *Genome Announc*. 2016;4(6):e01396–16.
32. Unterweger D, Miyata ST, Bachmann V, Brooks TM, Mullins T, Kostiuik B, et al. The *Vibrio cholerae* type VI secretion system employs diverse effector modules for intraspecific competition. *Nat Commun*. 2014;5:3549 Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3988814&tool=pmcentrez&rendertype=abstract>.
33. Labbate M, Orata FD, Petty NK, Jayatilake ND, King WL, Kirchberger PC, et al. A genomic island in *Vibrio cholerae* with VPI-1 site-specific recombination characteristics contains CRISPR-Cas and type VI secretion modules. *Sci Rep*. 2016;6:1–7 Available from: <https://doi.org/10.1038/srep36891>.
34. Russell AB, Leroux M, Hathazi K, Agnello DM, Ishikawa T, Wiggins PA, et al. Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature*. 2013;496(7446):508–12 Available from: <http://www.nature.com/doi/10.1038/nature12074>.
35. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10(6):845–58 Available from: <http://www.nature.com/articles/nprot.2015.053>.
36. H Hu, H Zhang, Z Gao, D Wang, G Liu, J Xu, K Lana and Y Dong. Structure of the type VI secretion phospholipase effector Tle1 provides insight into its hydrolysis and membrane targeting. *Acta Cryst*. 2014;D70:2175–85. Available from: <https://doi.org/10.1107/S1399004714012899>.
37. Palmer T, Berks BC. The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol*. 2012;10(7):483–96 Available from: <https://doi.org/10.1038/nrmicro2814>.
38. Suarez G, Sierra JC, Sha J, Wang S, Erova TE, Fadl AA, et al. Molecular characterization of a functional type VI secretion system from a clinical isolate of *Aeromonas hydrophila*. *Microb Pathog*. 2008;44(4):344–61.
39. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci*. 2009;106(45):19126–31 Available from: <http://www.pnas.org/content/106/45/19126.abstract>.
40. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol*. 2007;57(Pt 1):81–91 Available from: [syncii://DNA-DNA hybridization values a.pdf](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1929400).
41. Konstantinidis KT, Ramette A, Tiedje JM. The bacterial species definition in the genomic era. *Philos Trans R Soc London B Biol Sci*. 2006;361(1475):1929–40 Available from: <http://rspb.royalsocietypublishing.org/content/royptb/361/1475/1929.full.pdf>.
42. Ali A, Chen Y, Johnson JA, Redden E, Mayette Y, Rashid MH, et al. Recent clonal origin of cholera in Haiti. *Emerg Infect Dis*. 2011;17(4):699–701 Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3377427/>.
43. Bakhshi B, Mahmoodi-Aznavah A, Salimi-Khorashad A. Clonal dissemination of a single *Vibrio cholerae* O1 biotype El Tor strain in Sistan–Baluchestan province of Iran during 2013. *Curr Microbiol*. 2015;71(2):163–9 Available from: <https://doi.org/10.1007/s00284-015-0806-x>.
44. Paul FH, Deirdre C, Marinjho HJ, Alexander R, Samir RD, Rosheila D, et al. Clonal origins of *Vibrio cholerae* O1 El Tor Strains, Papua New Guinea, 2009–2011. *Emerg Infect Dis J*. 2011;17(11):2063 Available from: <http://wwwnc.cdc.gov/eid/article/17/11/11-0782>.
45. Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, Fiebig A, et al. Complete genome sequence of DSM 30083T, the type strain (U5/41T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci*. 2014;9(1):1–19 Available from: <https://doi.org/10.1186/1944-3277-9-2>.
46. Meibom KL, Blokesch M, Dolganov NA, Wu CY, Schoolnik GK. *Microbiology*. 2005;310(5755):1824–7.
47. Esteves K, Mosser T, Aujoulat F, Hervio-Heath D, Monfort P, Jumas-Bilak E. Highly diverse recombining populations of *Vibrio cholerae* and *Vibrio parahaemolyticus* in French Mediterranean coastal lagoons. *Front Microbiol*. 2015;6:708 Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4503927/>.
