



ARTICLE

<https://doi.org/10.1038/s41467-019-11546-6>

OPEN

A modular effector with a DNase domain and a marker for T6SS substrates

Biswanath Jana^{1,3}, Chaya M. Fridman^{1,3}, Eran Bosis²  & Dor Salomon¹ 

Bacteria deliver toxic effectors via type VI secretion systems (T6SSs) to dominate competitors, but the identity and function of many effectors remain unknown. Here we identify a *Vibrio* antibacterial T6SS effector that contains a previously undescribed, widespread DNase toxin domain that we call PoNe (Polymorphic Nuclease effector). PoNe belongs to a diverse superfamily of PD-(D/E)xK phosphodiesterases, and is associated with several toxin delivery systems including type V, type VI, and type VII. PoNe toxicity is antagonized by cognate immunity proteins (PoNi) containing DUF1911 and DUF1910 domains. In addition to PoNe, the effector contains a domain of unknown function (FIX domain) that is also found N-terminal to known toxin domains and is genetically and functionally linked to T6SS. FIX sequences can be used to identify T6SS effector candidates with potentially novel toxin domains. Our findings underline the modular nature of bacterial effectors harboring delivery or marker domains, specific to a secretion system, fused to interchangeable toxins.

¹Department of Clinical Microbiology and Immunology, Sackler Faculty of Medicine, Tel Aviv University, 6997801 Tel Aviv, Israel. ²Department of Biotechnology Engineering, ORT Braude College of Engineering, 2161002 Karmiel, Israel. ³These authors contributed equally: Biswanath Jana, Chaya M. Fridman. Correspondence and requests for materials should be addressed to E.B. (email: bois@braude.ac.il) or to D.S. (email: dorsalomon@mail.tau.ac.il)

Bacteria are social organisms that constantly interact with neighboring cells. To ensure sufficient nutrients and space, bacteria must be able to outcompete their nonkin neighbors. The type VI secretion system (T6SS) is a widespread mechanism used by Gram-negative bacteria to antagonize competitors^{1,2}. The T6SS is a multi-protein machine that is often encoded by large gene clusters³. It delivers toxins, called effectors, across the bacterial cell envelope and into neighboring cells^{4,5}. The T6SS is composed of three sub-complexes: the membrane complex, the baseplate, and the tail⁶. The membrane complex recruits the baseplate to the cell membrane⁷. Consequently, the baseplate serves as a platform for assembling the tail tube. This secreted tube is composed of stacked hexameric rings of a protein called Hcp; these rings are capped by a spike composed of a VgrG trimer and a PAAR repeat-containing protein that sharpens the tip^{1,2,8–10}. Effectors that mediate T6SS activity decorate the secreted tube. Contraction of an outer sheath propels the inner tube, with the effectors, out of the cell¹¹.

Although the T6SS was originally identified as a virulence mechanism^{1,5}, it is now commonly acknowledged that T6SSs mainly play a role in interbacterial competition⁴. Indeed, most T6SS effectors that have been identified mediate antibacterial activities¹². Such antibacterial effectors are paired, as bi-cistronic units, with cognate immunity proteins that inhibit the activity of the effectors and thus protect the bacterium from self- or kin intoxication¹³. To date, various antibacterial effector families have been described. Some effectors were shown to target vital bacterial components such as nucleic acids^{14,15}, cell walls^{4,13}, cell membranes^{16–18}, and the essential dinucleotides NAD(+) and NADP(+) ^{19,20}, whereas others target conserved proteins such as FtsZ, which is involved in bacterial cell division²¹. Noteworthy, various T6SS effectors targeting eukaryotic cell processes and components have also been described²².

T6SS effectors have been identified mainly by examining genes within or adjacent to T6SS gene clusters²³, their homology to known effectors^{13,16}, or by the presence of an N-terminal delivery domain^{5,10,24,25}. N-terminal domains associated with T6SS effectors can be either T6SS-secreted tail components (e.g., VgrG⁵, Hcp²⁶, or PAAR¹⁰) or specialized domains such as Rhs²⁷ and MIX²⁴; the latter can serve as a marker for T6SS substrates. Notably, many toxins that have been identified as T6SS effectors do not contain a recognizable delivery domain or signal, suggesting that additional delivery domains remain to be discovered.

Vibrio parahaemolyticus, an emerging pathogen found in marine environments²⁸, harbors two T6SSs^{29,30}. T6SS1 mediates antibacterial activity and is predominantly associated with pathogenic isolates^{24,30,31}, whereas T6SS2 is found in all *V. parahaemolyticus* isolates²⁹. We previously characterized T6SS1 in the clinical isolate RIMD 2210633 and identified three antibacterial effectors²⁴. Two effectors, VP1388 and VP1415, are encoded within the T6SS1 gene cluster and are common to T6SS1 clusters in *V. parahaemolyticus*. VP1388 contains a MIX domain, whereas VP1415 contains a PAAR domain. The third effector, VPA1263, is a MIX-effector present only in a subset of *V. parahaemolyticus* isolates³². Several reports revealed that members of the same bacterial species often harbor different T6SS effector arsenals^{33–35}. We recently showed that MIX effectors can account for some of the effector diversity in members of the *Vibrionaceae* family^{24,25,32}. However, we found that many T6SS-encoding *Vibrionaceae* strains, including *V. parahaemolyticus* isolates, do not contain multiple MIX effectors²⁵. Thus, other mechanisms of effector diversification probably exist in vibrios.

Here, we set out to identify non-MIX effectors that diversify T6SS1 repertoires in *V. parahaemolyticus*. To this end, we focus on *V. parahaemolyticus* 12-297/B, an isolate from Vietnam that causes acute hepatopancreatic necrosis disease in shrimp³¹.

Vibrio parahaemolyticus 12-297/B does not encode MIX effectors, aside from a VP1388-homolog that is encoded within the T6SS1 gene cluster³¹. Therefore, we hypothesize that it harbors non-MIX effectors that diversify its T6SS1 effector arsenal. Indeed, we find an auxiliary T6SS module that contains a VgrG-encoding gene, which we name *vgrG1b*. We show that the two additional genes found in this module constitute a T6SS1 effector/immunity pair. In addition, we demonstrate that the identified effector contains a novel DNase toxin domain, which we name PoNe (Polymorphic Nuclease effector), that is widespread in various antibacterial toxin secretion systems of both Gram-negative and Gram-positive bacteria. Moreover, we describe PoNi (Polymorphic Nuclease immunity), cognate immunity proteins for these DNase toxins that contain DUF1911 and DUF1910 domains. The newly identified effector further reveals an N-terminal domain, which we name FIX (Found in type sIX effector). We show that FIX is genetically and functionally linked to T6SS, and that it is found in many known and predicted T6SS effectors. We propose that FIX can serve as a marker for T6SS substrates.

Results

T6SS1 auxiliary module in *V. parahaemolyticus* 12-297/B. Effectors found outside of the main T6SS gene clusters in vibrios often belong to auxiliary T6SS modules that encode at least one of the tail tube components: VgrG, Hcp, or PAAR^{33,34}. Searching for such auxiliary T6SS modules, we performed TBLASTN analysis for genes encoding VgrG, Hcp, or PAAR proteins in the genome of *V. parahaemolyticus* 12-297/B. Our search revealed a gene encoding an orphan VgrG, B5C30_RS14470. This gene was annotated as a pseudogene, yet examination of the open reading frame revealed an intact gene (located between positions 22692 and 20311 in reference sequence NZ_MYFG01000470.1). This sequence is predicted to encode a VgrG protein homologous to VgrG1 of *V. parahaemolyticus*, with an additional C-terminal extension of 101 amino acids (Supplementary Fig. 1). Thus, we named it VgrG1b. *vgrG1b* appears to be the first in a three-gene operon flanked by a tRNA gene and predicted transposases (Fig. 1a), suggesting that this module is mobile. We found the VgrG1b module in the genomes of other *V. parahaemolyticus* strains isolated around the world (Fig. 1b and Supplementary Table 1).

New T6SS1 antibacterial effector/immunity pair. Since VgrG1b is very similar to VgrG1, we predicted that the VgrG1b module plays a role in T6SS1 activity. Before testing this, we first confirmed, using secretion and competition assays, that like the previously characterized *V. parahaemolyticus* POR1 (a RIMD 2210633 derivative) T6SS1^{24,30}, *V. parahaemolyticus* 12-297/B T6SS1 is functional and mediates antibacterial activity under warm marine-like conditions in the presence of surface-sensing activation (Supplementary Fig. 2).

Often, *vgrG*-containing modules encode downstream effectors. Therefore, we hypothesized that the two genes found downstream of *vgrG1b*, namely, B5C30_RS14465 (encoding WP_029857615.1) and B5C30_RS14460 (encoding WP_024703575.1) (hereafter named *v12_14465* and *v12_14460*, respectively), encode an effector/immunity (E/I) pair of T6SS1 (Fig. 1a). Consistent with this premise, deletion of the genes encoding the putative E/I pair ($\Delta v12_14465-0$) resulted in the loss of immunity against T6SS1-mediated self-intoxication. However, immunity was restored upon exogenous expression of V12_14460, thus confirming its role as an immunity protein (Fig. 1c). Self-intoxication, but not overall T6SS1-mediated antibacterial activity against *Escherichia coli* prey, was dependent on *v12_14465*, confirming that it is a *bona*

