Multiple T6SSs, mobile auxiliary modules, and effectors revealed in a systematic analysis of the *Vibrio parahaemolyticus* pan-genome

Biswanath Jana[1], Kinga Keppel[1], Chaya Mushka Fridman[1], Eran Bosis[2,#], & Dor Salomon[1,#]

[1] Department of Clinical Microbiology and Immunology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[2] Department of Biotechnology Engineering, Braude College of Engineering, Karmiel, Israel

[#] Corresponding authors. Tel: +972 3 6408583; E-mail: dorsalomon@mail.tau.ac.il (DS); Tel: +972 4 9901927; E-mail: bosis@braude.ac.il (EB)

Running title: Pan-genome repertoire of *Vibrio parahaemolyticus* T6SS

Key words: T6SS, antibacterial, competition, secretion, Vibrio

Abstract

Type VI secretion systems (T6SSs) play a major role in interbacterial competition and in bacterial interactions with eukaryotic cells. The distribution of T6SSs and the effectors they secrete vary between strains of the same bacterial species. Therefore, a pan-genome investigation is required to better understand the T6SS potential of a bacterial species of interest. Here, we performed a comprehensive, systematic analysis of T6SS gene clusters and auxiliary modules found in the pan-genome of *Vibrio parahaemolyticus*, an emerging pathogen widespread in marine environments. We identified four different T6SS gene clusters within genomes of this species; two systems appear to be ancient and widespread, whereas the other two systems are rare and appear to have been more recently acquired via horizontal gene transfer. In addition, we identified diverse T6SS auxiliary modules containing putative effectors with either known or predicted toxin domains. Many auxiliary modules are possibly horizontally shared between *V. parahaemolyticus* genomes, since they are flanked by DNA mobility genes. We further investigated a DUF4225-containing protein encoded on an Hcp auxiliary module, and we showed that it is an antibacterial T6SS effector that exerts its toxicity in the bacterial periplasm, leading to cell lysis. Computational analyses of DUF4225 revealed a widespread toxin domain associated with various toxin delivery systems. Taken together, our findings reveal a diverse repertoire of T6SSs and auxiliary modules in the *V. parahaemolyticus* pan-genome, as well as novel T6SS effectors and toxin domains that can play a major role in the interactions of this species with other cells.

Importance

Gram-negative bacteria employ toxin delivery systems to mediate their interactions with neighboring cells. *Vibrio parahaemolyticus*, an emerging pathogen of humans and marine animals, was shown to deploy antibacterial toxins into competing bacteria via the type VI secretion system (T6SS). Here, we analyzed 1,727 *V. parahaemolyticus* genomes and revealed the pan-genome T6SS repertoire of this species, including the T6SS gene clusters, horizontally shared auxiliary modules, and toxins. We also identified a role for a previously uncharacterized domain, DUF4225, as a widespread antibacterial toxin associated with diverse toxin delivery systems.

Introduction

During competition, bacteria deliver toxic cocktails of effectors using specialized, contact-dependent protein secretion systems (1). Gram-negative bacteria often employ the type VI secretion system (T6SS) to gain a competitive advantage over their rivals (2). This system comprises 13-14 core components, in addition to accessory components that differ between systems and may play different regulatory roles (3–6). The structural and regulatory components of T6SS are encoded within large gene clusters that also often encode effectors; effectors are also encoded within auxiliary modules containing T6SS core components (7–10), or within orphan cassettes (11, 12). The effectors are loaded onto a missile-like structure comprising a tube consisting of stacked rings of Hcp hexamers; the tube is capped by a spike complex composed of a VgrG trimer sharpened by a PAAR repeat-containing protein (hereafter referred to as PAAR) (13). The loaded missile is propelled outside of the cell by a contracting sheath that engulfs it (14); the contraction provides sufficient force to penetrate into a neighboring cell and deploy the effectors (5).

Many effector families that contain toxin domains that mediate antibacterial activities have been reported. Effectors often target conserved bacterial components in the cytoplasm or periplasm, such as nucleic acids (nucleases) (7, 15–19), the membrane (phospholipases and pore-forming toxins) (10, 11, 20–22), or the peptidoglycan layer (amidases, glycoside hydrolases, and carboxy- and transpeptidases) (12, 23, 24). Additional activities mediated by T6SS effectors include the following: altering the energy balance by hydrolyzing $NAD(P)^+$ (25, 26), ADP-ribosylating the conserved protein FtsZ to inhibit cell division (27), ADP-ribosylating the 23S rRNA to inhibit translation (28), targeting the transamidosome to inhibit protein synthesis (29), deaminating the target cell's DNA (30), and synthesizing the toxic molecule (p)ppApp (31). Antibacterial T6SS effectors are encoded adjacent to cognate immunity proteins that protect against self- or kin-intoxication (2, 32). Several experimental and computational approaches have been used to identify effector and immunity (E/I) pairs (2, 7, 11, 12, 29, 33–36). Nevertheless, because T6SS effectors employ diverse mechanisms for secretion (13, 37–39), and therefore lack a canonical secretion signal, it is estimated that many more effectors still await discovery.

Vibrio parahaemolyticus is an emerging pathogen that inhabits marine and estuarine environments (40). Previous works revealed that all investigated *V. parahaemolyticus* isolates contain a T6SS on chromosome 2, named T6SS2, whereas pathogenic isolates encode another T6SS on chromosome 1, named T6SS1 (41–43). The presence of additional T6SSs in this species remains to be investigated. T6SS2 was recently shown to mediate antibacterial activities; however, its effector repertoire remains unstudied (11, 44). T6SS1, which has been studied in several isolates (7, 11, 36, 42, 43), also mediates antibacterial activities. Each isolate contains conserved antibacterial E/I pairs, which are encoded by the main T6SS1 gene cluster, as well as "accessory" E/I pairs that differ between isolates and diversify the effector repertoire (36, 45). To date, accessory T6SS1 E/I pairs were found in auxiliary T6SS modules containing a gene encoding VgrG (7), or as orphan operons that often reside next to DNA mobility elements (11, 36, 42). Notably, since most vibrios are naturally competent (46, 47), horizontal gene transfer (HGT) may play a role in the acquisition and evolution of the T6SS E/I pair repertoire (48, 49).

In this study, we sought to reveal the collective repertoire of T6SS gene clusters and auxiliary modules in *V. parahaemolyticus*, as well as to identify new effectors. By systematically analyzing 1,727 *V. parahaemolyticus* genomes, we identified four types of T6SS gene clusters and many distinct, widespread auxiliary modules predicted to encode diverse effectors; the vast majority of the auxiliary modules, as well as two of the T6SS gene clusters, are found next to DNA mobility genes, suggesting that they were acquired via HGT. Intriguingly, most *hcp*-containing auxiliary modules encode a previously undescribed effector with a C-terminal domain of unknown function

4225 (DUF4225). We experimentally showed that this effector is toxic upon delivery to the bacterial periplasm, where it leads to cell lysis. We also identified a downstream-encoded cognate immunity protein that antagonizes the effector's toxic effect. Surprisingly, although several strains of marine bacteria were intoxicated by this effector during competition, others, including two *V. parahaemolyticus* strains that do not contain homologs of the cognate immunity protein, were resistant to the attack. Further analysis revealed that DUF4225 is a widespread toxin domain that is present in polymorphic toxins associated with several protein secretion systems.

Results

Four T6SS gene clusters are found in the *V. parahaemolyticus* pan-genome

We first set out to identify the T6SS gene clusters found in the *V. parahaemolyticus* pan-genome. To this end, we retrieved the sequences of the conserved T6SS core component, TssB (3), from 1,727 available RefSeq *V. parahaemolyticus* genomes (**Supplementary Dataset S1**). Analysis of the phylogenetic distribution of TssB revealed four groups (**Fig. 1A**) corresponding to four distinct T6SS gene clusters in *V. parahaemolyticus* genomes (**Fig. 1B** and **Supplementary Dataset S2,S3**).

T6SS1 and T6SS2 were previously investigated and found to mediate interbacterial competition (7, 11, 36, 43, 44). In agreement with previous analyses on smaller genome datasets (7, 11, 43), we identified T6SS2 in nearly all *V. parahaemolyticus* genomes (99%), whereas T6SS1 was identified in 68.3% of the genomes (**Fig. 1b**, **Fig. 2**, and **Supplementary Dataset S2,S3**). We did not identify known or potential effectors encoded within the T6SS2 gene clusters; however, the T6SS1 clusters contain two antibacterial effectors, corresponding to VP1390 and VP1415 in the type strain RIMD 2210633 (36, 37) (**Table 1**). Interestingly, we observed some diversity at the end of T6SS1 gene clusters; we found what appears to be duplications (between one and six copies) of the PAAR-containing specialized effector, corresponding to VP1415 in the type strain RIMD 2210633 (36).

Two additional T6SS gene clusters, which we named T6SS3 and T6SS4 (**Fig. 1b**), have a limited distribution in *V. parahaemolyticus* genomes (0.8% and 1.8%, respectively) (**Fig. 2**, and **Supplementary Dataset S2,S3**). T6SS3 is similar to the previously reported T6SS3 of *V. proteolyticus*, which was suggested to have anti-eukaryotic activity (50). T6SS4 could be further divided into three sub-groups, a-c, with minor differences in gene sequence and content (**Fig. 1b** and **Supplementary Dataset S3**). We identified potential effectors encoded in both the T6SS3 and T6SS4 gene clusters (**Fig. 1b** and **Table 1**). Notably, T6SS3 and T6SS4, which were not previously described in *V. parahaemolyticus*, are flanked by transposases and other DNA mobility elements (**Fig. 1**), suggesting that they have been acquired via HGT. Taken together, these results reveal that the *V. parahaemolyticus* pan-genome contains two widespread T6SSs and two T6SSs with limited distribution.