48. Li J, Yao Y, Xu HH, Hao L, Deng Z, Rajakumar K, et al. SecRE6: a web-based resource for type VI secretion systems found in bacteria. *Environ Microbiol*. 2015;17(7):2196–202 Available from: <https://doi.org/10.1111/1462-2920.12794>.
49. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods*. 2014;12(1):7–8 Available from: <http://www.nature.com/doi/10.1038/nmeth.3213>.
50. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46(W1):W296–303.
51. Ho BT, Fu Y, Dong TG, Mekalanos JJ. *Vibrio cholerae* type VI secretion system effector trafficking in target bacterial cells 2017;1–6.
52. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*. 2010;26(13):1608–15.
53. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*. 2019;37(4):420–3.
54. Thomas J, Watve SS, Ratcliff WC, Hammer BK. Horizontal gene transfer of functional type VI killing genes by natural transformation. *MBio*. 2017;8(4):1–11.
55. Kirchberger PC, Unterweger D, Provenzano D, Pukatzki S, Boucher Y. Sequential displacement of type VI secretion system effector genes leads to evolution of diverse immunity gene arrays in *Vibrio cholerae*. *Sci Rep*. 2017;7:45133 Available from: <http://www.nature.com/articles/srep45133>.
56. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*. 2017;27(5):722–36.
57. Boetzer M, Pirovano W. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics*. 2014;15(1):1–9.
58. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. 2014;9(11):112963.
59. Tatusova T, Ciufo S, Fedorov B, O'Neill K, Tolstoy I. RefSeq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Res*. 2014;42(database issue):D553–9 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24316578>.
60. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Ostell J, Pruitt KD, et al. GenBank. *Nucleic Acids Res*. 2018;46(D1):D41–7.
61. Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, et al. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res*. 2014;42(database issue):D581–91 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24225323>.
62. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389–402 <http://www.ncbi.nlm.nih.gov/pubmed/9254694>.
63. Team RC. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2018.
64. Wickham H. ggplot2. Vol. 9, International Journal of ChemTech Research. New York: Springer New York; 2009. p. 407–11.
65. Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res*. 2016;44(D1):D279–85 Available from: <http://nar.oxfordjournals.org/content/44/D1/D279.abstract>.
66. Marchler-Bauer A, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, et al. CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res*. 2003;31(1):383–7. <https://doi.org/10.1093/nar/gkg087>.
67. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res*. 2015;43(database issue):D222–6 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25414356>.
68. RStudio I. shiny: web application framework for R. R package; 2013.
69. Crisan CV, Chande AT, Williams K, Raghuram V, Rishishwar L, Steinbach G, Watve SS, Yunker P, Jordan IK, Hammer BK. T6SS Predictor. Github. <https://git.vcholerae.com/arch/vibriocholeera.com>. 2019.
70. Skorupski K, Taylor RK. Positive selection vectors for allelic exchange. *Gene*. 1996;169(1):47–52.
71. Watve SS, Bernardy EE, Hammer BK. *Vibrio cholerae*: measuring natural transformation frequency. *Curr Protoc Microbiol*. 2014;2014(April):6A.4.1–6A.4.12. <https://doi.org/10.1002/9780471729259.m06a04s35>

72. Chande A, Watve S, Rishishwar L, Bernardy E, Marino-Ramirez L, Jordan K, Hammer BK. *Vibrio cholerae* genome sequencing and assembly. NCBI BioProject PRJNA339415. <https://www.ncbi.nlm.nih.gov/bioproject/339415>.
73. Crisan CV, Chande AT, Williams K, Raghuram V, Rishishwar L, Steinbach G, Watve SS, Yunker P, Jordan IK, Hammer BK. T6SS Predictor. Zenodo. <https://zenodo.org/record/3273422>. 2019.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