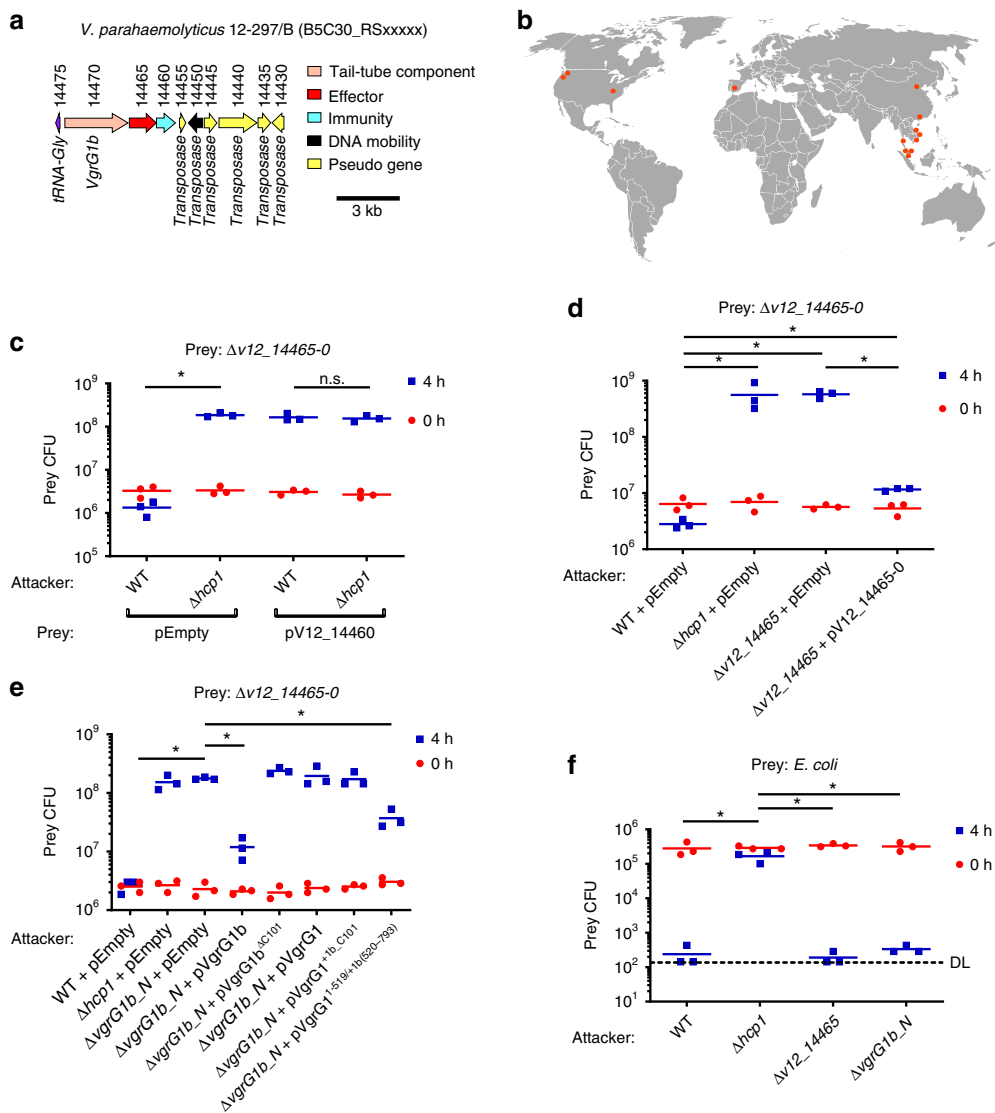


Fig. 1 An auxiliary T6SS1 module in *V. parahaemolyticus* 12-297/B. **a** T6SS1 VgrG1b auxiliary module. Genes are represented by arrows indicating the direction of translation. Locus tags (B5C30_RSxxxx) are shown above. **b** Locations in which the *V. parahaemolyticus* isolates containing the VgrG1b module have been isolated are denoted by red circles. Viability counts of *V. parahaemolyticus* 12-297/B (**c-e**) or *E. coli* (**f**) prey before (0 h) and after (4 h) co-incubation with the indicated *V. parahaemolyticus* 12-297/B attackers on media containing 3% NaCl at 30 °C. L-Arabinose was added (0.1% in **c**, 0.01% in **e**) to induce expression from the Pbad promoter. $\Delta hcp1$ is used as T6SS1⁻ control. Asterisks denote statistical significance between samples at the 4 h timepoint by unpaired, two-tailed Student's *t* test (**P* < 0.05); n.s. no significant difference, WT wild-type, DL detection limit. Source data are provided as a source data file

fide T6SS1 antibacterial effector (Fig. 1d, f). Notably, exogenous over-expression of V12_14465 alone proved toxic to expressing cells, and we therefore used the *v12_14465-0* E/I bi-cistronic unit (pV12_14465-0) to complement *v12_14465* deletion.

Since the effector V12_14465 is encoded adjacent to *vgrG1b*, we hypothesized that its delivery is dependent on the presence of VgrG1b. Indeed, deletion of *vgrG1b* ($\Delta vgrG1b_N$, deletion of nucleotides encoding the first 724 amino acids) hampered V12_14465-mediated self-intoxication, which was comparable to inactivation of T6SS1 by *hcp1* deletion (Fig. 1e). However, the deletion of *vgrG1b* did not affect the overall T6SS1-mediated antibacterial activity against *E. coli* prey, indicating that other T6SS1 effectors were still delivered in its absence (Fig. 1f). Notably, deletions of *hcp1*, *v12_14465*, or *vgrG1b* did not affect *V. parahaemolyticus* 12-297/B growth (Supplementary Fig. 3).

We further hypothesized that the 101 residues C-terminal extension, found in VgrG1b but absent in VgrG1 (Supplementary

Fig. 1), plays a role in the delivery of the downstream effector. Indeed, this C-terminal extension was required for effector delivery since ectopic expression of VgrG1b in which we removed the 101 C-terminal residues (VgrG1b^{ΔC101}) was unable to complement the loss of bacterial killing in the $\Delta vgrG1b_N$ strain (Fig. 1e). Yet, this C-terminal extension was not sufficient to allow delivery of the effector, as fusing this extension to the T6SS1 cluster-encoded VgrG1 (VgrG1^{+1b_C101}) did not complement the $\Delta vgrG1b_N$ strain. We noted that considerable differences are also found in the sequences corresponding to residues 520–692 of the two VgrG homologs (Supplementary Fig. 1). Therefore, we tested whether additional information required for effector delivery is found in this region of VgrG1b. As shown in Fig. 1e, swapping residues 520–692 of VgrG1 with residues 520–793 of VgrG1b (VgrG1^{1-519/+1b(520-793)}) resulted in partial complementation of the $\Delta vgrG1b_N$ strain. Taken together, our results indicate that *v12_14465-0* encodes a T6SS1 antibacterial E/I pair

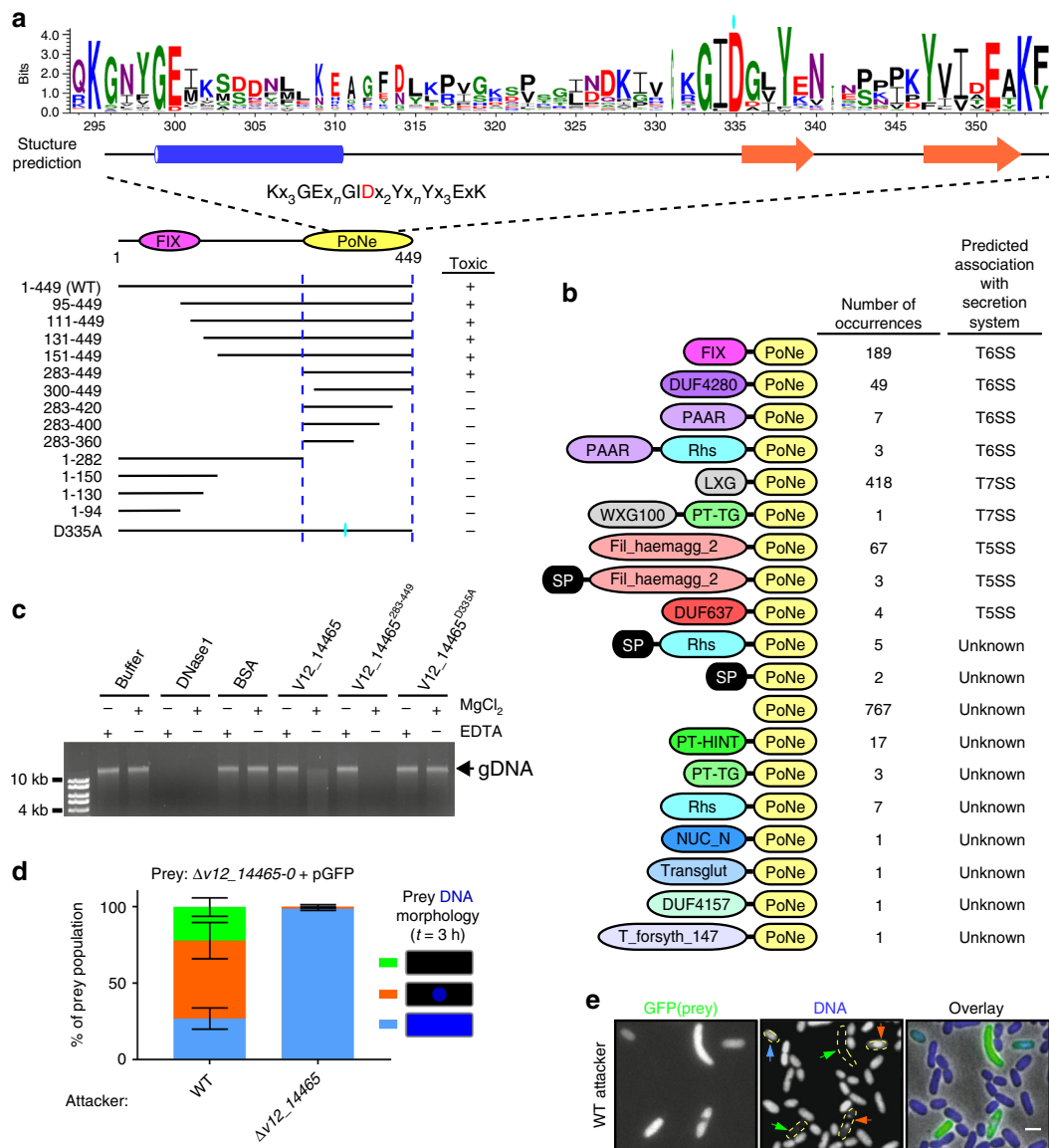


Fig. 2 Effector V12_14465 contains a new DNase domain. **a** Schematic representation of full-length and truncated V12_14465 forms tested for toxicity upon expression in *E. coli*. The region necessary and sufficient for toxicity is denoted by dashed vertical blue lines (based on the results shown in Supplementary Fig. 5). The region of the toxic C-terminal domain (PoNe) in which a conserved motif was identified (residues 294–354) is illustrated using WebLogo 3. Secondary-structure prediction (Jpred) and the conserved motif are provided. Alpha helix is denoted as blue cylinder, and beta strands as orange arrows. A cyan oval above the WebLogo denotes the catalytic aspartic acid at position 335. **b** Domain architectures of PoNe-containing proteins. SP, signal peptide. **c** In vitro DNase activity assay. Purified *E. coli* genomic DNA (gDNA) was co-incubated with buffer or with the indicated purified proteins in the presence (+) or absence (–) of Mg²⁺ or EDTA at 37 °C for 30 min. The integrity of gDNA was visualized on 1% agarose gel. **d** Quantification of *V. parahaemolyticus* 12-297/B $\Delta v12_14465-0$ prey cells, constitutively expressing GFP from a plasmid (pGFP), that present DNA morphologies: DNA across entire cell area (cyan), DNA nucleoid or foci (orange), or no DNA (green). DNA was detected using Hoechst dye fluorescence signal after 3 h of co-incubating prey with the indicated attacker strains. Results shown are average \pm S.D. of three independent biological replicates. In each experiment, 100 prey cells were evaluated per treatment. **e** Sample images of prey cells described in **d** after 3 h of co-incubation with WT attackers. Dashed yellow shapes in DNA channel encircle prey cells detected in the GFP channel. Colored arrows denote prey cells presenting each of the three DNA morphologies described in **d**. Bar = 2 μ m. WT wild-type. Source data are provided as a source data file

whose delivery is dependent on VgrG1b. Since all tested VgrG forms were stably expressed in *V. parahaemolyticus* (Supplementary Fig. 4), we also conclude that the C-terminal end of VgrG1b contains information required and sufficient for delivery of its downstream effector.