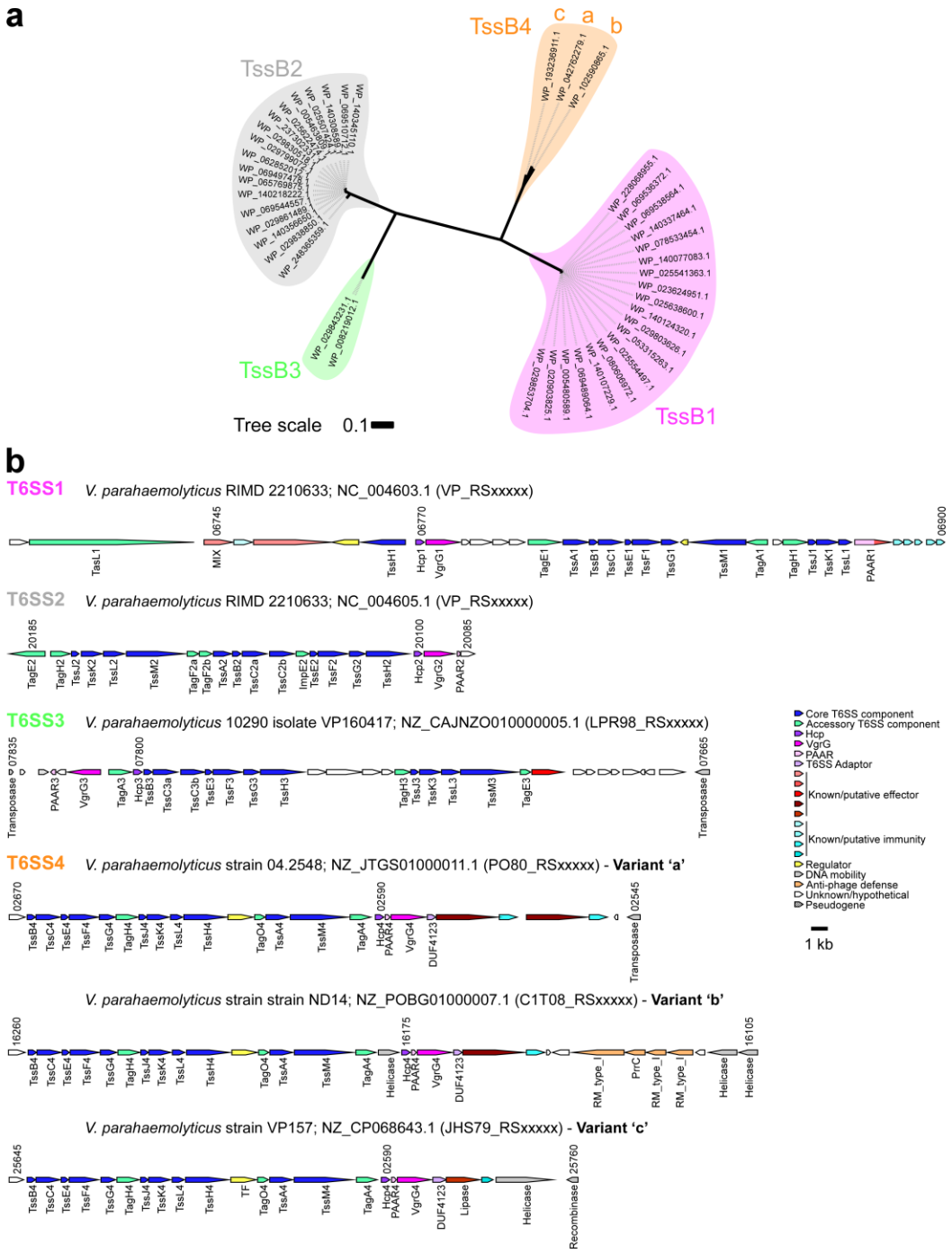

Figure 1. The *V. parahaemolyticus* pan-genome harbors four T6SSs. (a) Phylogenetic distribution of the T6SS core component TssB encoded within *V. parahaemolyticus* genomes. The evolutionary history was inferred using the neighbor-joining method. The phylogenetic tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method; they are in the units of the number of amino acid substitutions per site. **(b)** Representative T6SS gene clusters found in *V. parahaemolyticus* genomes. The strain names, the GenBank accession numbers, and the locus tag annotation patterns are provided. Genes are denoted by arrows indicating the direction of transcription. Locus tags are denoted above, and the names of encoded proteins or domains are denoted below.

128

Diverse and widespread T6SS auxiliary modules contain effectors

Although T6SS effectors are often encoded within the main T6SS gene clusters, auxiliary modules containing at least one secreted T6SS component (i.e., Hcp, VgrG, or PAAR) and downstream-encoded effectors are also common (7–10). Therefore, to identify auxiliary T6SS modules, we searched the *V. parahaemolyticus* pan-genome for Hcp, VgrG, and PAAR encoded outside the four main T6SS gene clusters described above. We found diverse auxiliary module types widely distributed among the different *V. parahaemolyticus* genomes (**Fig. 2**, **Supplementary Fig. S1**, and **Supplementary Dataset S2,S4**). These modules are predominantly found next to DNA mobility genes, such as integrases, recombinases, transposases, phage proteins, or plasmid mobility elements. Notably, some modules contain more than one secreted T6SS component, and some genomes harbor multiple module types (up to seven modules in one genome).

Figure 2. Distribution of T6SS gene clusters and auxiliary modules in *V. parahaemolyticus*. The phylogenetic tree was based on DNA sequences of *rpoB* coding for DNA-directed RNA polymerase subunit beta. The evolutionary history was inferred using the neighbor-joining method. The height of the blue bars denoted the number of T6SS auxiliary modules per genome.

PAAR proteins encoded within auxiliary modules are often specialized effectors containing known (e.g., AHH, Ntox15, and NUC nucleases) or predicted C-terminal toxin domains (**Table 1**), followed by a downstream gene that possibly encodes a cognate immunity protein (**Supplementary Fig. S1** and **Supplementary Dataset S4**). Homologous PAAR proteins are also encoded within similar auxiliary module configurations in which the toxin domain is encoded by a separate gene as a possible cargo effector (**Table 1** and **Supplementary Dataset S4**). In

contrast, we did not identify auxiliary VgrG proteins containing C-terminal toxin domains; instead, auxiliary VgrG modules carry cargo effectors, some with known activities (e.g., PoNe DNase, NucA/B nuclease, and Lip2 lipase). These cargo effectors are mostly encoded downstream of a gene encoding a T6SS adaptor protein, such as DUF4123, DUF2169, and DUF1795 (39, 51) (**Supplementary Fig. S1** and **Supplementary Dataset S4)**.

Hcp-encoding auxiliary modules contain a few, previously uncharacterized, putative cargo effectors of unknown function (**Table 1**, **Supplementary Fig. S1** and **Supplementary Dataset S4**). Interestingly, in most Hcp-containing auxiliary modules that we found, which are located in diverse syntenies, a DUF4225-containing protein is encoded immediately downstream of *hcp* (**Supplementary Fig. S2**). Some DUF4225-encoding genes have an adjacent, small gene encoding a protein predicted to contain transmembrane helices (according to a Phobius server analysis (52)) (**Supplementary Dataset S4**). Based on these findings, we hypothesized that these DUF4225-encoding genes and their downstream adjacent genes are antibacterial T6SS E/I pairs.

Table 1. Predicted effectors in *V. parahaemolyticus* T6SS clusters and auxiliary modules.

Example accession number	Example gene locus	Found in T6SS or module type	Effector type	Predicted toxin domain	Predicted activity	Homologs in other polymorphic toxin classes	Ref.
WP_005480617.1	VP_RS06755 (VP1390)	T6SS1	Cargo	Unknown	Cell lysis	No	(37)
WP_005493834.1	VP_RS06875 (VP1415)	T6SS1, PAAR	Specialized / Cargo	AHH	Nuclease	Yes	(36, 53)
WP_029843206.1	LPR98_RS07710	T6SS3	Cargo	Unknown	Unknown	Yes	
WP_042762256.1	PO80_RS02570	T6SS4a,b	Cargo	Unknown	Unknown	No	
WP_193237005.1	JHS79_RS25745	T6SS4c, Hcp+VgrG	Cargo	Lip2	Lipase	Yes	
WP_065771704.1	AKH09_RS09365	PAAR+VgrG+Hcp	Cargo	Unknown	Unknown	Yes	
WP_238790300.1	K6U37_RS14065	PAAR+VgrG	Cargo	NucA/B	Nuclease	Yes	(53)
WP_020841305.1	H9J98_RS02420	PAAR	Specialized / Cargo	Ntox15	Nuclease	Yes	(15, 53)
WP_083135234.1	GPY55_RS17385	PAAR	Specialized / Cargo	NUC	Nuclease	Yes	(53)
WP_102591288.1	C1T08_RS26340	PAAR	Specialized	Unknown	Unknown	Yes	
WP_102591220.1	C1S85_RS24675	PAAR	Specialized	Tme	Membrane-disrupting	Yes	(11)
WP_102591225.1	C1S85_RS24700	PAAR	Specialized	Unknown	Unknown	Yes	
WP_238789479.1	K6U37_RS04455	PAAR	Specialized (truncated)	Unknown	Unknown	Yes	
WP_129147717.1	EGL73_RS17180	VgrG	Cargo	Unknown	Unknown	No	(54)
WP_029857615.1	B5C30_RS14465	VgrG	Cargo	PoNe	DNase	Yes	(7)
WP_238790289.1	K6U37_RS13990	VgrG	Cargo	Unknown	Unknown	Yes	(45)
WP_086585359.1	JHS88_RS14235	Hcp	Cargo	DUF4225	Cell lysis	Yes	This work
WP_195433156.1	K6U37_RS12660	Hcp	Cargo	Unknown	Unknown	Yes	
WP_228085946.1	IB292_RS21975	Hcp	Cargo	Unknown	Unknown	Yes	

Constructing an "effectorless" surrogate T6SS platform

A surrogate T6SS platform can be used as a tool to study putative E/I pairs. The major advantage of the surrogate platform is that it only requires constructing a single plasmid to express the putative E/I pair in question (11). Since we have previously reported that the T6SS1 of POR1, a *V. parahaemolyticus* strain RIMD 2210633 derivative, can be used as a surrogate platform to deliver and investigate effectors and modules belonging to T6SS1 from other *V. parahaemolyticus*

strains (11), we decided to employ this strategy to investigate the putative DUF4225-containing effector.