V12_14465 belongs to a widespread family of new DNase toxins. Analysis of V12_14465 revealed no sequence homology to known toxin domains. Therefore, we set out to decipher the

antibacterial mechanism of action and target of this effector. Exogenous expression of V12_14465 in *E. coli* was detrimental (Fig. 2a and Supplementary Fig. 5). Truncation analysis indicated that the V12_14465 C-terminal end, encompassing residues 283–449, was necessary and sufficient for the observed toxicity in *E. coli* (Fig. 2a and Supplementary Fig. 5). We named this toxic domain PoNe (as explained below). Using position-specific iterated BLAST (PSI-BLAST), we identified 1546 proteins containing a PoNe domain (Supplementary Data 1); these

proteins were widely distributed in the genomes of both Gram-negative and Gram-positive bacteria (Supplementary Fig. 6). Many PoNe domains were fused to N-terminal polymorphic toxin delivery domains of T6SS (e.g., PAAR¹⁰ and DUF4280³⁶), type VII secretion system (T7SS; e.g., LXG³⁷ and WXG100³⁸), and type V secretion system (T5SS) contact-dependent inhibition (e.g., DUF637 and Fil_{haemagg}^{239,40}); they were even found fused to predicted Sec pathway (T2SS) signal peptide (Fig. 2b and Supplementary Data 1). These results suggest that V12_14465 belongs to a widespread family of antibacterial effectors associated with various protein secretion systems.

When we analyzed the multiple sequence alignment of the identified PoNe family members, we discovered a conserved motif: K_xGEX_nGIDx₂Y_xY_xEXK (Fig. 2a). We noted that the PoNe conserved motif and its position within the predicted secondary structure resembled the conserved motif of members of the diverse PD-(D/E)xK phosphodiesterase superfamily, many of which target DNA^{41–43}. Therefore, we hypothesized that PoNe constitutes a new family of phosphodiesterases within this superfamily, and that V12_14465 is a nuclease that targets DNA. Indeed, purified V12_14465 or its PoNe domain (V12_14465^{283–449}) exhibited magnesium-dependent DNase activity in vitro (Fig. 2c). We reasoned that if V12_14465 acts as a DNase inside bacterial cells, then its activity should activate the SOS response, a mechanism used by bacteria to cope with DNA damage. As predicted, expression of V12_14465, or its PoNe domain, induced the expression of a fluorescent SOS response reporter in *E. coli* (Supplementary Fig. 7), consistent with PoNe inducing DNA damage. Elevated SOS signals were also detected under repressing conditions, possibly due to leaky expression which was also evident by visible toxicity (Supplementary Fig. 5a). Moreover, substitution of the conserved V12_14465 aspartic acid 335 for alanine (D335A), which is expected to play a role in the catalytic activity of the PD-(D/E)xK phosphodiesterase superfamily⁴¹, abolished bacterial toxicity (Fig. 2a and Supplementary Fig. 5), in vitro DNase activity (Fig. 2c), and SOS response activation (Supplementary Fig. 7). Noteworthy, when we extracted total RNA from *E. coli* cells expressing V12_14465 we did not observe differences compared with control cells, suggesting that PoNe targets DNA rather than RNA (Supplementary Fig. 8).

To further test the hypothesis that PoNe-containing V12_14465 functions as a DNase during bacterial competition, we used fluorescence microscopy. We visualized the DNA in prey cells lacking the *v12_14465-0* E/I pair ($\Delta v12_14465-0$) when competed against WT *V. parahaemolyticus* 12-297/B attacker or against a mutant in which the *v12_14465* effector was deleted ($\Delta v12_14465$). As shown in Fig. 2d, e, prey cells in which DNA was no longer visible were detected after 3 h incubation with the WT attacker, but not when prey were competed against an attacker lacking the effector *v12_14465*. Taken together, these results demonstrate that V12_14465 is a *bona fide* antibacterial DNase toxin.

Next, we aimed to determine whether other members of the PoNe family also possess DNase activity. To this end, we expressed in *E. coli* the *Bacillus cereus* ATCC 14579 protein BC3021 (NP_832767.1). This protein harbors a C-terminal PoNe domain fused to an N-terminal region containing a T7SS LXG delivery domain. As we observed for V12_14465, the PoNe-containing *B. cereus* BC3021 was detrimental to *E. coli*, activated the SOS response, and induced DNA degradation in vivo (Supplementary Fig. 9). Taking the above results together, we conclude that PoNe constitutes a new family of DNases that is widespread in antibacterial toxins associated with various secretion systems.

DUF1911-containing proteins provide immunity against PoNe. V12_14460, the immunity protein encoded downstream of effector *v12_14465*, contains a domain of unknown function 1911 (DUF1911). Out of 1546 PoNe-encoding genes, 1389 (89.84%) are found immediately upstream of a gene that encodes a protein containing DUF1911. DUF1911 is often also accompanied by a DUF1910 domain (Supplementary Data 1). This observation implied that DUF1911- and DUF1911-DUF1910-containing proteins are PoNe immunity proteins (hereafter referred to as PoNi). We investigated whether the DUF1911-containing V12_14460, which was required for immunity against V12_14465-mediated self-intoxication in *V. parahaemolyticus* (Fig. 1c), was sufficient to provide immunity against V12_14465 expressed in *E. coli*. As expected, expression of V12_14460 (PoNi^{VP}) specifically abrogated the toxicity mediated by the V12_14465 PoNe domain (residues 283–449; V12_14465^{PoNe}) in *E. coli*, but did not provide immunity against VPA1263, a T6SS1 effector from *V. parahaemolyticus* RIMD 2210633 with a predicted colicin DNase domain²⁴ (Fig. 3a). Next, we examined whether PoNi^{VP} provided immunity by direct interaction with the V12_14465 PoNe domain. Indeed, direct interaction between PoNi^{VP} and PoNe domain-containing variants of V12_14465, but not between PoNi^{VP} and the N-terminal portion of V12_14465 (residues 1–282) was apparent in both pull-down and bacterial two-hybrid (BACTH) assays (Fig. 3b, c). Accordingly, the DUF1911-containing PoNi BC3020 (accession number: NP_832766.1; hereafter referred to as PoNi^{BC}), encoded downstream of the *B. cereus* PoNe-containing toxin BC3021, was sufficient to rescue *E. coli* expressing the *B. cereus* PoNe domain (BC3021^{PoNe}) or the full-length BC3021 protein (Supplementary Fig. 9b). Notably, PoNi proteins appear to provide immunity specifically against their adjacently-encoded PoNe-containing effector, since the *Bacillus* and *Vibrio* PoNi were unable to cross-protect against a noncognate PoNe toxin (Supplementary Fig. 9e). These results indicate that DUF1911 domains, which are often accompanied by DUF1910, provide immunity against PoNe-containing DNase toxins, most likely by directly interacting with them.

The N-terminus of V12_14465 contains a new T6SS marker.

Next, we set out to characterize the N-terminal portion of effector V12_14465. Since PoNe is often fused to N-terminal effector delivery domains (Fig. 2b), we reasoned that the N-terminal part of V12_14465 contains a yet-unknown domain that is required for its T6SS-mediated delivery into recipient cells. Consistent with this premise, V12_14465 N-terminal truncations were unable to complement a deletion in *v12_14465* during self-competition against a $\Delta v12_14465-0$ prey (Fig. 4a), even though these truncations were still functional toxins as evident by their toxicity upon expression in *E. coli* (Supplementary Fig. 5). This indicated that the N-terminal part of V12_14465 is required for its T6SS1-mediated delivery. Using PSI-BLAST, we identified 2881 proteins containing a region homologous to the V12_14465 N-terminus. Interestingly, the identified proteins are almost exclusively encoded by Gram-negative bacteria (Supplementary data 2). Multiple sequence alignment of the homologous sequences revealed a possible domain of ~80 amino acids, corresponding to positions 43–121 in V12_14465, containing a conserved motif (Fig. 4b). Since this domain was first identified in T6SS effector V12_14465, we named it FIX (Found in type sIX effector).

Analysis of FIX-containing proteins showed that FIX is often fused to various C-terminal nuclease toxin domains, including PoNe, or to C-terminal extensions of unknown function (Fig. 4c and Supplementary Data 2). Interestingly, FIX is also present in various established T6SS-secreted effectors that have an

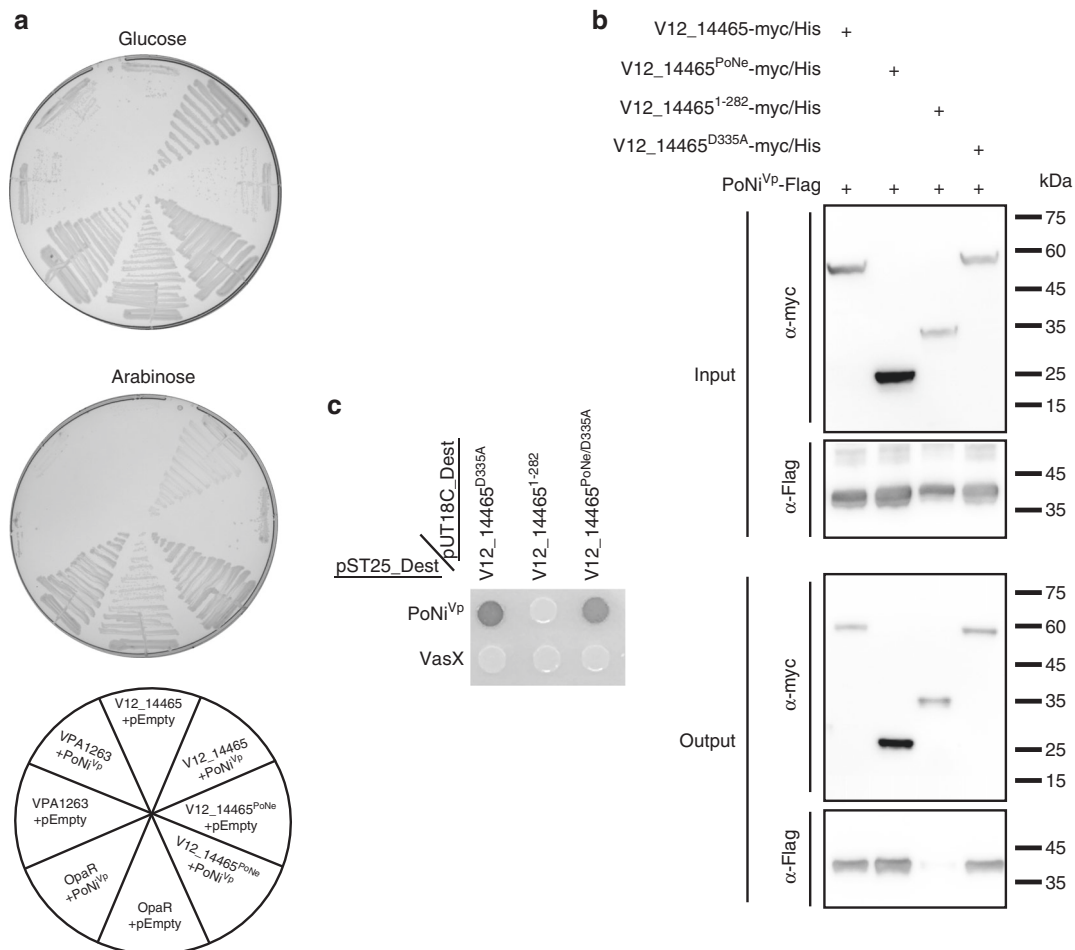


Fig. 3 V12_14460 provides immunity against PoNe domain. **a** Toxicity of V12_14465 variants expressed in *E. coli* BL21(DE3) from arabinose-inducible expression plasmid with or without V12_14460 (PoNi^{Vp}). OpaR, the *V. parahaemolyticus* high cell density quorum sensing master regulator, was used as a nontoxic control. VPA1263, a T6SS1 effector with a predicted colicin DNase domain, was used as a non-PoNe toxic control. **b** Ni-Sepharose resin pull-down of myc/His-tagged V12_14465 forms with PoNi^{Vp}-Flag after co-expression in *E. coli* BL21(DE3). **c** Bacterial two-hybrid assay. The indicated plates were expressed in *E. coli* BTH101 reporter strain fused to the T18 or T25 domain of the *Bordetella* adenylate cyclase. Bacteria were spotted on plates supplemented with the chromogenic substrate X-Gal. A dark-colored colony indicates protein-protein interaction. The *V. cholerae* T6SS effector VasX was used as a negative control. PoNi^{Vp}, V12_14460; V12_14465^{PoNe}, V12_14465(283–449); V12_14465^{PoNe/D335A}, V12_14465(283–449/D335A); pEmpty, empty expression vector. Source data are provided as a source data file

N-terminal VgrG or PAAR domain (e.g., *V. parahaemolyticus* T6SS1 AHH nuclease domain-containing effector VP1415²⁴) (Fig. 4c and Supplementary Data 2). The presence of T6SS-related domains such as PAAR and VgrG in FIX-containing proteins, but not of domains associated with other effector secretion systems, suggested that FIX may be specifically linked to T6SS. Indeed, FIX-containing proteins are encoded almost exclusively (>96%) in bacterial genomes that harbor T6SS (Fig. 4d and Supplementary Data 3). In agreement with this observation, we identified genes encoding T6SS components, such as the structural components VgrG⁵ and TssL⁷, or the adapter DUF4123^{44,45}, in 70.14% of the genes found directly upstream to FIX (Fig. 4e and Supplementary Data 2). An additional 3.71% of FIX-containing proteins can also be directly linked to T6SS because they contain PAAR or PAAR-like (i.e., DUF4280) domains (Supplementary Data 2). These findings suggested that FIX plays a role in delivery of T6SS effectors. In support of such a role, we found that deletion of the central FIX region encompassing residues 70–89 of the effector V12_14465 (V12_14465^{Δ70–89}) abrogated the effector's ability to mediate toxicity during bacterial competition (Fig. 4f). Noteworthy, V12_14465^{Δ70–89} retained its toxic activity when exogenously expressed in *E. coli* (Supplementary Fig. 5).

Based on these results, we propose that FIX plays a role in T6SS effector delivery, and that it can serve as a new marker for T6SS-delivered proteins to enable the identification of novel T6SS substrates.

Discussion

In this work, we identified a VgrG module containing a new T6SS antibacterial effector/immunity pair in *V. parahaemolyticus*. We revealed a new bacterial DNase toxin family, which we named PoNe, that belongs to the PD-(D/E)xK superfamily of phosphodiesterases^{42,43}. PoNe is widespread and found in various Gram-negative and Gram-positive bacterial species. Because of the extreme divergence in sequence and structure of PD-(D/E)xK superfamily members, it was impossible to simply identify PoNe by homology searches⁴². Like previously reported toxins belonging to the Tox-REase class^{46,47}, PoNe possesses the quintessential catalytic D-(D/E)xK residues in secondary-structure positions that are typical of members of this superfamily. However, the broad conserved motif that we identified in PoNe, which surrounds these core residues, is different from motifs found in Tox-REase toxins, indicating that PoNe represents a new and distinct toxin family. This conclusion is further supported by the

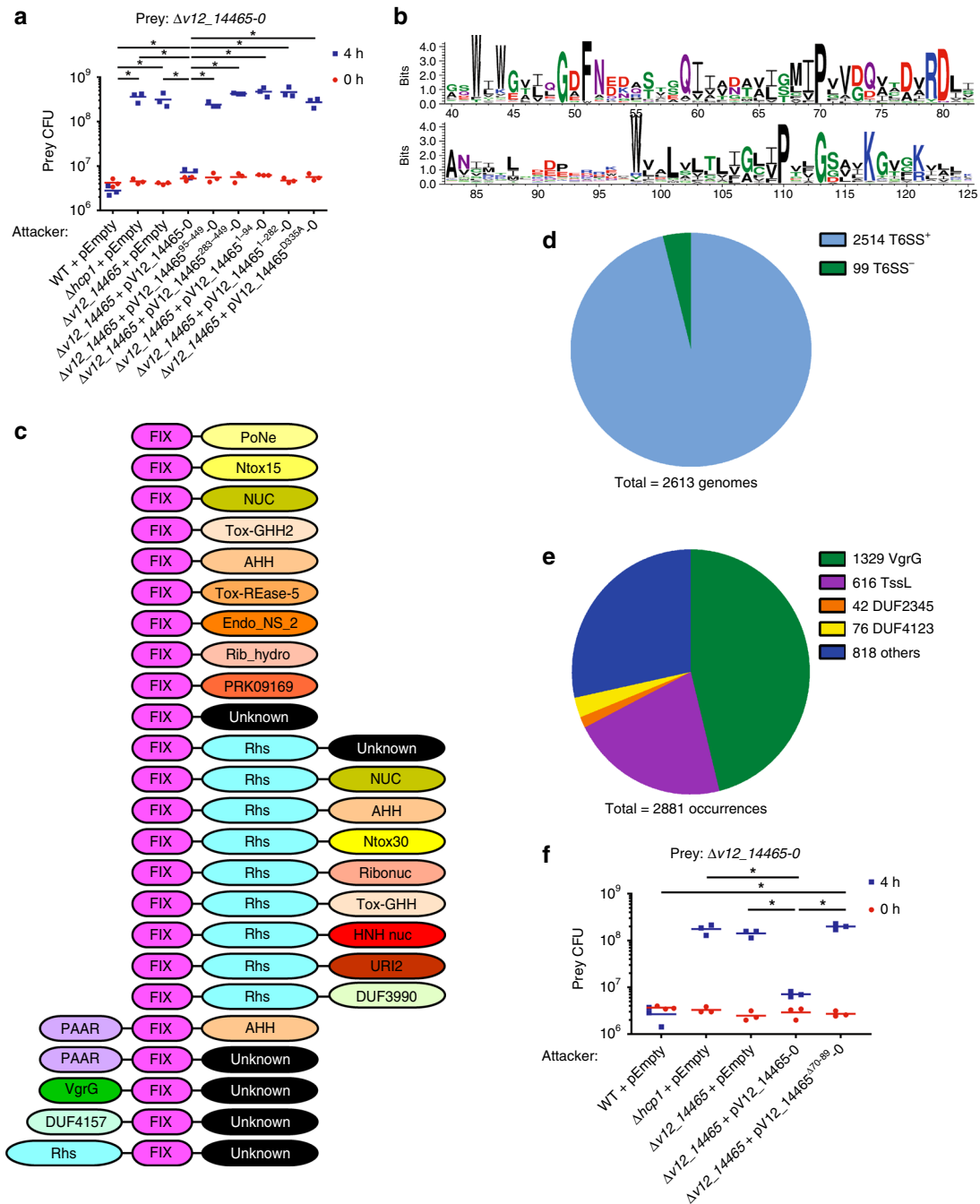


Fig. 4 V12_14465 contains a marker for T6SS substrates. **a** Viability counts of *V. parahaemolyticus* 12-297/B prey before (0 h) and after (4 h) co-incubation with the indicated *V. parahaemolyticus* 12-297/B attackers at 30 °C on media containing 3% NaCl. **b** Conservation pattern found in FIX homologs. The region shown corresponds to positions 40–125 in V12_14465. The image was produced using WebLogo 3. **c** Representative domain architectures of FIX-containing proteins. **d** Pie chart of FIX-containing bacterial genomes that encode a T6SS (encoding at least 9 T6SS core components). **e** Pie chart of domains encoded by genes found immediately upstream of FIX. The number of occurrences for each domain is listed next to its name. **f** Viability counts of *V. parahaemolyticus* 12-297/B prey before (0 h) and after (4 h) co-incubation with the indicated *V. parahaemolyticus* 12-297/B attackers at 30 °C on media containing 3% NaCl. Asterisks denote the statistical significance between samples at the 4 h timepoint by unpaired, two-tailed Student’s *t* test (**P* < 0.05); WT wild-type. Source data are provided as a source data file

wide distribution of PoNe domains in bacterial species, and by its association with various effector secretion systems, such as T5SS, T6SS, and T7SS. Importantly, to the best of our knowledge, PoNe is the first toxin family of PD-(D/E)xK phosphodiesterases that has been experimentally demonstrated to have DNase activity.