A drawback of our previously reported surrogate platform was the presence of the endogenous T6SS1 effectors of the RIMD 2210633 strain (36), which prevented the use of a possibly sensitive prey strain and thus required the use of a RIMD 2210633-derived strain containing the cognate immunity proteins as prey during competition assays. To enable the use of a surrogate platform during competition against diverse prey strains, we set out to construct an "effectorless" version. To this end, we deleted the genes encoding the reported effector VPA1263 (36) and the co-effector VP1388 (37), and we replaced two residues in the predicted active site of the specialized effector VP1415 (36) with alanine, as previously reported (55). In addition, we deleted *vp1133*, which encodes an H-NS protein that represses T6SS1 activity in *V. parahaemolyticus* (56), to constitutively activate T6SS1 in the surrogate strain. The resulting platform, which we named VpT6SS1^{Surrogate}, is active at 30°C in media containing 3% (wt/vol) NaCl, as evident by the expression and secretion of the hallmark VgrG1 protein in a T6SS1-dependent manner (**Supplementary Fig. S3A**). Furthermore, VpT6SS1^{Surrogate} enables interbacterial killing of susceptible *V. natriegens* prey during competition mediated by a plasmid-borne VgrG1b auxiliary module belonging to T6SS1 of *V. parahaemolyticus* strain 12-297/B, containing the PoNe DNase effector (7) (**Supplementary Fig. S3B**).

A DUF4225-containing protein is a T6SS1 antibacterial effector

To determine whether DUF4225-encoding genes and their adjacent downstream genes are antibacterial T6SS E/I pairs, we chose to investigate an Hcp auxiliary module from *V. parahaemolyticus* strain CFSAN018764 (**Supplementary Fig. S2**; RefSeq sequence NZ_LHBG01000025.1) encoding an Hcp and a DUF4225-containing protein, hereafter referred to as DUF4225^{18764} (accession numbers WP_065788327.1 and WP_065788326.1, respectively). This strain harbors both T6SS1 and T6SS2 (**Supplementary Dataset S2**). Since the module's Hcp is more similar to Hcp1 than to Hcp2 (**Supplementary Fig. S4** and **Supplementary Dataset S5**), we reasoned that this Hcp auxiliary module is probably associated with T6SS1; we therefore named this Hcp as Hcp1b.

The Hcp1b module lacked an annotated gene immediately downstream of the gene encoding the putative effector DUF4225^{18764}. Nevertheless, when we performed a manual analysis of the nucleotide sequence, we identified an open reading frame immediately downstream of the putative effector (positions 296 to 27 in NZ_LHBG01000025.1) (**Supplementary Fig. S2**). The identified gene encodes an 89 amino acid-long protein containing three predicted transmembrane helices. Therefore, we predicted that this gene encodes a cognate DUF4225^{18764} immunity protein, and we named it Imm4225^{18764}.

Using VpT6SS1^{Surrogate}, we next set out to investigate the ability of the Hcp1b auxiliary module from strain CFSAN018764 to mediate T6SS1-dependent competition. As shown in **Figure 3**, an arabinose-inducible plasmid encoding the three proteins of the Hcp1b module (pModule; i.e., Hcp1b, DUF4225^{18764}, and Imm4225^{18764}), but not Hcp1b alone (pHcp1b), mediated the T6SS1-dependent intoxication of *V. natriegens* prey. Expression of Imm4225^{18764} from a plasmid (pImm) rescued the *V. natriegens* prey strain from this attack. Taken together, these results indicate that the Hcp1b auxiliary module of *V. parahaemolyticus* strain CFSAN018764 carries a T6SS1 antibacterial E/I pair.

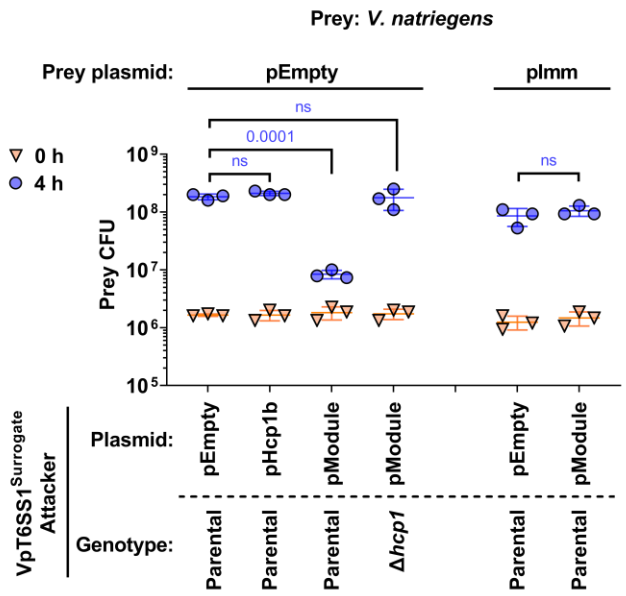

Figure 3. A DUF4225-containing Hcp auxiliary module contains a T6SS1 effector and immunity pair. Viability counts (CFU) of *V. natriegens* prey strains containing an empty plasmid (pEmpty) or a plasmid for the arabinose-inducible expression of Imm4225^{18764} (pImm) before (0 h) and after (4 h) co-incubation with the surrogate T6SS1 platform strain (VpT6SS1^{Surrogate}) or its T6SS1^{-} derivative (Δ*hcp1*) carrying an empty plasmid (pEmpty) or a plasmid for the arabinose-inducible expression of Hcp1b (pHcp1b) or the three-gene Hcp1b module (pModule) from *V. parahaemolyticus* strain CFSAN018764. The statistical significance between samples at the 4 h timepoint was calculated using an unpaired, two-tailed Student's *t*-test; ns, no significant difference ($p > 0.05$). Data are shown as the mean ± SD; n = 3.

A DUF4225-containing effector leads to cell lysis upon delivery to the periplasm

Next, we investigated the toxicity mediated by DUF4225^{18764}. The arabinose-inducible expression of the effector in the periplasm of *E. coli* (by fusion to an N-terminal PelB signal peptide), but not in the cytoplasm, led to a clear reduction in the optical density (OD_{600}) of the bacterial culture over time (**Fig. 4A**); the phenotype was similar to the lytic effect of the amidase effector Tse1 from *Pseudomonas aeruginosa* (57), and dissimilar to the effect of the membrane-disrupting effector Tme1 from *V. parahaemolyticus* (11). Lysis was also observed when *E. coli* cells expressing the periplasmic version of DUF4225^{18764} were visualized under a fluorescence microscope. Approximately 90 minutes after inducing DUF4225^{18764} expression, cells began to shrink and appeared to have lost their cytoplasmic content. Concomitantly, the leakage of DNA from these cells became apparent, manifesting as the fluorescence of propidium iodide (PI), a non-permeable DNA dye that was added to the media, around them (**Fig. 4B** and **Supplementary Movie S1**). These phenotypes are characteristic of cell lysis (58). Sometimes, mostly at later stages of the time course, we observed cells stained from the inside by PI, indicative of the slow permeabilization of the membranes. Similar phenotypes were not seen in *E. coli* cells containing an empty expression plasmid (**Supplementary Fig. S5** and **Supplementary Movie S1**). Taken together, our results suggest that the activity of DUF4225^{18764} in the bacterial periplasm can lead to cell lysis. Importantly, the co-expression of Imm4225^{18764} in *E. coli* rescued the cells from the toxicity mediated by periplasmic DUF4225^{18764} (**Fig. 4C**), further supporting its role as the cognate immunity protein of this effector. The expression of the non-toxic cytoplasmic version of DUF4225^{18764} was detected in immunoblots, whereas the toxic periplasmic version of DUF4225^{18764} was only detected when Imm4225^{18764} was co-expressed (**Supplementary Fig. S6**).