We further showed that proteins containing DUF1911 and DUF1910, which we named PoNi, function as specific immunity proteins that bind and antagonize PoNe. Notably, DUF1911 and

DUF1910 can be found downstream of proteins in which the PoNe domain was not detected. This can be explained, at least in part, by the presence of PoNe-like toxin domains (which did not pass our conservative threshold for PoNe family members; see the Methods section) that employ PoNi-like immunity proteins. In other instances, PoNi may be an orphan immunity gene, as was previously demonstrated for various T6SS immunity genes^{13,48}.

The modular nature of PoNe-containing toxins employing delivery domains of diverse secretion systems prompted us to investigate the N-terminal end of effector V12_14465. This led us to identify a region of ~80 residues that we named FIX. Like the previously reported T6SS marker MIX, FIX is genetically linked to T6SS. FIX is encoded almost exclusively by bacteria that harbor T6SS (>96% of genomes encoding FIX also harbor T6SS, whereas only ~25% of Gram-negative bacterial genomes are expected to encode T6SS⁴⁹). FIX-encoded genes are often neighbors of T6SS component-encoding genes such as VgrG or TssL, and FIX is found fused to known T6SS delivery domains such as PAAR or VgrG. Moreover, FIX is found N-terminal to diverse known toxin domains. Interestingly, most of these toxin domains, like PoNe, are predicted to target nucleic acids. These findings led us to propose that FIX can be used as a new marker for T6SS substrates. Therefore, it is likely that many, if not all, FIX-containing proteins are T6SS effectors. Importantly, aside from MIX, FIX is the only currently available marker for T6SS-secreted proteins that is not a T6SS structural component (i.e., Hcp, VgrG, and PAAR). Future studies of FIX proteins containing C-terminal domains of unknown function may lead to the discovery of novel toxin domains.

We noted that the T6SS markers, MIX²⁴ and FIX, appear to be mutually exclusive because we did not find proteins containing both of them. This observation implies different or competing mechanisms of T6SS association or roles in the delivery process for these two markers. Future work will determine whether FIX plays either a structural or functional role in T6SS-mediated delivery, such as binding to T6SS-secreted structural components. Although we showed that FIX is required for delivery of a T6SS effector, we have not been able to determine whether FIX is also a *bona fide* T6SS secretion domain sufficient for T6SS-mediated delivery, or whether it can mediate binding directly to a T6SS tail component. Thus, elucidating FIX's exact function will be addressed in future studies.

In conclusion, we identified a new DNase toxin family and its cognate immunity domains, and we also revealed a new marker for T6SS substrates. Our findings illustrate the modular nature of secreted bacterial effectors, and demonstrate the potential to use such modular domains for discovering new toxins.

Methods

Strains and media. *Vibrio parahaemolyticus* isolates 12-297/B³¹ (gift from Donald Lightner at the University of Arizona), RIMD 2210633 (gift from Kim Orth at the University of Texas Southwestern Medical Center), PORI³⁰ (RIMD 2210633 Δ *tdhA5* derivative), and their derivatives were routinely grown in Marine Lysogeny broth (MLB; LB media supplemented with NaCl to a final concentration of 3% wt/vol) or on marine minimal media (MMM) agar plates (2% wt/vol NaCl, 0.4% wt/vol galactose, 5 mM MgSO₄, 5 mM K₂SO₄, 77 mM K₂HPO₄, 35 mM KH₂PO₄, 20 mM NH₄Cl, 1.5% wt/vol agar) at 30 °C.

Escherichia coli strain DH5 α (λ pir) (gift from Eric V. Stabb at the University of Georgia) was used for plasmid maintenance and amplification. *E. coli* strain DH5 α was used as prey in bacterial competition assays. *E. coli* strain BL21(DE3) was used for protein expression and for toxicity assays. *E. coli* strain BTH101 (from BACTH system kit) was used for BACTH experiments. *E. coli* were routinely grown in 2xYT media (1.6% wt/vol tryptone, 1% wt/vol yeast extract, 0.5% wt/vol NaCl), or on LB (1% wt/vol NaCl) agar plates at 37 °C. When necessary, media were supplemented with 10 μ g/mL chloramphenicol, 30 μ g/mL (for *E. coli*) or 250 μ g/mL (for *V. parahaemolyticus*) kanamycin, 100 μ g/mL ampicillin, or 100 μ g/mL spectinomycin. To induce the expression of genes from a plasmid, 0.01, 0.05, or 0.1% (wt/vol) l-arabinose was included in the media.

Plasmid construction. For arabinose-inducible expression in bacteria, DNA fragments corresponding to the full-length or truncated forms of B5C30_RS14465 (V12_14465) or BC3021 were amplified from *V. parahaemolyticus* 12-297/B or *B. cereus* ATCC 14579 genomic DNA, respectively. PCR fragments were inserted into the multiple cloning site (MCS) of the pBAD/Myc-His vector (Invitrogen) harboring a kanamycin-resistance cassette³⁰, in-frame with the C-terminal Myc-6xHis tag.

For arabinose-inducible expression of immunity genes, the coding sequence of B5C30_RS14460 (V12_14460, PoNi^VP) or BC3020 (PoNi^BC) was amplified with an

in-frame C-terminal Flag tag from *V. parahaemolyticus* 12-297/B or *B. cereus* ATCC 14579 genomic DNA, respectively. PCR fragments were inserted into the MCS of pBAD33.1 (Addgene).

For complementation of the *v12_14465* deletion strain, the indicated *v12_14465* variants were cloned 5' of the *v12_14460* coding sequence, and 3' of the predicted native promoter region of the *vgrG1b-v12_14465-v12_14460* operon (426 bp upstream of the *vgrG1b*, B5C30_RS14470, start site). l-Arabinose was not added to induce expression from these plasmids in *Vibrio*, since expression was mediated by the *vgrG1b* native promoter. For complementation of *vgrG1b* deletion, the *vgrG1* coding sequence (B5C30_RS15295), the *vgrG1b* coding sequence (found between positions 22692 and 20311 in reference sequence NZ_MYFG01000470.1) or its truncated versions, and fusions of VgrG1 with VgrG1b were cloned into the MCS of pBAD/Myc-His vector (Invitrogen) harboring a kanamycin-resistance cassette, with a C-terminal Flag tag.

For bacterial expression of Tse1 (NP_250535.1), a peptidoglycan-hydrolase T6SS effector⁴, its coding sequence was amplified from *Pseudomonas aeruginosa* PAO1 genomic DNA and inserted into pPER5²⁵ in-frame with a N-terminal PelB signal peptide and a C-terminal Myc-6xHis tag. The Gibson assembly method⁵¹ was employed to generate the above plasmids.

For bacterial expression of VPA1263, a T6SS1 effector with predicted colicin-like DNase domain²⁴, its coding sequence was amplified from *V. parahaemolyticus* RIMD 2210633 genomic DNA and inserted between the SacI and KpnI sites in the MCS of pBAD2³⁰ in-frame with a C-terminal Myc-6xHis tag. Construction of the plasmid for expression of OpaR was previously reported³⁰.

To generate V12_14465^{D335A} mutant, site-directed mutagenesis was performed using the above-mentioned pBAD/Myc-His plasmids harboring *v12_14465* as a template to generate pV12_14465^{D335A}. This construct served as a template for amplification and the subsequent cloning of the D335A variant into subsequently described plasmids.

Plasmids used for the BACTH assay were constructed using the Gateway cloning method. The indicated *v12_14465* forms were amplified from pV12_14465^{D335A}, whereas the *Vibrio cholerae* effector VasX (NP_232421.1) was amplified from *V. cholerae* N16961 genomic DNA. The genes were introduced into the entry plasmid pENTER-D-TOPO using the pENTER-D-TOPO cloning kit (Invitrogen) following the manufacturer's instructions, followed by transfer to the BACTH destination plasmids pST25_DEST or pUT18c_DEST⁵² using the Gateway LR clonase II enzyme mix kit (Invitrogen), according to the manufacturer's instructions.

All constructs were confirmed by DNA sequencing.

Construction of deletion strains. For in-frame, genomic deletions of *hcp1* (B5C30_RS15290), *vgrG1b* (B5C30_RS14470), *v12_14465* (B5C30_RS14465), or *v12_14465-0* (B5C30_RS14465 - B5C30_RS14460) in *V. parahaemolyticus* 12-297/B, 1 kb sequences upstream and downstream of each specified gene were cloned into pDM4, a Cm^rOriR6K suicide plasmid⁵³. pDM4 constructs were introduced into *V. parahaemolyticus* 12-297/B via tri-parental mating using *E. coli* DH5 α (λ pir) and a conjugation helper. Trans-conjugants were selected on MMM agar plates containing chloramphenicol (10 μ g/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. In-frame, genomic deletion of *V. parahaemolyticus* RIMD 2210633 *hcp1* was performed as described, using pDM4 plasmid described previously³⁰. Construction of in-frame, genomic deletion of PORI *hcp1* was reported previously³⁰.

Bacterial competition assay. Attacker and prey strains were grown overnight in appropriate broth (MLB for *V. parahaemolyticus* and 2xYT for *E. coli*) with the addition of antibiotics when maintenance of plasmids was required. Competition assays were performed as previously described³⁰. Briefly, cultures were normalized to OD₆₀₀ = 0.5 and were mixed at a 4:1 ratio (attacker:prey). Triplicates of mixtures were incubated for 4 h at 30 °C on MLB agar plates or on MLB plates containing 0.1 or 0.01% (wt/vol) l-arabinose (when required to induce expression from plasmids). CFU of prey were calculated after they were grown on selective plates at 0 and 4 h. Assays were repeated at least three times with similar results, and the results from representative experiments are shown.

Protein secretion assay. Expression and secretion of *V. parahaemolyticus* VgrG1 were determined as previously described³¹. Briefly, overnight cultures of *V. parahaemolyticus* were normalized to OD₆₀₀ = 0.18 in 5 mL of MLB media, and grown for 5 h at 30 °C. Phenamil (20 μ M) was added to induce T6SS1³⁰ (phenamil is an inhibitor of the polar flagella; therefore, it mimics surface-sensing activation⁵⁴). For expression fractions (cells) 1.0 OD₆₀₀ units of cells were collected and resuspended in (2x) tris-glycine SDS sample buffer (Novex, Life Sciences). Supernatants of volumes equivalent to 10 OD₆₀₀ units (media) were filtered (0.22 μ m) and precipitated with deoxycholate and trichloroacetic acid⁵⁵. Precipitated proteins were washed twice with ice-cold acetone prior to re-suspension in 20 μ L of 10 mM Tris-HCl pH = 8.0, followed by the addition of 20 μ L of 2x protein sample buffer. Samples were resolved on TGX stain-free gels (Bio-Rad), transferred onto polyvinylidene difluoride (PVDF) membranes, and immunoblotted with custom-made anti-VgrG1 antibodies³¹ at 1:1000 dilution. Loading of total protein lysates was

visualized by analyzing trihalo compounds' fluorescence of the immunoblot membrane. Protein signals were visualized by enhanced chemiluminescence (ECL). Experiments were repeated at least three times with similar results, and the results from representative experiments are shown.

To verify expression of VgrG1, VgrG1b, and their fusion versions in *V. parahaemolyticus*, cultures were grown as described above, and induced in MLB media containing L-arabinose (0.01%) and appropriate antibiotics for 5 h at 30 °C. A total of 0.5 OD₆₀₀ units of cells were collected and treated as described above. Protein expression was detected in SDS-PAGE using Flag antibodies (DYKDDDDK Tag (D6W5B) Rabbit mAb #14793; Cell Signaling Technology) at 1:1000 dilution.

Bacterial toxicity assay. To assess the toxicity of antibacterial effectors, *E. coli* BL21(DE3) were transformed with the indicated pBAD-based L-arabinose-inducible expression vectors. *E. coli* transformants were streaked onto either repressing (containing 0.2% wt/vol glucose) or inducing (containing 0.05% wt/vol L-arabinose) LB agar plates containing kanamycin (30 µg/mL). To assess the protection conferred by immunity proteins, pBAD33.1 L-arabinose-inducible expression vectors carrying predicted immunity genes were co-transformed with the plasmids mentioned above into *E. coli* BL21(DE3). Transformants were streaked onto either repressing or inducing LB agar plates containing kanamycin (30 µg/mL) and chloramphenicol (10 µg/mL). Bacterial growth was assessed after overnight incubation at 37 °C.

Vibrio growth assay. Overnight-grown cultures of *V. parahaemolyticus* were normalized to an OD₆₀₀ of 0.01 in MLB media and transferred to 96-well plates (200 µL per well; *n* = 10). Cultures were grown at 30 °C in a BioTek SYNERGY H1 microplate reader with continuous shaking at 205 cpm. OD₆₀₀ readings were acquired every 10 min. Experiments were performed at least three times with similar results. Results from a representative experiment are shown.

SOS response activity. *E. coli* BL21(DE3) were co-transformed with the SOS response reporter plasmid pL(*lexO*)-GFP⁵⁶ and the indicated arabinose-inducible effector expression vector. Overnight cultures that were grown in 2xYT media containing the appropriate antibiotics and 0.2% (wt/vol) glucose (to repress expression from Pbad promoter) were washed with fresh 2xYT. Washed suspensions were normalized to OD₆₀₀ = 0.5 in 900 µL of 2xYT supplemented with appropriate antibiotics and either 0.2% (wt/vol) glucose or 0.1% (wt/vol) L-arabinose, to repress or induce gene expression, respectively. Cultures were grown at 37 °C for 2.5 h with agitation (215 rpm), before cells were pelleted and washed with 1 mL of M9 media. Cells were then resuspended in 800 µL of M9 media, and 200 µL of cell suspension was transferred, in triplicate, into wells of a black 96-well plate with a clear optical bottom. GFP fluorescence (excitation 479 nm and emission 520 nm) and OD₆₀₀ were measured using a BioTek SYNERGY H1 microplate reader. Data are presented as arbitrary units resulting from the division of fluorescence measurements by OD₆₀₀ per well. The experiment was repeated three times with similar results. Results from a representative experiment are shown.

Protein expression in *E. coli*. To test the expression of C-terminally myc-tagged V12_14465 variants from pBAD-based L-arabinose-inducible expression vectors in *E. coli*, nontoxic proteins were transformed into *E. coli* BL21(DE3). Toxic variants were co-transformed into *E. coli* BL21(DE3) together with pBAD33.1 containing the immunity gene, *v12_14460*, C-terminally tagged with Flag, to provide protection against self-intoxication and allow expressing cells to grow. Overnight bacterial cultures were diluted 100-fold into 5 mL of fresh 2xYT media supplemented with appropriate antibiotics and grown for 2 h at 37 °C with agitation (215 rpm). Protein expression was induced by addition of L-arabinose (0.1% wt/vol) to the media and an additional incubation at 30 °C for 4 h. Next, 1 mL of each culture was pelleted and the cells were resuspended in 100 µL of (2x) tris-glycine SDS sample buffer followed by boiling at 95 °C for 5 min. Samples were resolved on precast gels (GenScript), transferred onto PVDF membranes, and immunoblotted with c-Myc (9E10) antibodies (sc-40, Santa Cruz Biotechnology) at 1:1000 dilution. Loading of total protein lysates was visualized by Ponceau S staining of the immunoblot membrane. Protein signals were visualized by ECL. Experiments were repeated at least three times with similar results; the results from representative experiments are shown.

In vitro DNase assay. Overnight cultures of *E. coli* BL21(DE3) containing plasmids for arabinose-induced expression of the indicated Myc-6xHis-tagged V12_14465 variants and the Flag-tagged V12_14460 (the immunity protein) required to allow the toxic V12_14465 variants inside the cells to accumulate) were diluted 100-fold into 100 mL 2xYT media supplemented with kanamycin and chloramphenicol and incubated at 37 °C with agitation (215 rpm). Protein expression was induced by adding 0.1% (wt/vol) L-arabinose when cultures reached an OD₆₀₀ of ~1.0, followed by incubation at 30 °C for 4 h with agitation (215 rpm). Cells were harvested by centrifugation and then resuspended in 5 mL lysis buffer A (20 mM Tris-Cl pH 7.5, 500 mM NaCl, 5% vol/vol glycerol, 10 mM imidazole, 1 mM PMSF, and 6 M urea). Urea was used to denature the proteins so that

V12_14460 (the immunity protein; see above) can be released from V12_14465. Cells were disrupted using a high-pressure cell disruptor (Constant system One Shot cell disruptor, model code: MC/AA). Cell debris was removed by centrifugation for 20 min at 13,300 × *g* at 4 °C. Supernatant fractions of lysed cells containing the denatured His-tagged V12_14465 variants were mixed with 100 µL Ni-Sepharose resin (50% slurry; GE healthcare) and incubated for 1 h at 4 °C with constant shaking. The suspensions (cell lysate with Ni-Sepharose resin) were loaded onto a column. Immobilized resin was washed with 10 mL wash buffer A (20 mM Tris-Cl pH 7.5, 500 mM NaCl, 5% vol/vol glycerol, 40 mM imidazole, and 6 M urea). Bound proteins were eluted from the column using 1 mL of elution buffer A (20 mM Tris-Cl pH 7.5, 500 mM NaCl, 5% vol/vol glycerol, 500 mM imidazole, and 6 M urea). Buffers were kept at 4 °C. The presence and purity of the eluted proteins were confirmed by SDS-PAGE.

To refold the denatured, purified proteins before *in vitro* DNase activity assays, a step-wise refolding procedure was applied. First, the eluted proteins were diluted twofold in ice-cold refolding buffer (20 mM Tris-Cl pH 7.5, 500 mM NaCl, 5% vol/vol glycerol) and incubated for 10 min on ice. This step was repeated, and then the 4 mL suspension was concentrated to 300 µL using a Spin-X UF concentrator column (Corning; 30 kDa or 5 kDa columns were used for full-length or truncated versions of V12_14465, respectively). The concentrated protein suspension was diluted in 3 mL of ice-cold DNase assay buffer (10 mM Tris-Cl pH 7.5, and 50 mM NaCl). The concentration and dilution in DNase assay buffer steps were repeated twice, resulting in 300 µL of concentrated protein in DNase assay buffer. The purified proteins were quantified by the Bradford method using 5X Bradford reagent (Bio-rad).

For *in vitro* DNase activity, *E. coli* BL21(DE3) isolated genomic DNA (100 ng) was incubated with 1 µg of purified V12_14465 variants (as explained above) for 30 min at 37 °C in DNase assay buffer supplemented with either 1 mM EDTA or 2 mM MgCl₂. The reactions were stopped by adding 6.65 µg of Proteinase K and the samples were incubated at room temperature for 10 more minutes. Samples were analyzed on 1.0% agarose-gel electrophoresis. For positive and negative controls, 1 µg DNase I (Bio-world) and 1 µg BSA (Sigma) were used, respectively.

In vivo DNase assay. *Escherichia coli* BL21(DE3) containing the indicated pBAD/Myc-His plasmids for expression of BC3021 variants were grown overnight in 2xYT media supplemented with kanamycin and 0.2% (wt/vol) glucose. Overnight cultures were washed with 2xYT and normalized to an OD₆₀₀ of 1.0 in 3 mL of fresh 2xYT supplemented with kanamycin and 0.1% (wt/vol) L-arabinose (to induce protein expression). Cultures were grown for 2 more hours with agitation (215 rpm) at 37 °C before 1.0 OD₆₀₀ units were pelleted. Genomic DNA was isolated from each sample using the EZ spin column genomic DNA kit (Bio lase) and eluted with 30 µL of sterile distilled water (Milli-Q). Equal genomic DNA elution volumes were analyzed by 0.7% (wt/vol) agarose gel electrophoresis. The experiments were repeated three times with similar results. Results from a representative experiment are shown.

In vivo RNase assay. *E. coli* BL21(DE3) containing the indicated pBAD/Myc-His plasmids for expression of BC3021 and V12_14465 variants were grown and induced as described above for the *in vivo* DNase assay, except that the growth media were supplemented with 0.4% (wt/vol) glucose, and that arabinose induction time was 1 h. Total RNA was purified from 0.5 OD₆₀₀ units of cells using the EZgene™ systems bacterial RNA kit (BIOMIGA) following the manufacturer's instructions. RNA was eluted with 100 µL DEPC-treated water, and 50 µL was further treated with 1 U of DNase I for 30 min. Twenty microliter samples were analyzed by 1% (wt/vol) agarose gel electrophoresis. The experiment was repeated three times with similar results. Results from a representative experiment are shown.