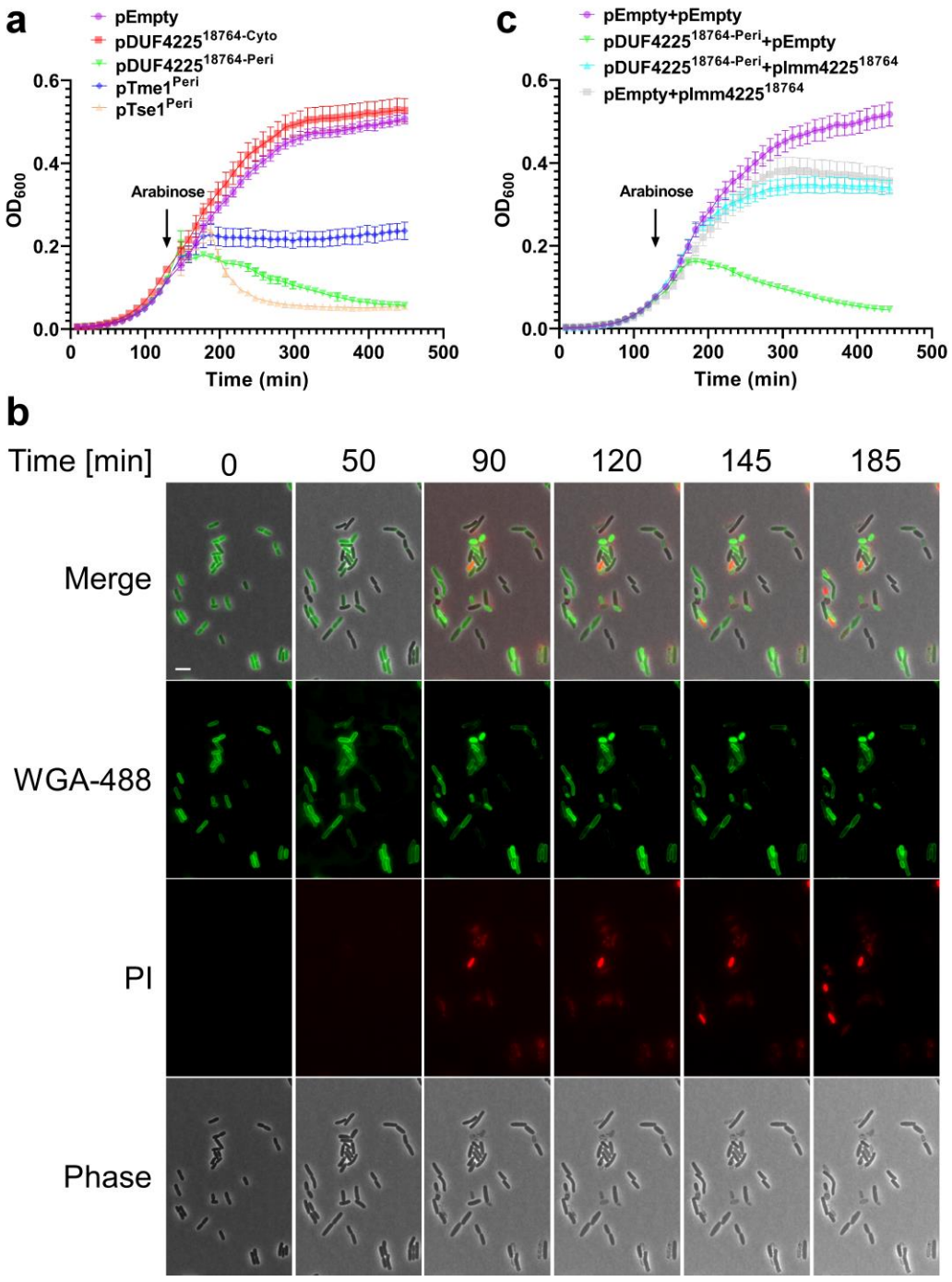

Figure 4. DUF4225^{18764} induces cell lysis upon delivery to the periplasm. (a,c) Growth of *E. coli* BL21 (DE3) containing arabinose-inducible plasmids, either empty (pEmpty) or expressing the indicated proteins in the cytoplasm (Cyto) or periplasm (Peri; fused to an N-terminal PelB signal peptide). In 'c', bacteria contain a second plasmid, either empty (pEmpty) or encoding Imm4225^{18764}. An arrow denotes the timepoint at which arabinose (0.05%) was added to the media. **(b)** Time-lapse microscopy of *E. coli* MG1655 cells stained with Wheat Germ Agglutinin Alexa Fluor 488 conjugate (WGA-488) and that express a periplasmic DUF4225^{18764} from an arabinose-inducible plasmid, grown on agarose pads supplemented with chloramphenicol (to maintain the plasmid) and 0.2% arabinose (to induce expression), and propidium iodide (PI). WGA-488 (green), PI (red), phase contrast and merged channels are shown. Size bar = 5 µm.

237

Vibrio strains without a cognate immunity protein can resist DUF4225^{18764} toxicity

We and others previously reported that certain T6SS effectors appear to have a specific toxicity range; they can intoxicate some but not all bacterial prey strains during T6SS-mediated attacks (55, 59). Since some resistant strains do not carry a homolog of the intoxicating effector's cognate immunity protein, it had been suggested that non-immunity protein-mediated defense mechanisms play a role in such resistance (60). In light of these recent observations, we set out to examine the toxicity range of a T6SS attack mediated by DUF4225^{18764}. To this end, we monitored the viability of several marine bacteria prey strains during competition against the VpT6SS1^{Surrogate} platform delivering DUF4225^{18764}. For prey strains with antibacterial T6SSs known or predicted to be active under the tested conditions, we used mutants in which the T6SS was inactivated by deleting a T6SS core component (e.g., *hcp* or *tssB*), as indicated, to avoid counterattacks during competition. Interestingly, whereas *V. campbellii* ATCC 25920, *V. coralliilyticus* ATCC BAA-450, and *Aeromonas jandaei* DSM 7311 were susceptible to a DUF4225^{18764}-mediated attack, *V. parahaemolyticus* 12-297/B was only mildly susceptible to the attack and *V. vulnificus* CMCP6 and *V. parahaemolyticus* RIMD 2210633 were not affected by it (**Fig. 5**). The three latter strains do not carry a homolog of Imm4225^{18764}. Importantly, all prey strains except for *V. parahaemolyticus* 12-297/B and *Aeromonas jandaei* DSM 7311, which contain a PoNi immunity protein, were susceptible to intoxication by a PoNe DNase effector from *V. parahaemolyticus* strain 12-297/B (7), when it was delivered by the VpT6SS1^{Surrogate} platform (**Fig. 5**); this result indicates that the platform can deliver effectors into the tested prey strains. Therefore, our results reveal that certain bacteria can resist intoxication by DUF4225^{18764} even in the absence of a cognate immunity protein.

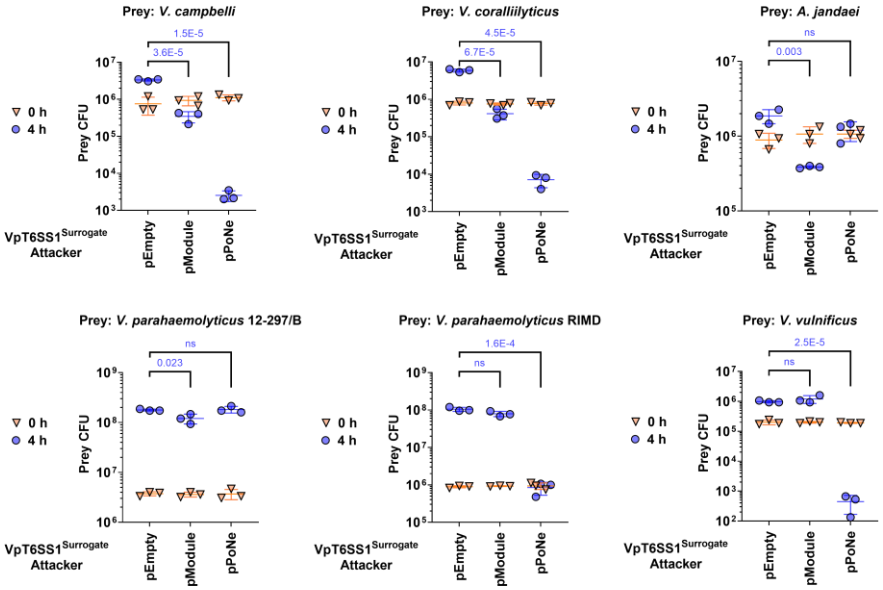

Figure 5. Varying sensitivity of marine bacteria to DUF4225^{18764} T6SS-mediated attacks. Viability counts (CFU) of the indicated prey strains before (0 h) and after (4 h) co-incubation with the surrogate T6SS1 platform strain (VpT6SS1^{Surrogate}) carrying an empty plasmid (pEmpty) or a plasmid for the arabinose-inducible expression of the three-gene Hcp1b module (pModule) from *V. parahaemolyticus* strain CFSAN018764 or the PoNe DNase-containing VgrG1b module from *V. parahaemolyticus* 12-297/B (pPoNe). Statistical significance between samples at the 4 h timepoint was calculated using an unpaired, two-tailed Student's *t*-test; ns, no significant difference ($P > 0.05$). Data are shown as the mean ± SD; n=3. The prey strains used were *V. campbellii* ATCC 25920 Δ*hcp1*, *V. coralliilyticus* ATCC BAA-450 Δ*hcp1*, *Aeromonas jandaei* DSM 7311 Δ*tssB*, *V. parahaemolyticus* 12-297/B Δ*hcp1*, *V. parahaemolyticus* RIMD 2210633 Δ*hcp1*, and *V. vulnificus* CMCP6.

DUF4225 is a widespread toxin domain associated with diverse secretion systems

We next investigated the distribution of the DUF4225 toxin domain in bacterial genomes. We found that the homologs of DUF4225^{18764} are widespread in bacterial genomes, almost exclusively belonging to the Pseudomonadota (formerly, Proteobacteria) phylum (**Fig. 6A** and **Supplementary Dataset S6**). Interestingly, DUF4225 is found in polymorphic toxins containing N-terminal domains associated with T6SS (e.g., PAAR, VgrG, and Hcp), type V secretion system (e.g., Fil_haemagg_2 and DUF637 (61)), and others (e.g., SpvB (53), RHS_repeat (16), and Sec system signal peptides), although most of the homologs do not contain an identifiable N-terminal domain fused to DUF4225 (**Fig. 6B** and **Supplementary Dataset S6**). Sometimes the association of the DUF4225-containing protein with a specific secretion system can be inferred from adjacent genes encoding known components of protein secretion systems, as is the case with DUF4225^{18764}, which is encoded downstream of *hcp1b*. Taken together, these results reveal that DUF4225 is a toxin domain that is widespread in secreted polymorphic effectors.

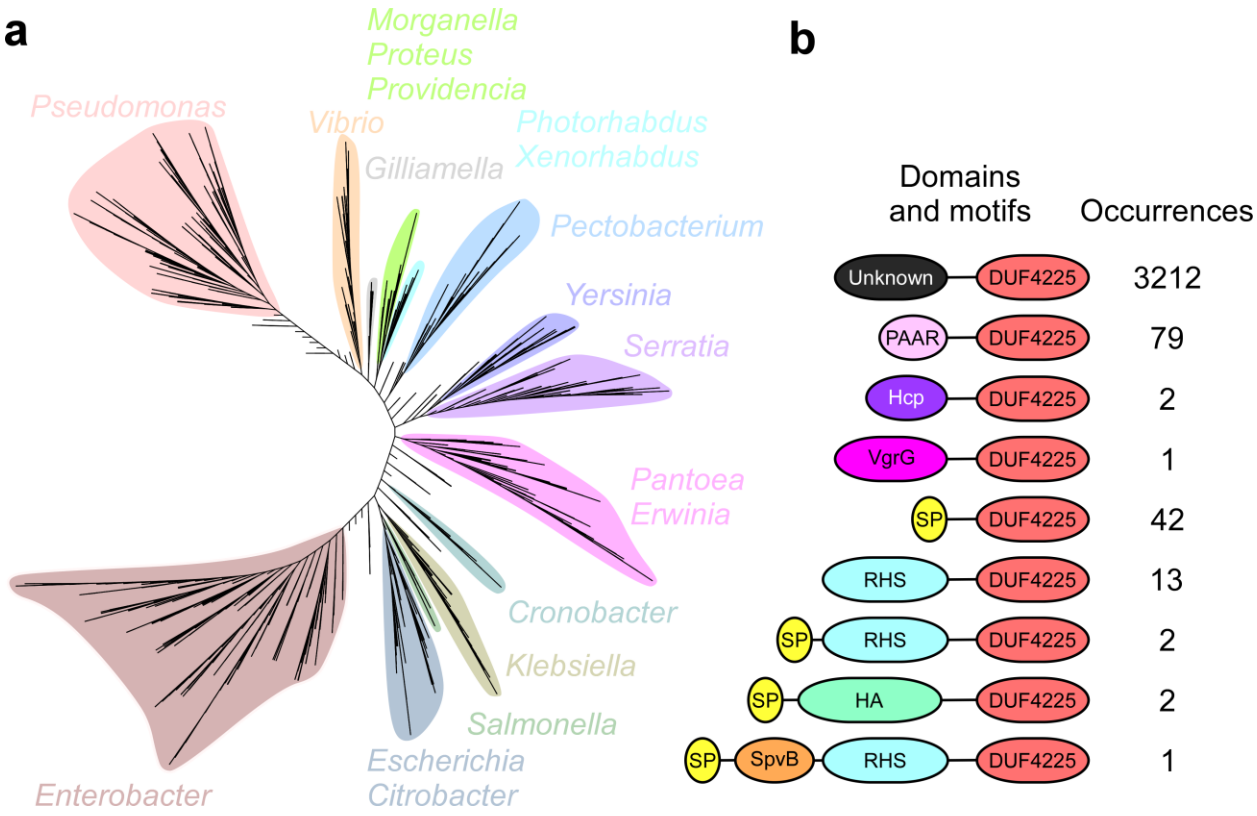

Figure 6. DUF4225 is a widespread toxin domain. (a) Phylogenetic distribution of bacteria encoding a DUF4225 homolog was based on the DNA sequences of *rpoB* coding for DNA-directed RNA polymerase subunit beta. The evolutionary history was inferred using the neighbor-joining method. **(b)** The domain architecture and the number of occurrences of DUF4225-containing proteins. SP, signal peptide; HA, Haemagg_act + Fil_haemagg_2 + DUF637. Domain sizes are not to scale.

Discussion

V. parahaemolyticus is an emerging marine pathogen responsible for gastroenteritis in humans (40) and for the economically devastating acute hepatopancreatic necrosis disease in shrimp (62). Like many other vibrios, this species employs T6SSs to manipulate and outcompete its rivals (7,

11, 43, 45). Here, we performed a systematic analysis of all available RefSeq *V. parahaemolyticus* genomes, and we revealed the pan-genome repertoire of T6SS gene clusters, auxiliary modules, and the effectors therein.

We identified four T6SS gene cluster types in *V. parahaemolyticus* genomes. Two systems appear to be ancient and widespread, and two seem to have been more recently acquired. T6SS1 and T6SS2, which were previously shown to mediate antibacterial activities (11, 43), are the most common systems in this species. T6SS2 is omnipresent; therefore, it is probably an ancient system that plays an important role in the *V. parahaemolyticus* life cycle. T6SS1, which is present in 68.3% of the genomes, is possibly also an ancient system; however, since T6SS1 appears to be absent from certain lineages (**Fig. 2**), we propose that it had been lost several times during the evolution of this species. This may be because T6SS1 serves a specialized purpose that is not beneficial for some strains. Indeed, we and others previously proposed that T6SS1 is present predominantly in pathogenic isolates of *V. parahaemolyticus* (41–43); therefore, it is plausible that non-pathogenic strains have lost this system, since it may only be required during colonization of a host. T6SS3 and T6SS4 are found in <3% of the genomes combined. Since DNA mobility genes flank these gene clusters, we propose that they have been recently acquired via HGT. This seems possible in light of the ability of vibrios to horizontally acquire large DNA fragments (46, 63, 64). The role and activity mediated by these two systems remains to be investigated. Nevertheless, we propose that T6SS3 mediates anti-eukaryotic activity, since it is similar to the previously reported T6SS3 in *V. proteolyticus* (50), and since we identified a putative effector within the cluster that lacks an identifiable potential immunity protein, suggesting that it does not mediate antibacterial activity. We also propose that T6SS4 mediates antibacterial activities, since we identified putative antibacterial effectors within it, which have a downstream adjacent gene that possibly encodes a cognate immunity protein (**Fig. 1b** and **Table 1**).

Our analysis also revealed diverse T6SS auxiliary modules encoding at least one of the secreted proteins, Hcp, VgrG, and PAAR. In most of these auxiliary modules, we found known or putative effectors that can diversify the toxic repertoires of this species' T6SSs. The identity of these genes as effectors is supported by their homology to known effectors, by their position downstream of genes encoding T6SS adaptors, by the presence of domains associated with T6SS-secreted proteins (e.g., MIX (36) and FIX (7)), and by the presence of homologs of their C-terminus in polymorphic toxins (**Table 1**). Interestingly, several effectors, identified within auxiliary modules, do not resemble previously studied toxins, and they may therefore employ novel mechanisms of action. Notably, although this was not directly addressed in this work, additional orphan effectors have been previously reported in *V. parahaemolyticus* genomes (7, 11, 42, 45); many are found next to DNA mobility genes, suggesting that they may also be horizontally shared within this species (42, 48). These orphan effectors further diversify the T6SS effector repertoire of this species.

In this work, we investigated the role of a gene encoding a DUF4225-containing protein; we showed that it is an antibacterial T6SS1 effector, and that its cognate immunity protein is encoded directly downstream. To the best of our knowledge, this is the first activity described for this domain of unknown function. Notably, DUF4225-encoding genes are common in Hcp-containing auxiliary modules, suggesting that they too are T6SS effectors. Our analysis also revealed that DUF4225 is widespread outside vibrios, where it is associated with T6SSs, as well as with other secretion systems that deliver polymorphic toxins.

We found that DUF4225 exerts its toxic activity in the bacterial periplasm. However, its mechanism of action and its cellular target remain unknown. Since its expression in the periplasm of *E. coli* cells led to cell lysis, we hypothesize that it affects the stability of the peptidoglycan or the membrane. Future biochemical and structural work will be required to address these open questions. Interestingly, we identified marine bacteria that resist intoxication by a DUF4225-

containing effector. Since we did not identify homologs of the cognate immunity protein encoded within the genomes of these resistant strains, we predict that they employ a yet-unknown non-immunity protein-mediated defense mechanism that counteracts the toxicity of this effector.

The effectorless surrogate T6SS platform constructed in this work is an important tool, allowing us to rapidly identify and investigate T6SS effectors encoded by any *V. parahaemolyticus* strain. Although we previously reported the construction of a similar surrogate system (11), the effectorless version reported here is superior: 1) the toxicity mediated by a putative effector can be tested against diverse prey strains, thus reducing the possibility of false negatives due to the lack of toxicity against a *V. parahaemolyticus* RIMD 2210633 prey (in hindsight, DUF4225^{18764} would have been a false negative in the previous version of the surrogate platform); and 2) no endogenous effectors compete with the investigated putative effector for loading onto the secreted tube and spike, thus increasing the probability that it will be delivered by the surrogate system.

In conclusion, we present the first comprehensive analysis of the T6SS repertoire in the *V. parahaemolyticus* pan-genome. Our results reveal four T6SSs found within this species; they also indicate that mobile auxiliary modules probably contribute greatly to diversifying the T6SS effector repertoires in various strains. We also describe a role for the widespread DUF4225 as an antibacterial toxin domain, and we identify additional putative effectors that await investigation.

Materials and Methods

Strains and Media: For a complete list of strains used in this study, see **Supplementary Table S1**. *E. coli* strains were grown in 2xYT broth (1.6% wt/vol tryptone, 1% wt/vol yeast extract, and 0.5% wt/vol NaCl) or on Lysogeny Broth (LB) agar plates containing 1% wt/vol NaCl at 37°C, or at 30°C when harboring effector expression plasmids. The media were supplemented with chloramphenicol (10 µg/mL) or kanamycin (30 µg/mL) to maintain plasmids, and with 0.4% wt/vol glucose to repress protein expression from the arabinose-inducible promoter, P*bad*. To induce expression from P*bad*, L-arabinose was added to the media at 0.1-0.2% (wt/vol), as indicated.

Vibrio parahaemolyticus, *V. natriegens, V. coralliilyticus*, and *V. vulnificus* strains were grown in MLB media (LB media containing 3% wt/vol NaCl) or on marine minimal media (MMM) agar plates (1.5% wt/vol agar, 2% wt/vol NaCl, 0.4% wt/vol galactose, 5 mM $MgSO_4$, 7 mM K_2SO_4, 77 mM K_2HPO_4, 35 mM KH_2PO_4, and 2 mM NH_4Cl). *V. campbellii* were grown in MLB media and on MLB agar plates. *Aeromonas jandaei* were grown in LB media and on LB agar plates. When vibrios or *A. jandaei* contained a plasmid, the media were supplemented with kanamycin (250 µg/mL), chloramphenicol (10 µg/mL), or streptomycin (100 µg/mL) to maintain the plasmid. Bacteria were grown at 30°C. To induce expression from P*bad*, L-arabinose was added to media at 0.05% (wt/vol).

Plasmid construction: For a complete list of plasmids used in this study, see **Supplementary Table S2**. The DNA sequence of the Hcp1b auxiliary module from *V. parahaemolyticus* strain CFSAN018764 (positions 296 to 27 in NZ_LHBG01000025.1) was synthesized by Twist Bioscience (USA). The entire module sequence or the sequences of genes within it were PCR amplified and then inserted into the multiple cloning site (MCS) of pBAD^K/Myc-His, pPER5, or pBAD33.1^F, in-frame with C-terminal Myc-His or FLAG tags, using the Gibson assembly method (65). For the expression of Imm4225^{18764} in *V. natriegens*, the region spanning the *araC* gene to the rrnT1 terminator was amplified from pBAD33.1^F containing the gene in its MCS, and then inserted into the *Not*I restriction site of pCLTR plasmid using restriction-digestion and ligation.