Fluorescence detection of DNase activity during competition. For assessment of prey DNA morphology during bacterial competition, the indicated *V. parahaemolyticus* 12-297/B strains were grown overnight and mixed as described above for the bacterial competition assay. The *Δv12_14465-0* prey harbored a pGFP plasmid (a high copy number plasmid for constitutive expression of GFP⁵⁷) to allow for their distinction from attacker cells. Attacker:prey mixtures were spotted on MLB agar plates and incubated at 30 °C for 3 h. Bacteria were then scraped from the plates, washed with M9 media, and incubated at room temperature for 10 min in M9 media containing the Hoechst 33342 (Invitrogen) DNA dye at a final concentration of 1 µg/µL. The cells were then washed and resuspended in 10 µL of M9 media. One microliter of bacteria suspensions was spotted onto M9 agar (1.5% wt/vol) pads and allowed to dry for 2 min before the pads were placed face-down in sterile 35 mm glass bottom Cellview cell culture dishes. Bacteria were imaged in a Nikon Eclipse Ti-E inverted motorized microscope equipped with a CFI PLAN apochromat DM 100X oil lambda PH-3 (NA, 1.45) objective lens, Lumencor SOLA SE II 395 light source, and ET-DAPI (#49028, used to visualize Hoechst signal) and ET-EGFP (#49002, used to visualize GFP signal) filter sets. Images were acquired using a DS-QI2 Mono cooled digital microscope camera (16 MP), and were postprocessed using Fiji ImageJ suite⁵⁸. Cells exhibiting the indicated DNA morphologies were manually counted (100 GFP-expressing prey cells per treatment in each experiment). The experiment was repeated three times.

Pull-down assay. Overnight cultures of *E. coli* BL21(DE3) containing plasmids for arabinose-induced expression of the indicated Myc-6xHis-tagged V12_14465 variants and the Flag-tagged V12_14460 (the immunity protein required to allow accumulation of the toxic V12_14465 variants inside the cells) were diluted 100-fold into 100 mL 2xYT media supplemented with kanamycin and chloramphenicol and incubated at 37 °C with agitation (215 rpm). Protein expression was induced by adding 0.1% (wt/vol) L-arabinose when cultures reached an OD₆₀₀ of ~1.0, followed by incubation at 30 °C for 4 h with agitation (215 rpm). Cells were harvested by centrifugation and then resuspended in 5 mL lysis buffer B (20 mM Tris-Cl pH 7.5, 500 mM NaCl, 5% vol/vol glycerol, 10 mM imidazole, and 1 mM PMSF). Cells were disrupted using a high-pressure cell disruptor (Constant system One Shot cell disruptor, model code: MC/AA). Cell debris was removed by centrifugation for 20 min at 13,300 × g at 4 °C. Next, 50 µL of supernatant fractions of lysed cells was taken and mixed with 50 µL of 2× protein sample buffer, boiled at 95 °C for 5 min, and kept for subsequent input protein analysis. Then 200 µL of supernatant fractions of lysed cells (400 µL was used for V12_14465^{1–282} due to its lower expression level compared with other V12_14465 variants) was mixed with 30 µL Ni-Sepharose resin (50% slurry; GE healthcare) and incubated for 1 h at 4 °C with constant shaking. Resin was collected by centrifugation at 2000 × g at 4 °C, and washed with 2.5 mL wash buffer B (20 mM Tris-Cl pH 7.5, 500 mM NaCl, 5% vol/vol glycerol, and 40 mM imidazole). Bound proteins were eluted from the resin using 200 µL of elution buffer B (20 mM Tris-Cl pH 7.5, 500 mM NaCl, 5% vol/vol glycerol, and 500 mM imidazole). Buffers were kept at 4 °C.

The eluted fractions (output) were mixed with 200 µL 2× protein sample buffer and boiled at 95 °C for 5 min. Samples were resolved on TGX stain-free gels (Bio-Rad) and transferred onto PVDF membranes. For V12_14465^{1–282}, we loaded twice the volume of input and output samples to equalize the levels of V12_14465 variants. The presence of V12_14465 variants was detected using c-Myc (9E10) antibodies (sc-40, Santa Cruz Biotechnology), and the presence of V12_14460 was detected using Flag antibodies (DYKDDDDK Tag (D6W5B) Rabbit mAb #14793; Cell Signaling Technology). Protein signals were visualized by ECL. The experiment was repeated at least three times with similar results, and the results from a representative experiment are shown.

BACTH assay. The adenylate cyclase-based BACTH Gateway system⁵² was used. Gateway-compatible plasmids producing the indicated proteins fused to either the T18 or T25 catalytic domains of *Bordetella* adenylate cyclase were transformed into the *E. coli* BTH101 reporter strain. Three independent colonies of each transformant were grown overnight in 2xYT media supplemented with appropriate antibiotics and IPTG (0.5 mM) at 30 °C. The following day, 5 µL of each overnight culture was spotted onto LB agar plates supplemented with ampicillin, spectinomycin, bromo-chloro-indolyl-galactopyrannoside (X-gal, 40 mg/mL), and IPTG (0.5 mM). Plates were incubated for 24 h at 30 °C. The experiment was repeated three times with similar results. Results from a representative experiment are shown.

TBLASTN search for auxiliary T6SS modules. TBLASTN was performed against the nucleotide sequence of *V. parahaemolyticus* 12-297/B using the protein sequences of *V. parahaemolyticus* RIMD 2210633-secreted components as queries: NP_797772.1 (Hcp1), NP_800537.1 (Hcp2), NP_797773.1 (VgrG1), NP_800536.1 (VgrG2), NP_797794.1 (amino acids 37–143, PAAR1) and NP_800535.1 (PAAR2). The expect value threshold was set to 10⁻⁶.

Identification of isolates containing the VgrG1b module. The nucleotide sequences of *V. parahaemolyticus* isolates were downloaded from NCBI on December 28, 2018, and TBLASTN was performed using the protein sequences of VgrG1b, WP_029857615.1 (V12_14465), and WP_024703575.1 (V12_14460) as queries. The expect value threshold was set to 10⁻⁶. The results were merged and sorted by their nucleotide accession and alignment start position, then manually inspected to identify the VgrG1b modules.

Identification of FIX- and PoNe-containing proteins. Position-Specific Scoring Matrices of FIX and PoNe were constructed using amino acids 1–282 and 283–449 of V12_14465, respectively. Five iterations of PSI-BLAST⁵⁹ were performed against the reference protein database (a maximum of 500 hits with a threshold of 10⁻⁶ were used in each iteration). Bacterial genomes containing FIX and PoNe were identified and their protein sequences and feature tables were downloaded from NCBI on November 27–29, 2018.

Reverse position-specific BLAST⁵⁹ was used to identify the FIX- and PoNe-containing proteins in the bacterial genomes. The results were filtered using an expect value threshold of 10⁻⁹. With PoNe, the protein length had to be at least 80 aa. Unique accessions located at the ends of the contigs were removed. Unique accessions appearing in the same organism in more than one contig were removed if the same downstream gene existed at the same distance.

The proteins were clustered using CD-HIT v4.17⁶⁰. The clustering threshold was set to 40% identity with a minimal 50% length difference (i.e., the shorter protein must be at least 50% of the longer protein), and the sequences were clustered to the most similar cluster that met the threshold. FIX-containing proteins were clustered according to their C-terminal sequences (i.e., sequences

located downstream of FIX), and PoNe-containing proteins were clustered according to their N-terminal sequences (i.e., sequences located upstream of PoNe).

The proteins located upstream of the FIX- and PoNe-containing proteins on the same strand were defined as upstream proteins, whereas proteins located downstream on the same strand were defined as downstream proteins. Protein sequences were analyzed using the NCBI Batch Conserved Domains-Search Tool⁶¹. Transmembrane topology and signal peptides were predicted using the Phobius server⁶². Signal peptide cleavage sites were predicted using the SignalP 4.1 server⁶³. Proteins were considered as containing a signal peptide only if they were designated as such by both Phobius and SignalP 4.1 predictions.

Construction of the phylogenetic tree. The DNA sequences of *rpoB* coding for DNA-directed RNA polymerase subunit beta were downloaded from NCBI on December 8, 2018 for bacterial strains encoding PoNe. Partial sequences were removed. The sequences were aligned using MAFFT v7.408 FFT-NS-2^{64,65}. The evolutionary history was inferred using the neighbor-joining method⁶⁶ with the Jukes–Cantor substitution model (JC69). The analysis involved 1303 nucleotide sequences and 2535 conserved sites. Evolutionary analyses were conducted using the MAFFT server (<https://mafft.cbrc.jp/alignment/server/>), and the tree was visualized using MEGA 7⁶⁷.

Illustration of conserved residues using WebLogo 3. FIX- and PoNe-sequences were aligned using MUSCLE⁶⁸, and aligned columns not represented in WP_029857615.1 were removed. Conserved residues were illustrated using the WebLogo 3 server (<http://weblogo.threeplusone.com/>). Amino acid numbering was based on the sequence of WP_029857615.1.

Identification of FIX-containing genomes encoding T6SS. RPS-BLAST⁵⁹ was employed to identify the T6SS core components in the FIX-containing bacterial genomes. The proteins were aligned against 11 COGs (COG3516, COG3517, COG3157, COG3521, COG3522, COG3455, COG3523, COG3518, COG3519, COG3520, and COG3515) shown to specifically predict T6SS and VgrG (COG3501), which was correlated with T6SS-associated organisms³. Bacterial genomes encoding at least nine out of the eleven T6SS core components were regarded as harboring T6SS.

VgrG sequence alignment. To perform the VgrG sequence alignment, protein sequences of *V. parahaemolyticus* 12-297/B VgrG1 (B5C30_RS15295) and VgrG1b (translated from the nucleotide sequence found between positions 22692 and 20311 in reference sequence NZ_MYFG01000470.1) were aligned using Clustal Omega⁷⁰, and the alignment was visualized using JalView⁷¹.

Secondary-structure prediction. Secondary structure prediction of the PoNe conserved region of V12_14465 was performed using the Jpred 4 server⁷².

Statistical analysis. Data were statistically analyzed with GraphPad Prism 6, using the unpaired, two-tailed Student *t* test. Differences were considered significant at *P* < 0.05; n.s. denotes not significant.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The experimental and computational data that support the findings of this research are available in this article and its Supplementary Information files, or upon request from the corresponding authors. The source data underlying Figs. 1c–f, 2c–d, 3b, 4a, f, and Supplementary Figs. 2a, b, 3, 4, 5b, c, 7, and 9c are provided as a Source data file.

Received: 15 January 2019 Accepted: 16 July 2019

Published online: 09 August 2019

References

- Pukatzki, S. et al. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the Dictyostelium host model system. *Proc. Natl. Acad. Sci. USA* **103**, 1528–1533 (2006).
- Mougous, J. D. et al. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* **312**, 1526–1530 (2006).
- 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 Genom.* **10**, 104 (2009).