Plasmids were introduced into *E. coli* using electroporation. Transformants were grown on agar plates supplemented with 0.4% wt/vol glucose to repress expression from the P*bad* promotor

during the subcloning steps. Plasmids were introduced into vibrios and *A. jandaei* via conjugation. Trans-conjugants were grown on MLB agar plates for *V. campbellii*, LB agar plates for *A. jandaei*, or MMM agar plates for all other vibrios. Plates were supplemented with appropriate antibiotics to maintain the plasmids.

Construction of deletion strains: For in-frame deletions of *hcp1* in *V. coralliilyticus* (VIC_RS16330) or of *hcp1* in *V. campbellii* (A8140_RS17660), 1 kb sequences upstream and downstream of each gene were subcloned into pDM4, a Cm^{r}OriR6K suicide plasmid (66). Next, pDM4 constructs were introduced into the respective *Vibrio* strain via conjugation. Trans-conjugants were selected on MMM agar plates containing chloramphenicol (10 µg/mL). The resulting trans-conjugants were grown on MMM agar plates containing sucrose (15% wt/vol) for counter-selection and loss of the SacB-containing pDM4.

The VpT6SS1^{Surrogate} strain and its Δ*hcp1* derivative were generated by consecutive deletions or mutations of the relevant genes using previously reported pDM4 plasmids.

Toxicity assays in *E. coli*: To assess the toxicity mediated by DUF4225^{18764}, pBAD^K/Myc-His (for cytoplasmic expression) and pPER5 (for periplasmic expression fused to an N-terminal PelB signal peptide) plasmids, either empty or encoding DUF4225^{18764} were transformed into *E. coli* BL21 (DE3). *E. coli* transformants were grown overnight in 2xYT media supplemented with kanamycin (30 µg/mL) under P*bad* repressing conditions (0.4% wt/vol glucose). Overnight cultures were washed to remove residual glucose, and normalized to OD_{600} = 0.01 in 2xYT media supplemented with kanamycin. Then, 200 µL of each bacterial culture were transferred into 96-well plates in quadruplicate. The cultures were grown at 37 °C with agitation (205 cpm) in a microplate reader (BioTek SYNERGY H1). After 2 h of growth, L-arabinose was added to each well at a final concentration of 0.1% (wt/vol), to induce protein expression. OD_{600} readings were recorded every 10 min for 7 h.

To test the ability of Imm4225^{18764} to antagonize the toxicity of DUF4225^{18764}, a pBAD33.1^F plasmid, either empty or encoding Imm4225^{18764}, was co-transformed with a pPER5 plasmid, either empty or encoding DUF4225^{18764}, into *E. coli* BL21 (DE3). The growth of these strains was determined as described above. Growth assays were performed at least four times with similar results. Results from a representative experiment are shown.

Protein expression in *E. coli*: To determine the expression of C-terminal Myc-His-tagged DUF4225^{18764}, *E. coli* BL21 (DE3) bacteria carrying a single arabinose-inducible expression plasmid, either empty or encoding a cytoplasmic or a periplasmic DUF4225^{18764}, or bacteria carrying two plasmids, one for expression of Imm4225^{18764} and the other either empty or expressing the periplasmic version of DUF4225^{18764}, were grown overnight in 2xYT media supplemented with the appropriate antibiotics to maintain plasmids, and glucose to repress expression from *Pbad*. The cultures were washed twice with fresh 2xYT medium to remove residual glucose, and then diluted 100-fold in 5 mL of fresh 2xYT medium supplemented with appropriate antibiotics and grown for 2 h at 37°C. To induce protein expression, 0.1% (wt/vol) L-arabinose was added to the media. After 4 additional hours of growth at 30°C, 1.0 OD_{600} units of cells were pelleted and resuspended in 100 µL of (2X) Tris-Glycine SDS sample buffer (Novex, Life Sciences). Samples were boiled for 5 min, and cell lysates were resolved on Mini-PROTEAN TGX Stain-Free™ precast gels (Bio-Rad). For immunoblotting, α-Myc antibodies (Santa Cruz Biotechnology, 9E10, mouse mAb; sc-40) were used at 1:1,000 dilution. Protein signals were visualized in a Fusion FX6 imaging system (Vilber Lourmat) using enhanced chemiluminescence (ECL) reagents. Experiments were performed at least three times with similar results; the results from representative experiments are shown.

Bacterial competition assays: Bacterial competition assays were performed as previously described (37), with minor changes. Briefly, cultures of the indicated attacker and prey strains

were grown overnight. Bacterial cultures were then normalized to OD_{600} = 0.5 and mixed at a 10:1 (attacker: prey) ratio. The mixtures were spotted (25 µL) on MLB agar plates supplemented with 0.05% (wt/vol) L-arabinose, and incubated for 4 h at 30°C. Colony forming units (CFU) of the prey strains were determined at the 0 and 4-hour timepoints. The experiments were performed at least three times with similar results. Results from a representative experiment are shown.

Fluorescence microscopy: Cell morphology and membrane permeability during the expression of $DUF4225^{18764}$ in *E. coli* was assessed as previously described (37). Briefly, overnight-grown *E. coli* MG1655 cells carrying a pPER5 plasmid, either empty or encoding $DUF4225^{18764}$, were diluted 100-fold into 5 mL of fresh LB media supplemented with kanamycin and 0.2% (wt/vol) glucose. Bacterial cultures were grown for 2 h at 37°C, and then cells were washed with 0.15 M NaCl to remove residual glucose. Bacterial cultures were normalized to OD_{600} = 0.5 in 0.15 M NaCl solution. To visualize the cell wall of *E. coli*, 20 µL of bacterial cultures were incubated with Wheat Germ Agglutinin Alexa Fluor 488 Conjugate (Biotium; Catalogue no. 29022–1) at a final concentration of 0.1 mg/mL, and incubated for 10 min at room temperature (RT). Next, 1 µL was spotted on LB agarose pads (1% wt/vol agarose supplemented with 0.2% wt/vol L-arabinose) onto which 1 µL of the membrane-impermeable DNA dye, propidium iodide (PI; 1 mg/mL; Sigma-Aldrich) had been pre-applied. After the spots had dried (1–2 min at RT), the agarose pads were mounted, facing down, on 35 mm glass bottom CELLview™ cell culture dishes (Greiner). Cells were then imaged every 5 min for 4 h under a fluorescence microscope, as detailed below. The stage chamber (Okolab) temperature was set to 37°C. Bacteria were imaged in a Nikon Eclipse Ti2-E inverted motorized microscope equipped with a CFI PLAN apochromat DM 100X oil lambda PH-3 (NA, 1.45) objective lens, a Lumencor SOLA SE II 395 light source, and ET-EGFP (#49002, used to visualize the Alexa Fluor 488 signal), and an RFP filter cube (#49005, used to visualize the PI signal) filter sets. Images were acquired using a DS-QI2 Mono cooled digital microscope camera (16 MP) and were post-processed using Fiji ImageJ suite. The experiments were performed three times. Results from a representative experiment are shown.

VgrG1 secretion assays: *V. parahaemolyticus* $VpT6SS1^{Surrogate}$ and its Δ*hcp1* derivative strain were grown overnight at 30°C in MLB media. Bacterial cultures were normalized to OD_{600} = 0.18 in 5 mL of MLB media, and after 5 h of incubation at 30°C with agitation (220 rpm), expression fractions (cells) and secretion fractions (media) were collected and processed as previously described (37).

Identifying T6SS gene clusters in *V. parahaemolyticus*: A local database containing the RefSeq bacterial nucleotide and protein sequences was generated (last updated on June 11, 2022). *V. parahaemolyticus* genomes were retrieved from the local database and OrthoANI was performed as described previously (11, 67). Two genomes (assembly accessions GCF_000591535.1 and GCF_003337295.1) with OrthoANI values <95% were removed from the dataset.

The presence of T6SS gene clusters in *V. parahaemolyticus* genomes was determined by following the two-step procedure described below. In the first step, BLASTN (68) was employed to align *V. parahaemolyticus* nucleotide sequences against the nucleotide sequences of representative T6SS clusters (**Figure 1** and **Supplementary Dataset S3**). The expect value threshold was set to 10^{-12} and the minimal alignment length was 500 bp. The results were then sorted by their nucleotide accession numbers and bit score values (from largest to smallest), and the best alignments for each nucleotide accession number were saved. This step resulted in a list of *V. parahaemolyticus* nucleotide accession numbers and their best alignments to the representative T6SS gene clusters, including the positions of the alignments. In the second step, a two-dimensional matrix was generated for each T6SS gene cluster in which rows represented the *V. parahaemolyticus* genomes and columns represented the coordinates of the specific T6SS gene cluster. The matrices were then filled in with the percent identity values, based on the

positions of the alignments. Finally, the overall coverage was calculated for each T6SS gene cluster in each genome. *V. parahaemolyticus* genomes with at least 70% overall coverage of T6SS gene cluster were regarded as containing that T6SS gene cluster (**Supplementary Dataset S3**).

Identifying T6SS auxiliary modules: RPS-BLAST (69) was employed to identify proteins containing Hcp (COG3157), PAAR (DUF4150, PAAR_motif, PAAR_1, PAAR_2, PAAR_3, PAAR_4, PAAR_5, PAAR_RHS, PAAR_CT_1, PAAR_CT_2), and VgrG (COG3501) domains that were retrieved from the Conserved Domain Database (70), in *V. parahaemolyticus* genomes. Protein accessions located at the ends of contigs were removed. T6SS auxiliary modules were manually identified, based on the distance from the T6SS gene clusters, the genomic architecture, and the conserved domains in neighboring genes (**Supplementary Dataset S4**).

Identifying DUF4225 homologs with domain and neighborhood analysis: The Position-Specific Scoring Matrix (PSSM) of the DUF4225 domain was reconstructed using amino acids 105-243 of DUF4225^{18764} from *V. parahaemolyticus* strain CFSAN018764 (WP_065788326.1). Five iterations of PSI-BLAST were performed against the reference protein database (a maximum of 500 hits with an expect value threshold of 10^{-6} and a query coverage of 70% were used in each iteration). RPS-BLAST was then performed to identify DUF4225-containing proteins. The results were filtered using an expect value threshold of 10^{-8} and a minimal coverage of 70%. The genomic neighborhoods of DUF4225-containing proteins were analyzed as described previously (11, 45). Duplicated protein accessions appearing in the same genome in more than one genomic accession were removed if the same downstream protein existed at the same distance (**Supplementary Dataset S6**).

Constructing phylogenetic trees: The nucleotide sequences of *rpoB* coding for DNA-directed RNA polymerase subunit beta were retrieved from the local RefSeq database. Partial and pseudogene sequences were not included in the analyses. In the case of bacterial genomes encoding DUF4225 homologs, the *rpoB* sequences were first clustered using CD-HIT to remove identical sequences (100% identity threshold). Phylogenetic analyses of bacterial genomes were conducted using the MAFFT 7 server (mafft.cbrc.jp/alignment/server/). The *rpoB* sequences were aligned using MAFFT v7 FFT-NS-2 (71, 72). In the case of *V. parahaemolyticus* genomes, the evolutionary history was inferred using the neighbor-joining method (73) with the Jukes-Cantor substitution model (JC69). The analysis included 1,694 nucleotide sequences and 3,964 conserved sites. In the case of bacterial genomes encoding DUF4225, the evolutionary history was inferred using the average linkage (UPGMA) method and included 2,816 nucleotide sequences.

The protein accessions of TssB and Hcp from *V. parahaemolyticus* genomes were retrieved and unique sequences were aligned using CLUSTAL Omega (74). The evolutionary history was inferred using the Neighbor-Joining method (73). The analysis of TssB included 42 amino acid sequences and 166 conserved sites. The analysis of Hcp included 27 amino acid sequences and 133 conserved sites. Evolutionary analyses in both cases were conducted in MEGA X (75).

Identifying effectors in T6SS gene clusters and auxiliary modules: The presence of effectors in T6SS gene clusters and auxiliary modules was determined by homology to previously studied effectors in *V. parahaemolyticus*, by the location within auxiliary modules downstream of secreted core components (i.e., Hcp, VgrG, or PAAR) or of known T6SS adaptor-encoding genes (i.e., DUF4123, DUF1795, or DUF2169), and by the presence of potential C-terminal toxin domains identified using NCBI's Conserved Domain Database (76). The presence of homologs of C-terminal toxin domains in other polymorphic toxin classes was determined using Jackhmmer (77).

Acknowledgments

This project received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement no. 714224 to D Salomon), and from the Israel Science Foundation (grant no. 920/17 to D Salomon, and grant no. 1362/21 to D Salomon and E Bosis). CM Fridman was supported by scholarships from the Clore Israel Foundation and from the Manna Center Program in Food Safety and Security at Tel Aviv University, as well as by a scholarship for outstanding doctoral students from the Orthodox community from the Council for Higher Education. We thank Hila Saar, Nikol Dvir, and Nely Altshul for their excellent technical assistance, and members of the Salomon and Bosis laboratories for helpful discussions and suggestions.

Author Contributions

B Jana: conceptualization, investigation, methodology, and writing—review and editing.

K Keppel: investigation and methodology

CM Fridman: investigation and methodology

E Bosis: conceptualization, funding acquisition, investigation, methodology, and writing—— original draft.

D Salomon: conceptualization, supervision, funding acquisition, investigation, methodology, and writing—original draft.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Klein TA, Ahmad S, Whitney JC. 2020. Contact-dependent interbacterial antagonism mediated by protein secretion machines. Trends Microbiol. Elsevier Ltd https://doi.org/10.1016/j.tim.2020.01.003.

2. Hood RD, Singh P, Hsu FS, Güvener T, Carl MA, Trinidad RRS, Silverman JM, Ohlson BB, Hicks KG, Plemel RL, Li M, Schwarz S, Wang WY, Merz AJ, Goodlett DR, Mougous JD. 2010. A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. Cell Host Microbe 7:25–37.

3. Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I. 2009. 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 10:104.

4. Bingle LE, Bailey CM, Pallen MJ. 2008. Type VI secretion: a beginner's guide. Curr Opin Microbiol 11:3–8.

5. Wang J, Brackmann M, Castaño-Díez D, Kudryashev M, Goldie KN, Maier T, Stahlberg H, Basler M. 2017. Cryo-EM structure of the extended type VI secretion system sheath-tube complex. Nat Microbiol 2:1507–1512.

6. Shneider MM, Buth SA, Ho BT, Basler M, Mekalanos JJ, Leiman PG. 2013. PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. Nature 500:350–353.

7. Jana B, Fridman CM, Bosis E, Salomon D. 2019. A modular effector with a DNase domain and a marker for T6SS substrates. Nat Commun 10:3595.

8. Crisan C V., Chande AT, Williams K, Raghuram V, Rishishwar L, Steinbach G, Watve SS, Yunker P, Jordan IK, Hammer BK. 2019. Analysis of Vibrio cholerae genomes identifies new type VI secretion system gene clusters. Genome Biol 20.

9. Hachani A, Allsopp LP, Oduko Y, Filloux A. 2014. The VgrG proteins are “à la carte” delivery systems for bacterial type VI effectors. J Biol Chem 289:17872–84.

10. Russell AB, Leroux M, Hathazi K, Agnello DM, Ishikawa T, Wiggins PA, Wai SN, Mougous JD. 2013. Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. Nature 496:508–512.

11. Fridman CM, Keppel K, Gerlic M, Bosis E, Salomon D. 2020. A comparative genomics methodology reveals a widespread family of membrane-disrupting T6SS effectors. Nat Commun 11:1085.

12. Russell AB, Singh P, Brittnacher M, Bui NK, Hood RD, Carl MA, Agnello DM, Schwarz S, Goodlett DR, Vollmer W, Mougous JD. 2012. A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach. Cell Host Microbe 11:538–549.

13. Jana B, Salomon D. 2019. Type VI secretion system: a modular toolkit for bacterial dominance. Future Microbiol 14:fmb-2019-0194.

14. Basler M, Pilhofer M, Henderson GP, Jensen GJ, Mekalanos JJ. 2012. Type VI secretion requires a dynamic contractile phage tail-like structure. Nature 483:182–6.

15. Ma LS, Hachani A, Lin JS, Filloux A, Lai EM. 2014. Agrobacterium tumefaciens deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. Cell Host Microbe 16:94–104.

16. Koskiniemi S, Lamoureux JG, Nikolakakis KC, t’Kint de Roodenbeke C, Kaplan MD, Low DA, Hayes CS. 2013. Rhs proteins from diverse bacteria mediate intercellular competition. Proc Natl Acad Sci U S A 110:7032–7.

17. Pissaridou P, Allsopp LP, Wettstadt S, Howard SA, Mavridou DAI, Filloux A. 2018. The Pseudomonas aeruginosa T6SS-VgrG1b spike is topped by a PAAR protein eliciting DNA damage to bacterial competitors. Proc Natl Acad Sci U S A 115:12519–12524.

18. Alcoforado Diniz J, Coulthurst SJ. 2015. Intraspecies competition in Serratia marcescens is mediated by type VI-secreted Rhs effectors and a conserved effector-associated accessory protein. J Bacteriol 197:2350–60.

19. Ma J, Pan Z, Huang J, Sun M, Lu C, Yao H. 2017. The Hcp proteins fused with diverse extended-toxin domains represent a novel pattern of antibacterial effectors in type VI secretion systems. Virulence 8:1189–1202.

20. LaCourse KD, Peterson SB, Kulasekara HD, Radey MC, Kim J, Mougous JD. 2018. Conditional toxicity and synergy drive diversity among antibacterial effectors. Nat Microbiol 3:440–446.

21. Mariano G, Trunk K, Williams DJ, Monlezun L, Strahl H, Pitt SJ, Coulthurst SJ. 2019. A family of Type VI secretion system effector proteins that form ion-selective pores. Nat Commun 10:1–15.

22. Miyata ST, Unterweger D, Rudko SP, Pukatzki S. 2013. Dual expression profile of type VI secretion system immunity genes protects pandemic Vibrio cholerae. PLoS Pathog 9:1–18.

604 23. Whitney JC, Chou S, Russell AB, Biboy J, Gardiner TE, Ferrin MA, Brittnacher M,
605 Vollmer W, Mougous JD. 2013. Identification, structure, and function of a novel type VI
606 secretion peptidoglycan glycoside hydrolase effector-immunity pair. J Biol Chem
607 288:26616–24.

608 24. Sibinelli-Sousa S, Hespanhol JT, Nicastro GG, Matsuyama BY, Mesnage S, Patel A, de
609 Souza RF, Guzzo CR, Bayer-Santos E. 2020. A family of T6SS antibacterial effectors
610 related to l,d-transpeptidases targets the peptidoglycan. Cell Rep 31:107813.

611 25. Whitney JC, Quentin D, Sawai S, LeRoux M, Harding BN, Ledvina HE, Tran BQ,
612 Robinson H, Goo YA, Goodlett DR, Raunser S, Mougous JD. 2015. An interbacterial
613 NAD(P)+ glycohydrolase toxin requires elongation factor Tu for delivery to target cells.
614 Cell 163:607–619.

615 26. Tang JY, Bullen NP, Ahmad S, Whitney JC. 2018. Diverse NADase effector families
616 mediate interbacterial antagonism via the type VI secretion system. J Biol Chem
617 293:1504–1514.

618 27. Ting S-YY, Bosch DE, Mangiameli SM, Radey MC, Huang S, Park Y-JJ, Kelly KA, Filip
619 SK, Goo YA, Eng JK, Allaire M, Veesler D, Wiggins PA, Peterson SB, Mougous JD.
620 2018. Bifunctional immunity proteins protect bacteria against FtsZ-targeting ADP-
621 Ribosylating toxins. Cell 175:1380-1392.e14.