4. Russell, A. B. et al. Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475**, 343–349 (2011).
5. Pukatzki, S., Ma, A. T., Revel, A. T., Sturtevant, D. & Mekalanos, J. J. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl Acad. Sci. USA* **104**, 15508–15513 (2007).
6. Nguyen, V. S. et al. Towards a complete structural deciphering of Type VI secretion system. *Curr. Opin. Struct. Biol.* **49**, 77–84 (2018).
7. Zoued, A. et al. Structure–function analysis of the TssL cytoplasmic domain reveals a new interaction between the type VI secretion baseplate and membrane complexes. *J. Mol. Biol.* **428**, 4413–4423 (2016).
8. Wang, J. et al. Cryo-EM structure of the extended type VI secretion system sheath-tube complex. *Nat. Microbiol.* **2**, 1507–1512 (2017).
9. Nazarov, S. et al. Cryo-EM reconstruction of Type VI secretion system baseplate and sheath distal end. *EMBO J.* e201797103. <https://doi.org/10.15252/embj.201797103> (2017).
10. Shneider, M. M. et al. PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature* **500**, 350–353 (2013).
11. Basler, M., Pilhofer, M., Henderson, G. P., Jensen, G. J. & Mekalanos, J. J. Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature* **483**, 182–186 (2012).
12. Durand, E., Cambillau, C., Cascales, E. & Journet, L. VgrG, Tae, Tle, and beyond: the versatile arsenal of Type VI secretion effectors. *Trends Microbiol.* **22**, 498–507 (2014).
13. Russell, A. B. et al. A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach. *Cell Host Microbe* **11**, 538–549 (2012).
14. Ma, L. S., Hachani, A., Lin, J. S., Filloux, A. & Lai, E. M. Agrobacterium tumefaciens deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. *Cell Host Microbe* **16**, 94–104 (2014).
15. Pissaridou, P. et al. The *Pseudomonas aeruginosa* T6SS-VgrG1b spike is topped by a PAAR protein eliciting DNA damage to bacterial competitors. *Proc. Natl Acad. Sci. USA* **115**, 12519–12524 (2018).
16. Russell, A. B. et al. Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature* **496**, 508–512 (2013).
17. Miyata, S. T., Unterwieser, D., Rudko, S. P. & Pukatzki, S. Dual expression profile of type VI secretion system immunity genes protects pandemic *Vibrio cholerae*. *PLoS Pathog.* **9**, 1–18 (2013).
18. Jiang, F., Waterfield, N. R., Yang, J., Yang, G. & Jin, Q. A *Pseudomonas aeruginosa* type VI secretion phospholipase D effector targets both prokaryotic and eukaryotic cells. *Cell Host Microbe* **15**, 600–610 (2014).
19. Tang, J. Y., Bullen, N. P., Ahmad, S. & Whitney, J. C. Diverse NADase effector families mediate interbacterial antagonism via the type VI secretion system. *J. Biol. Chem.* **293**, 1504–1514 (2018).
20. Whitney, J. C. et al. An interbacterial NAD(P)⁺ glycohydrolase toxin requires elongation factor Tu for delivery to target cells. *Cell* **163**, 607–619 (2015).
21. Ting, S.-Y. et al. Bifunctional immunity proteins protect bacteria against FtsZ-targeting ADP-ribosylating toxins. *Cell*. <https://doi.org/10.1016/j.CELL.2018.09.037> (2018).
22. Hachani, A., Wood, T. E. & Filloux, A. Type VI secretion and anti-host effectors. *Curr. Opin. Microbiol.* **29**, 81–93 (2016).
23. English, G. et al. New secreted toxins and immunity proteins encoded within the Type VI secretion system gene cluster of *Serratia marcescens*. *Mol. Microbiol.* **86**, 921–936 (2012).
24. Salomon, D. et al. Marker for type VI secretion system effectors. *Proc. Natl Acad. Sci. USA* **111**, 9271–9276 (2014).
25. Dar, Y., Salomon, D. & Bosis, E. The antibacterial and anti-eukaryotic type VI secretion system MIX-effector repertoire in vibronaceae. *Mar. Drugs* **16**, 433 (2018).
26. Ma, J. et al. 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 (2017).
27. Koskiniemi, S. et al. Rhs proteins from diverse bacteria mediate intercellular competition. *Proc. Natl. Acad. Sci. USA* **110**, 7032–7037 (2013).
28. Lovell, C. R. Ecological fitness and virulence features of *Vibrio parahaemolyticus* in estuarine environments. *Appl. Microbiol. Biotechnol.* **101**, 1781–1794 (2017).
29. Yu, Y. et al. Putative type VI secretion systems of *Vibrio parahaemolyticus* contribute to adhesion to cultured cell monolayers. *Arch. Microbiol.* **194**, 827–835 (2012).
30. Salomon, D., Gonzalez, H., Updegraff, B. L. & Orth, K. *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 (2013).
31. Li, P. et al. 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 (2017).
32. Salomon, D. et al. Type VI secretion system toxins horizontally shared between marine bacteria. *PLoS Pathog.* **11**, 1–20 (2015).
33. Speare, L. et al. Bacterial symbionts use a type VI secretion system to eliminate competitors in their natural host. *Proc. Natl Acad. Sci. USA* **115**, E8528–E8537 (2018).
34. Unterwieser, D. et al. The *Vibrio cholerae* type VI secretion system employs diverse effector modules for intraspecific competition. *Nat. Commun.* **5**, 3549 (2014).
35. Verster, A. J. et al. The landscape of type VI secretion across human gut microbiomes reveals its role in community composition. *Cell Host Microbe* **22**, 411–419.e4 (2017).
36. Rigard, M. et al. Francisella tularensis IglG belongs to a novel family of PAAR-like T6SS proteins and harbors a unique N-terminal extension required for virulence. *PLoS Pathog.* **12**, 1–33 (2016).
37. Whitney, J. C. et al. A broadly distributed toxin family mediates contact-dependent antagonism between gram-positive bacteria. *Elife* **6**, e26938 (2017).
38. Poulsen, C., Panjikar, S., Holton, S. J., Wilmanns, M. & Song, Y.-H. WXG100 protein superfamily consists of three subfamilies and exhibits an α -helical C-terminal conserved residue pattern. *PLoS ONE* **9**, e89313 (2014).
39. Aoki, S. K. et al. A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* **468**, 439–442 (2010).
40. Hayes, C. S., Koskiniemi, S., Ruhe, Z. C., Poole, S. J. & Low, D. A. Mechanisms and biological roles of contact-dependent growth inhibition systems. *Cold Spring Harb. Perspect. Med.* **4**, a010025 (2014).
41. Steczkiewicz, K., Muszewska, A., Knizewski, L., Rychlewski, L. & Ginalski, K. Sequence, structure and functional diversity of PD-(D/E)XK phosphodiesterase superfamily. *Nucleic Acids Res.* **40**, 7016–7045 (2012).
42. Knizewski, L., Kinch, L. N., Grishin, N. V., Rychlewski, L. & Ginalski, K. Realm of PD-(D/E)XK nuclease superfamily revisited: detection of novel families with modified transitive meta profile searches. *BMC Struct. Biol.* **7**, 40 (2007).
43. Feder, M. & Bujnicki, J. M. Identification of a new family of putative PD-(D/E)XK nucleases with unusual phylogenomic distribution and a new type of the active site. *BMC Genom.* **6**, 21 (2005).
44. Liang, X. et al. Identification of divergent type VI secretion effectors using a conserved chaperone domain. *Proc. Natl Acad. Sci. USA* **112**, 9106–9111 (2015).
45. Unterwieser, D. et al. Chimeric adaptor proteins translocate diverse type VI secretion system effectors in *Vibrio cholerae*. *EMBO J.* **34**, 2198–2210 (2015).
46. Zhang, D., de Souza, R. F., Anantharaman, V., Iyer, L. M. & Aravind, L. Polymorphic toxin systems: comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol. Direct* **7**, 18 (2012).
47. Zhang, D., Iyer, L. M. & Aravind, L. A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems. *Nucleic Acids Res.* **39**, 4532–4552 (2011).
48. Kirchberger, P. C., Unterwieser, 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.* **7**, 1–12 (2017).
49. Bingle, L. E., Bailey, C. M. & Pallen, M. J. Type VI secretion: a beginner's guide. *Curr. Opin. Microbiol.* **11**, 3–8 (2008).
50. Park, K.-S. et al. Functional characterization of two type III secretion systems of *Vibrio parahaemolyticus*. *Infect. Immun.* **72**, 6659–6665 (2004).
51. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
52. Ouellette, S. P., Gauliard, E., Antosova, Z. & Ladant, D. A Gateway[®]-compatible bacterial adenylate cyclase-based two-hybrid system. *Environ. Microbiol. Rep.* **6**, 259–267 (2014).
53. O'Toole, R., Milton, D. L. & Wolf-Watz, H. Chemotactic motility is required for invasion of the host by the fish pathogen *Vibrio anguillarum*. *Mol. Microbiol.* **19**, 625–637 (1996).
54. Gode-Potratz, C. J., Kustusch, R. J., Breheny, P. J., Weiss, D. S. & McCarter, L. L. Surface sensing in *Vibrio parahaemolyticus* triggers a programme of gene expression that promotes colonization and virulence. *Mol. Microbiol.* **79**, 240–263 (2011).
55. Bensadoun, A. & Weinstein, D. Assay of proteins in the presence of interfering materials. *Anal. Biochem.* **70**, 241–250 (1976).
56. Kalderon, Z., Kumar, S. & Engelberg-Kulka, H. The SOS response is permitted in *Escherichia coli* strains deficient in the expression of the mazEF pathway. *PLoS ONE* **9**, e114380 (2014).
57. Ritchie, J. M. et al. Inflammation and disintegration of intestinal Villi in an experimental model for *Vibrio parahaemolyticus*-induced diarrhea. *PLoS Pathog.* **8**, e1002593 (2012).
58. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
59. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).

60. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
61. Marchler-Bauer, A. & Bryant, S. H. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* **32**, W327–W331 (2004).
62. Käll, L., Krogh, A. & Sonnhammer, E. L. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **338**, 1027–1036 (2004).
63. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
64. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* <https://doi.org/10.1093/bib/bbx108> (2017).
65. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
66. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
67. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
68. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
69. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
70. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
71. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
72. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* **43**, W389–W394 (2015).

Acknowledgements

This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant agreement no. 714224) (DS). DS is an Alon Fellow. We thank Hanna Engelberg-Kulka and Scot Ouellette for their generous gift of plasmids, Avigdor Eldar for providing bacterial strains, Kinga Keppel and Chen Yacobsohn for excellent technical assistance, and Udi

Qimron for his critical reading of the paper. We also thank members of the Salomon lab for fruitful discussions.

Author contributions

B.J., E.B. and D.S. designed the study. B.J., C.M.F. and D.S. performed the experiments. E.B. performed the computational analyses. B.J., C.M.F., E.B. and D.S. analyzed the data. D.S. wrote the paper, and the other authors contributed to the writing. All authors read and approved the final version of the paper.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-019-11546-6>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Peer review information: *Nature Communications* thanks John Whitney and other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019