622 28. Jurėnas D, Payelleville A, Roghanian M, Turnbull KJ, Givaudan A, Brillard J, Hauryliuk V,
623 Cascales E. 2021. Photorhabdus antibacterial Rhs polymorphic toxin inhibits translation
624 through ADP-ribosylation of 23S ribosomal RNA. Nucleic Acids Res 49.

625 29. Nolan LM, Cain AK, Clamens T, Furniss RCD, Manoli E, Sainz-Polo MA, Dougan G,
626 Albesa-Jové D, Parkhill J, Mavridou DAI, Filloux A. 2021. Identification of Tse8 as a type
627 VI secretion system toxin from Pseudomonas aeruginosa that targets the bacterial
628 transamidosome to inhibit protein synthesis in prey cells. Nat Microbiol 6.

629 30. de Moraes MH, Hsu F, Huang D, Bosch DE, Zeng J, Radey MC, Simon N, Ledvina HE,
630 Frick JP, Wiggins PA, Peterson SB, Mougous JD. 2021. An interbacterial dna deaminase
631 toxin directly mutagenizes surviving target populations. Elife 10.

632 31. Ahmad S, Wang B, Walker MD, Tran HKR, Stogios PJ, Savchenko A, Grant RA,
633 McArthur AG, Laub MT, Whitney JC. 2019. An interbacterial toxin inhibits target cell
634 growth by synthesizing (p)ppApp. Nature 575.

635 32. Ruhe ZC, Low DA, Hayes CS. 2020. Polymorphic toxins and their immunity proteins:
636 Diversity, evolution, and mechanisms of delivery. Annu Rev Microbiol 74:497–520.

637 33. Lien YW, Lai EM. 2017. Type VI secretion effectors: Methodologies and biology. Front
638 Cell Infect Microbiol https://doi.org/10.3389/fcimb.2017.00254.

639 34. Liang X, Moore R, Wilton M, Wong MJQ, Lam L, Dong TG. 2015. Identification of
640 divergent type VI secretion effectors using a conserved chaperone domain. Proc Natl
641 Acad Sci 112:9106–9111.

642 35. Dong TG, Ho BT, Yoder-Himes DR, Mekalanos JJ. 2013. Identification of T6SS-
643 dependent effector and immunity proteins by Tn-seq in Vibrio cholerae. Proc Natl Acad
644 Sci 110:2623–2628.

645 36. Salomon D, Kinch LN, Trudgian DC, Guo X, Klimko JA, Grishin N V., Mirzaei H, Orth K.
646 2014. Marker for type VI secretion system effectors. Proc Natl Acad Sci 111:9271–9276.

647 37. Dar Y, Jana B, Bosis E, Salomon D. 2021. A binary effector module secreted by a type VI

secretion system. EMBO Rep e53981.

38. Liang X, Pei TT, Li H, Zheng HY, Luo H, Cui Y, Tang MX, Zhao YJ, Xu P, Dong T. 2021. VgrG-dependent effectors and chaperones modulate the assembly of the type VI secretion system. PLoS Pathog 17:e1010116.

39. Manera K, Kamal F, Burkinshaw B, Dong TG. 2021. Essential functions of chaperones and adaptors of protein secretion systems in Gram-negative bacteria. FEBS J https://doi.org/10.1111/FEBS.16056.

40. Baker-Austin C, Oliver JD, Alam M, Ali A, Waldor MK, Qadri F, Martinez-Urtaza J. 2018. Vibrio spp. infections. Nat Rev Dis Prim 4:1–19.

41. Yu Y, Yang H, Li J, Zhang P, Wu B, Zhu B, Zhang Y, Fang W. 2012. Putative type VI secretion systems of Vibrio parahaemolyticus contribute to adhesion to cultured cell monolayers. Arch Microbiol 194:827–835.

42. Li P, Kinch LN, Ray A, Dalia AB, Cong Q, Nunan LM, Camilli A, Grishin N V, Salomon D, Orth K. 2017. Acute hepatopancreatic necrosis disease-causing Vibrio parahaemolyticus strains maintain an antibacterial type VI secretion system with versatile effector repertoires. Appl Environ Microbiol 83:e00737-17.

43. Salomon D, Gonzalez H, Updegraff BL, Orth K. 2013. Vibrio parahaemolyticus Type VI secretion system 1 Is activated in marine conditions to target bacteria, and is differentially regulated from system 2. PLoS One 8:e61086.

44. Metzger LC, Matthey N, Stoudmann C, Collas EJ, Blokesch M. 2019. Ecological implications of gene regulation by TfoX and TfoY among diverse *Vibrio* species. Environ Microbiol 21:2231–2247.

45. Dar Y, Salomon D, Bosis E. 2018. The antibacterial and anti-eukaryotic Type VI secretion system MIX-effector repertoire in Vibrionaceae. Mar Drugs 16:433.

46. Le Roux F, Blokesch M. 2018. Eco-evolutionary dynamics linked to horizontal gene transfer in vibrios. Annu Rev Microbiol 72:annurev-micro-090817-062148.

47. Sun Y, Bernardy EE, Hammer BK, Miyashiro T. 2013. Competence and natural transformation in vibrios. Mol Microbiol 89.

48. Salomon D, Klimko JA, Trudgian DC, Kinch LN, Grishin N V., Mirzaei H, Orth K. 2015. Type VI secretion system toxins horizontally shared between marine bacteria. PLoS Pathog 11:1–20.

49. Thomas J, Watve SS, Ratcliff WC, Hammer BK. 2017. Horizontal gene transfer of functional type VI killing genes by natural transformation. MBio 8:1–11.

50. Ray A, Schwartz N, Souza Santos M, Zhang J, Orth K, Salomon D, de Souza Santos M, Zhang J, Orth K, Salomon D. 2017. Type VI secretion system MIX-effectors carry both antibacterial and anti-eukaryotic activities. EMBO Rep 18:e201744226.

51. Unterweger D, Kostiuk B, Pukatzki S. 2017. Adaptor proteins of type VI secretion system effectors. Trends Microbiol 25:8–10.

52. Käll L, Krogh A, Sonnhammer EL. 2004. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 338:1027–1036.

53. Zhang D, de Souza RF, Anantharaman V, Iyer LM, Aravind L. 2012. Polymorphic toxin systems: Comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. Biol Direct 7:18.

691 54. Lopez J, Le N-H, Moon KH, Salomon D, Bosis E, Feldman MF. 2021. Formylglycine-
692 generating enzyme-like proteins constitute a novel family of widespread type VI secretion
693 system immunity proteins. J Bacteriol 203.

694 55. Jana B, Keppel K, Salomon D. 2021. Engineering a customizable antibacterial T6SS-
695 based platform in Vibrio natriegens. EMBO Rep 22:e53681.

696 56. Salomon D, Klimko JA, Orth K. 2014. H-NS regulates the Vibrio parahaemolyticus type VI
697 secretion system 1. Microbiol (United Kingdom) 160:1867–1873.

698 57. Russell AB, Hood RD, Bui NK, Leroux M, Vollmer W, Mougous JD. 2011. Type VI
699 secretion delivers bacteriolytic effectors to target cells. Nature 475:343–349.

700 58. Ringel PD, Hu D, Basler M. 2017. The role of type VI secretion system effectors in target
701 cell lysis and subsequent horizontal gene transfer. Cell Rep 21:3927–3940.

702 59. Hersch SJ, Watanabe N, Stietz MS, Manera K, Kamal F, Burkinshaw B, Lam L, Pun A, Li
703 M, Savchenko A, Dong TG. 2020. Envelope stress responses defend against type six
704 secretion system attacks independently of immunity proteins. Nat Microbiol 5:706–714.

705 60. Hersch SJ, Manera K, Dong TG. 2020. Defending against the Type Six Secretion
706 System: beyond Immunity Genes. Cell Rep 33:108259.

707 61. Aoki SK, Diner EJ, de Roodenbeke C t'Kint, Burgess BR, Poole SJ, Braaten BA, Jones
708 AM, Webb JS, Hayes CS, Cotter PA, Low DA. 2010. A widespread family of polymorphic
709 contact-dependent toxin delivery systems in bacteria. Nature 468:439–442.

710 62. Tran L, Nunan L, Redman R, Mohney L, Pantoja C, Fitzsimmons K, Lightner D. 2013.
711 Determination of the infectious nature of the agent of acute hepatopancreatic necrosis
712 syndrome affecting penaeid shrimp. Dis Aquat Organ 105:45–55.

713 63. Borgeaud S, Metzger LC, Scrignari T, Blokesch M. 2015. The type VI secretion system of
714 Vibrio cholerae fosters horizontal gene transfer. Science 347:63–7.

715 64. Matthey N, Stutzmann S, Stoudmann C, Guex N, Iseli C, Blokesch M. 2019. Neighbor
716 predation linked to natural competence fosters the transfer of large genomic regions in
717 vibrio cholerae. Elife 8.

718 65. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic
719 assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343–345.

720 66. O'Toole R, Milton DL, Wolf-Watz H. 1996. Chemotactic motility is required for invasion of
721 the host by the fish pathogen Vibrio anguillarum. Mol Microbiol 19:625–637.

722 67. Lee I, Ouk Kim Y, Park S-C, Chun J. 2016. OrthoANI: An improved algorithm and
723 software for calculating average nucleotide identity. Int J Syst Evol Microbiol 66:1100–
724 1103.

725 68. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.
726 2009. BLAST+: Architecture and applications. BMC Bioinformatics 10.

727 69. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997.
728 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.
729 Nucleic Acids Res 25:3389–402.

730 70. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK,
731 Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N,
732 Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH. 2017. CDD/SPARCLE:
733 functional classification of proteins via subfamily domain architectures. Nucleic Acids Res

45:D200–D203.

71. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–66.

72. Katoh K, Rozewicki J, Yamada KD. 2018. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160–1166.

73. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.

74. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539.

75. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549.

76. Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao L, He S, Hurwitz DI, Jackson JD, Ke Z, Krylov D, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang D, Bryant SH. 2007. CDD: a conserved domain database for interactive domain family analysis. Nucleic Acids Res 35:D237-40.

77. Zimmermann L, Stephens A, Nam S-Z, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J Mol Biol 430:2237–2243.